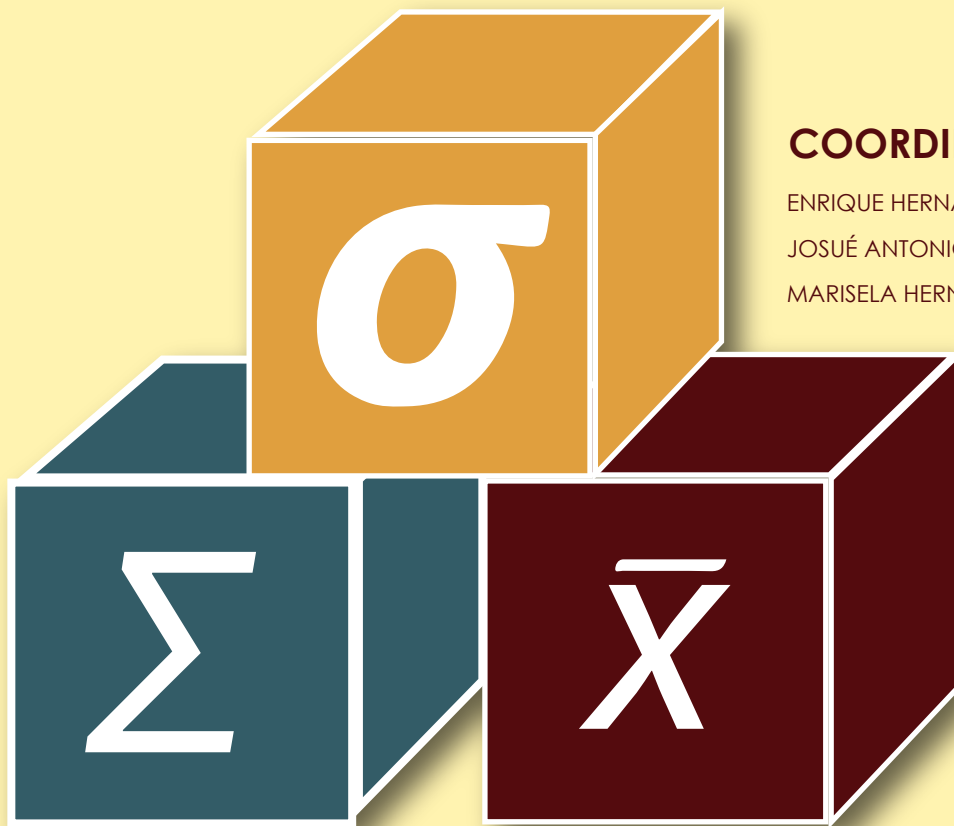


PRINCIPIOS DE

ESTADÍSTICA

PARA EL

EDUCADOR ESPECIAL



COORDINADORES

ENRIQUE HERNÁNDEZ ARTEAGA

JOSUÉ ANTONIO CAMACHO CANDIA

MARISELA HERNÁNDEZ GONZÁLEZ



PRINCIPIOS DE
ESTADÍSTICA
PARA EL
EDUCADOR ESPECIAL



Principios de estadística para el educador especial

Enrique Hernández Arteaga

Facultad de Ciencias para el Desarrollo Humano
Universidad Autónoma de Tlaxcala

Josué Antonio Camacho Candia

Facultad de Ciencias para el Desarrollo Humano
Universidad Autónoma de Tlaxcala

Marisela Hernández González

Instituto de Neurociencias
CUCBA, Universidad de Guadalajara

Coordinadores

Universidad Autónoma de Tlaxcala

Universidad de Guadalajara

2023



Universidad Autónoma de Tlaxcala

Serafín Ortiz Ortiz
Rectoría

Margarita Martínez Gómez
Secretaría Académica

Elvia Hernández Escalona
Secretaría Administrativa

Diana Selene Ávila Casco
Secretaría de Extensión Universitaria y Difusión Cultural

Alfredo Adán Pimentel
Secretaría de Investigación Científica y Posgrado

Roberto Carlos Cruz Becerril
Secretaría Técnica

Juan George Zecua
Secretaría de Autorrealización

Irving Eduardo Ortiz Gallardo
Coordinador de División de Ciencias y Humanidades

Josué Antonio Camacho Candia
Director de la Facultad de Ciencias para el Desarrollo Humano



Universidad de Guadalajara

Ricardo Villanueva Lomelí
Rectoría General

Héctor Raúl Solís Gadea
Vicerrectoría Ejecutiva

Guillermo Arturo Gómez Mata
Secretaría General

Graciela Gudiño Cabrera
Rectoría del Centro Universitario de Ciencias Biológicas y Agropecuarias

Carlos Bea Zárate
Dirección del Instituto de Neurociencias

Sayri Karp Mitastein
Dirección de la Editorial

Primera edición, 2023

Coordinación

Enrique Hernández Arteaga
Josué Antonio Camacho Candia
Marisela Hernández González

Textos

©Abril Zagnitte Gómez Méndez
Anders Agmo
Andrea Saldívar Reyes
Carolina Sotelo Tapia
Cid González González
Citlali López Esparza
Claudia del Carmen Amezcua Gutiérrez
Delia de Jesús Domínguez Morales
Edgar Oswaldo Zamora González
Enrique Hernández Arteaga
Fabiola Alejandra Iribe Burgos
Francisco Javier Aguilar Guevara
Jahaziel Molina del Río
José Carlos Hevia Orozco
Josué Antonio Camacho Candia
Juan Pablo García Hernández
Luz Berenice López Hernández
Manuel Alejandro Cruz Aguilar
Marai Pérez Hernández
María del Consuelo Pedraza Aguilar
María Isabel Meza Solís
Marisela Hernández González
Miguel Ángel Guevara Pérez
Myriam Nayeli Villafuerte Vega
Pedro Manuel Cortes Esparza
Rosa María Hidalgo Aguirre
Susana Angélica Chavira Castro



D.R. © 2023, Universidad Autónoma de Tlaxcala
Av. Universidad Núm. 1
Colonia Centro
90000 Tlaxcala, Tlaxcala.

ISBN: 978-607-545-093-3



D.R. © 2023, Universidad de Guadalajara
Editorial Universidad de Guadalajara
José Bonifacio Andrada 2679
Colonia Lomas de Guevara
44657 Guadalajara, Jalisco

ISBN: 978-607-581-036-2

Hecho en México
Made in Mexico

Los coordinadores de la obra “Principios de estadística para el educador especial” expresan su agradecimiento a todas(os) los autores y co-autores invitados, así como al Comité de Revisión y Dictaminación Técnica y Científica conformado por invitación para la presente obra.

ÍNDICE

Introducción	7
PARTE I. RELEVANCIA DEL ESTUDIO DE LA ESTADÍSTICA PARA EL EDUCADOR ESPECIAL	13
1. Métodos de investigación en educación especial	14
<i>Delia de Jesús Domínguez Morales & Andrea Saldívar Reyes</i>	
2. ¿Para qué cuantificar en educación especial?	51
<i>Josué Antonio Camacho Candia, Myriam Nayeli Villafuerte Vega, & María Isabel Solís Meza</i>	
3. El muestreo en investigación del área de educación especial	73
<i>Francisco Javier Aguilar Guevara</i>	
4. Introducción a la estadística	104
<i>Susana Angelica Castro Chavira & M. del Consuelo Pedraza Aguilar</i>	
PARTE II. ESTADÍSTICA DESCRIPTIVA	129
1. Implicación de las medidas de tendencia central y de dispersión en la educación especial	130
<i>Carolina Sotelo Tapia, Marisela Hernández González, Fabiola Alejandra Iribe Burgos, & Abril Zagnitte Gómez Méndez</i>	
2. Procedimientos para el cálculo de las medidas de posición	156
<i>Marai Pérez Hernández, Edgar Oswaldo Zamora González, & Luz Berenice López Hernández</i>	

3. La distribución de frecuencias y su representación gráfica en la educación especial	189
<i>Claudia Amezcua Gutiérrez, Rosa María Hidalgo Aguirre, Jahaziel Molina del Río</i>	
PARTE III. ESTADÍSTICA INFERENCIAL	232
1. El uso de la correlación	233
<i>Jorge Carlos Hevia Orozco & Cid Ramón González González</i>	
2. Uso de pruebas no paramétricas de significación estadística en educación especial	261
<i>Pedro Manuel Cortes Esparza, Citlali López Esparza, Juan Pablo García Hernández, & Miguel Ángel Guevara</i>	
3. Uso de pruebas paramétricas de significación estadística en educación especial	304
<i>Manuel Alejandro Cruz Aguilar, Enrique Hernández Arteaga, & Anders Ågmo</i>	
Autores(as) y coautores(as) de la obra	349
Comité de Revisión y Dictaminación Técnica y Científica de la Obra	356

INTRODUCCIÓN

“Principios de Estadística para el Educador Especial” surge como parte de un esfuerzo que se ha estado edificando desde hace algunos años entre docentes e investigadores de distintas universidades a nivel nacional e internacional. La obra particularmente nació como consecuencia de los trabajos del cuerpo académico en consolidación “Educación especial y procesos de aprendizaje” (UATLX-CA-213) y del doctorado en Ciencias Aplicadas a la Educación Especial, ambos de la Universidades Autónoma de Tlaxcala; así como de diversos grupos de investigación de la Universidad de Guadalajara. A partir de efectuar un análisis teórico-conceptual hacia algunas de las aplicaciones y posibles usos de la estadística por la educación especial, el propósito principal de esta compilación de capítulos fue enfatizar la importancia de la estadística para el desarrollo científico e interpretación de resultados obtenidos en la investigación en educación especial. Así, con la finalidad de obtener conclusiones mediante el análisis matemático que le permita al educador especial resolver casos. Si bien desde la perspectiva metodológica-matemática ya existen libros que han explicado el procesamiento matemático-estadístico de los datos, la relación directa entre la estadística y la educación especial a la fecha no ha quedado bien establecida.

Por lo anteriormente mencionado, se espera que esta incursión en ambos campos del conocimiento, generen un mayor impacto al ponerlos al alcance del campo profesional de la educación especial, la cual se ha visto en la necesidad de mejorar las bases científicas de la atención que brinda a la población, especialmente en Tlaxcala, ciudad que ha sido un semillero de educadores especiales desde 1979.

Por lo tanto, la presente no es una obra aislada, sino continuidad de otras que han sido elaboradas por diversos investigadores y docentes interesados en fortalecer

la educación especial desde sus propias disciplinas y campos profesionales, los cuales provienen de diversas universidades, como la Universidad Autónoma de Tlaxcala, la Universidad Veracruzana, la Universidad Nacional Autónoma de México, la Universidad de Guadalajara, la Universidad Autónoma de Yucatán, la Universidad de Córdoba y la Universidad de Almería principalmente. Desde luego, además de las obras mencionadas, se han publicado artículos y se han organizado eventos en Tlaxcala que han servido como espacios de reflexión y de vinculación con los docentes, estudiantes y profesionales dedicados a la educación especial. Uno de los eventos representativos, es el Coloquio Internacional de “Ciencias Aplicadas a la Educación y la Diversidad”, celebrado en mayo de forma anual desde 2019, en la Facultad de Ciencias para el Desarrollo Humano de la Universidad Autónoma de Tlaxcala.

La presente obra se compone de tres partes, las cuales contienen en total diez capítulos, con aportaciones de docentes de instituciones como la Universidad Nacional Autónoma de México, la Universidad de Guadalajara, la Universidad de Trømsø, el Instituto Nacional de Psiquiatría “Ramón de la Fuente Muniz”, la Universidad Anahuac Mayab, el Tecnológico Nacional de México y por supuesto, la Universidad Autónoma de Tlaxcala.

La primera parte, titulada *Relevancia del estudio de la estadística para el educador especial*, se compone de cuatro capítulos, donde los autores nos exponen de manera general cuál es la posición actual de la educación especial como una multidisciplina, así como el papel del educador especial y el posible uso del análisis matemático en estudios con relación a usuarios de la educación especial.

En el capítulo 1, Delia Domínguez y Andrea Saldívar se refieren a los tres enfoques que existen de la investigación: cualitativo, cuantitativo y mixto. Es así como ellas plantean algunos ejemplos de investigaciones desde el enfoque

cuantitativo aplicadas al campo de la educación, con la intención de que los estudiantes, docentes y educadores especiales en servicio encuentren en el enfoque cuantitativo una herramienta más para emprender proyectos de investigación educativa a partir de este enfoque de la investigación.

En el capítulo 2, Josué Camacho, Myriam Villafuerte e Isabel Solís discuten sobre la postura que señala que la educación especial debe ser vista como una disciplina científica multidimensional y multiparadigmática, en la que convergen diversas profesiones y disciplinas, por lo tanto, son varios los conocimientos y herramientas que deben usarse para mejorar la práctica educativa en la educación especial, como lo podría ser la estadística, la cual puede considerada como una herramienta fundamental para organizar e interpretar la información de evaluaciones objetivas de personas con discapacidad, con alteraciones del neurodesarrollo o con aptitudes sobresalientes.

En el capítulo 3, Francisco Aguilar reconoce a la Educación Especial como una interdisciplina que basa su conocimiento en teorías desarrolladas en diversas disciplinas científicas. Por lo tanto, es menester del educador especial considerar aspectos teóricos relacionados con el muestreo, entendido éste como el proceso de incorporación de participantes a estudios, abordando aspectos asociados tanto a estudios cuantitativos como cualitativos.

En el capítulo 4, Susana Castro y Consuelo Pedraza establecen que la estadística es una herramienta que permite organizar los datos obtenidos en el aula para su interpretación, y, por ende, para su uso en la toma de decisiones en el ámbito de la educación especial. Ellas describen el modelo estadístico enfocado en la solución de problemas PPDAC (Problema – Plan – Datos – Análisis – Conclusión y Comunicación), discutiendo su posible aplicabilidad en la exploración de los factores que intervienen en el proceso de enseñanza-aprendizaje de niños usuarios

de la educación especial. Resaltan también, la importancia de interpretar los resultados numéricos para comunicarlos mediante un flujo narrativo, de tal manera que sea útil para educadores, directores, e instituciones.

La segunda parte, titulada *Estadística descriptiva*, se compone de tres capítulos, donde los autores nos exponen de manera general los procedimientos básicos empleados en la estadística descriptiva, haciendo particular énfasis, en cómo podrían emplearse en el contexto de la investigación en ciencias aplicadas a la educación especial, así como su posible interpretación para que el educador especial que se encuentra en servicio, pueda interpretar la información que consulta en diversos artículos científicos al momento de buscar las bases para los métodos de intervención-atención educativa que emplea en la práctica.

En el capítulo 1, Carolina Sotelo, Marisela Hernández, Alejandra Iribe, y Abril Gómez describen el uso de las medidas de tendencia central (moda, mediana y media), así como las medidas de dispersión (varianza y desviación estándar), mencionando los cálculos para su cálculo, así como la manera de interpretarlos. Ellas resaltan el impacto que tiene el uso de la estadística descriptiva en los reportes científicos en investigaciones del área de la educación especial.

En el capítulo 2, Marai Pérez, Edgar Zamora, y Luz López abordan las llamadas medidas de posición, las cuales son herramientas de la estadística descriptiva que sirven para segmentar la distribución de datos, de tal manera que se puedan analizar conociendo la posición de un elemento dentro de un grupo de datos. Si bien, ellos mencionan los procedimientos para obtener dichas medidas, también ejemplifican el uso de éstas, recayendo en aspectos que le pudieran interesar al educador especial.

En el capítulo 3, Claudia Amezcua, Rosa Hidalgo y Jahaziel Molina establecen la importancia de contar con herramientas estadísticas, que faciliten la interpretación de datos, para que, mediante diferentes sistemas de representación gráfica, permitan que el educador especial evalúe objetivamente la información procedente de poblaciones o muestras de usuarios de la educación especial, esto con el fin de que pueda mejorar la intervención-atención educativa en la práctica profesional.

La tercera parte, titulada *Estadística inferencial*, se compone de tres capítulos, donde los autores nos exponen de manera general los procedimientos básicos empleados en la estadística inferencial, haciendo particular énfasis, en cómo podrían emplearse en el contexto de la investigación en ciencias aplicadas a la educación especial, así como la posible interpretación para que el educador especial que se encuentra en servicio, pueda utilizar la información acerca de la correlación o pruebas de significación estadística, que se presenta en diversos artículos científicos. Con la principal intención de que el educador especial sea capaz de basar su práctica en evidencia científica de métodos de intervención-atención educativa reportados en artículos de investigación.

En el capítulo 1, Jorge Hevia y Cid González nos mencionan que la correlación es una herramienta de la estadística inferencial que permite establecer el grado de asociación entre dos variables, por ejemplo, los procesos cognitivos. De tal manera que el educador especial podrá encontrar en estos análisis la asociación entre diversos tipos de intervención con el grado de avance del estudiante, obviamente, teniendo en cuenta que la correlación no indica una relación causal.

En el capítulo 2, Pedro Cortes, Citlali López, Juan García, y Miguel Guevara, plantean que es necesario recurrir a análisis de los datos subjetivos o cualitativos, donde la estadística puede ser una herramienta que permite hacer más objetivos

dichos análisis; partiendo de la cuantificación de escalas tipo Likert, permitirán desarrollar métodos de intervención basados en evidencia científica que sirvan para desarrollar de manera integral al alumno. Es entonces que surge la estadística inferencial no paramétrica, la cual provee de herramientas para el análisis estructurado de este tipo de datos. Ellos esperan que el educador especial sea capaz de refinar las propuestas de métodos de intervención y análisis de dichos resultados a partir del análisis objetivo de los casos.

Finalmente, en el capítulo 3, Manuel Cruz, Enrique Hernández y Anders Ågmo, discuten la posible aplicación de las pruebas de significación estadística paramétricas (t de Student y el Análisis de Varianza) en el contexto de la investigación en educación especial. En este capítulo, los autores pretenden que el educador especial sea capaz de identificar cuándo se habla de una diferencia significativa en un artículo de investigación, así como de la relevancia de dicho dato estadístico, para que pueda obtener sus propias conclusiones al respecto y emplear dicha información al momento de planear métodos de intervención-atención educativa basados en la evidencia científica para su práctica docente.

Esperamos que este libro sea un referente para los estudiantes, docentes y profesionales en servicio de la educación especial; que permita una reflexión conjunta sobre un problema o fenómeno común y sobre todo, que permita consolidar las bases científicas en el quehacer profesional educativo, especialmente en aquellos ámbitos que presentan retos a los procedimientos tradicionales, tal como ocurre en el caso de la educación especial.

Dr. Enrique Hernández Arteaga

Dr. Josué Antonio Camacho Candia

Dra. Marisela Hernández González

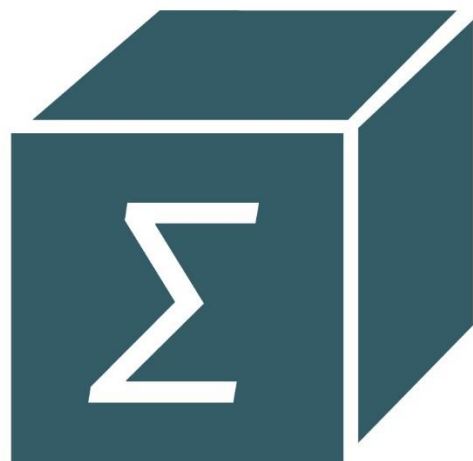
Parte I:

RELEVANCIA

DEL ESTUDIO DE LA

ESTADÍSTICA PARA EL

EDUCADOR ESPECIAL



MÉTODOS DE INVESTIGACIÓN EN EDUCACIÓN ESPECIAL

*Delia de Jesús Domínguez Morales
Andrea Saldívar Reyes ^Σ*

Resumen

Pensar que las cosas son dadas y que el sentido común es el elemento esencial para generar conjeturas sobre el mundo y sus realidades, lleva a realizar afirmaciones que pueden estar erradas. Poder establecer aproximaciones más cercanas y objetivas a la realidad, obliga al profesional a tener una mirada más crítica y reflexiva, para ello la investigación científica es una herramienta veraz y útil, pues tener un acercamiento desde diversas disciplinas y métodos conlleva a establecer criterios de rigurosidad y sistematicidad necesarios para comprender los fenómenos y hacer fundamentar suposiciones. En el presente capítulo se abordará de manera sucinta la investigación, las tres tradiciones metodológicas científicas, y se podrán encontrar ejemplos de investigaciones desde el enfoque cuantitativo aplicadas al campo de la educación especial, con la intención de que se pueda reconocer la importancia y vigencia de este enfoque. En espera que los educadores especiales en servicio y en formación, así como profesionales de áreas afines, encuentren en este capítulo una inspiración y una herramienta para emprender proyectos que tengan una envergadura para generar conocimientos nuevos.

Palabras clave: Investigación, cualitativo, cuantitativo, metodología.

^ΣFacultad de Ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala.
Carretera Tlaxcala-Puebla Km. 1.5, C.P. Tlaxcala, Tlax. asaldivarr_fcdh@uatx.mx

Introducción

Dentro de las muchas carreras universitarias a elegir para tener una formación profesional se encuentra la Educación Especial, la cual estudia a profundidad las necesidades educativas especiales de los alumnos, prácticas de inclusión, el desarrollo del individuo, las barreras para el aprendizaje y la participación, el currículo, políticas y programas educativos vigentes, por mencionar algunos, todo ello enmarcado en diversas multidisciplinas que se correlacionan como la psicología, pedagogía, neurociencias y demás. Tal vez en su elección de carrera profesional no se le ocurrió que podría necesitar de las matemáticas, o más específico aún, de la estadística y de métodos de investigación. Sin embargo, estos son pilares en toda ciencia y disciplina, y, por supuesto, la educación especial no es la excepción. Por ello, en las siguientes páginas se detalla la importancia de la investigación en la educación especial, los tipos de enfoques y los alcances de la investigación, el método de investigación, el escrito de la investigación científica, los distintos diseños de investigación que existen y ejemplos de estudios en educación especial y su relación con la estadística, con el fin de que el lector se acerque a conocer la importancia de los métodos de investigación en el campo de la Educación Especial.

La investigación en educación especial

La investigación, como actividad científica, permite el avance de la ciencia y el conocimiento en general; en educación especial ésta ha permitido renovar la educación en general y ha proporcionado un conocimiento de la realidad y de posibilidades educativas innovadoras, con el fin de mejorar la práctica educativa (Jurado de los Santos y Sanahuja Gavaldá, 1997).

Son varias las ventajas de la investigación en la educación especial, por ejemplo, contribuye a afianzar y reelaborar el conocimiento que tenemos sobre la enseñanza, permite fijar objetivos plausibles para los alumnos, proporciona un diagnóstico de la realidad para tomar decisiones y actuar con base en ello, así como mejorar la práctica de la enseñanza como su metodología, programas, materiales y la práctica del profesor.

El educador especial, al estar en continua práctica, puede investigar aquellos fenómenos que saltan a su vista experta en enseñanza-aprendizaje y que requieren de una explicación sistemática y verificable; de esta manera puede investigar con una metodología y marco teórico específico, y aplicar los conocimientos que se deriven.

Es decir, puede hacer ciencia; al respecto Villoro (2002) describe que la ciencia consiste en un conjunto de saberes compartibles por una comunidad epistémica determinada, de tal manera que las teorías y enunciados de observación comprobables intersubjetivamente se ponen en relación con un dominio de objetos.

La educación, y la educación especial se puede clasificar dentro de las ciencias empíricas, y como tal supone un conocimiento personal, de hechos observables verificados por la experiencia directa, de hechos determinados por un marco conceptual específico que aplica el científico, que responde a preguntas planteadas en ese marco y acorde a teorías vigentes (Villoro, 2002). El marco conceptual que orienta el conocimiento del científico es fundamental, ya que le sirve como una guía o brújula para buscar y apegarse a las características que le interesa buscar y excluir otras, de un fenómeno dado.

El científico, para hacer ciencia, debe enmarcarse en su actividad científica, que es la investigación; al respecto, ésta se considera como “un conjunto de procesos

sistemáticos, críticos y empíricos que se aplican al estudio de un fenómeno o problema” (Hernández Sampieri et. al., 2014, p. 4). De igual manera, para Kerlinger y Lee (2002) la investigación científica es una investigación sistemática, controlada, empírica, amoral, pública y crítica de fenómenos naturales, que se va a guiar por la teoría y las hipótesis sobre presuntas relaciones entre esos fenómenos.

La educación especial, al tener el cobijo de las ciencias empíricas, marca una excelente oportunidad para que el educador especial también pueda ser un científico, ya que amplía la mirada, pues si bien su quehacer profesional es vital dentro de una institución escolar y atendiendo a la población estudiantil que lo requiera, la investigación traerá consigo un corpus de conocimiento que alimentará aún más la práctica docente, pero que a su vez, posibilitará el desarrollo de un pensamiento más crítico y sistemático sobre los fenómenos que aquejan a la educación especial.

En ese sentido la educación especial, como una ciencia aplicada, está destinada a lograr un resultado práctico y admite descripciones variadas que aluden a un conocimiento personal del experto en su campo de estudio (Villoro, 2002), así, el profesional en educación especial, al ser un experto en su área y al tratar con una ciencia aplicada, puede ser también un científico, que hace investigación y que sus resultados le pueden ser útiles para su propia práctica y para compartir sus hallazgos con una comunidad epistémica, es decir, con más docentes interesados, con psicólogos y pedagogos expertos en necesidades educativas especiales, por mencionar algunos.

Y en estos momentos podría preguntarse ¿qué caracteriza que una investigación sea científica y no solo una investigación (en un sentido común de la palabra)? Cerda (2002) menciona que existe una investigación científica cuando hay un problema que resolver, describir, explicar o identificar.

Al respecto, Ander-Egg (1983) describe que el carácter científico de una investigación hace alusión a ciertos requisitos relacionados con el método, los fines y los objetivos, es una forma de plantear problemas y buscar soluciones mediante una búsqueda con un interés teórico o práctico, en la investigación científica se adquieren conocimientos acerca de un aspecto de la realidad para actuar sobre ella, es una exploración sistemática a partir de un marco teórico que sustenta los problemas o hipótesis, se requiere de una formulación precisa del problema que se quiere investigar y un diseño metodológico que exprese los procedimientos para buscar las respuestas de las preguntas planteadas, exige una comprobación y verificación del hecho o fenómeno que se estudia mediante la confrontación empírica, va más allá de situaciones o casos particulares para hacer inferencias con validez general, utiliza instrumentos metodológicos que son relevantes para obtener y comprobar datos pertinentes a los objetivos de investigación y, finalmente, se registra y expresa en un informe.

El enfoque de la investigación

La investigación también puede tomar matices muy variados, mismos que son dados por el enfoque, ya sea cualitativo, cuantitativo o mixto. Aun así, estos enfoques emplean procesos cuidadosos, metódicos y empíricos para generar conocimiento, como afirma Hernández Sampieri et. al. (2014) cada enfoque presenta diversas características; por ejemplo el enfoque cuantitativo es secuencial y probatorio, hay un orden para realizar cada etapa: se parte de una idea que deriva en objetivos y preguntas de investigación, se revisa la literatura para construir un marco teórico, de las preguntas se establecen hipótesis y variables, se elabora un diseño para probarlas, se miden las variables, se analizan las mediciones obtenidas utilizando métodos estadísticos y se extraen conclusiones respecto de las hipótesis, es decir, se utiliza la

recolección de datos para probar hipótesis basándose en la medición numérica y el análisis estadístico para establecer pautas de comportamiento y probar teorías.

El autor antes citado, señala que el proceso cuantitativo contempla varias fases: primero se establece la idea de la investigación, luego surge el planteamiento del problema, posteriormente se debe elaborar la revisión de la literatura y el marco teórico, la visualización del alcance del estudio, la elaboración de hipótesis y definición de variables, el desarrollo del diseño de investigación, la definición y selección de la muestra, la recolección de los datos, el análisis de los datos y, finalmente, la elaboración del reporte de resultados.

De igual manera, Cerda (2002) describe que lo cuantitativo en una investigación tradicional se reduce a medir variables en función de una magnitud (toda propiedad que puede ser medida), extensión (parte del espacio que ocupa una cosa) o cantidad; la cantidad es diferente a la cualidad, suele expresarse por el número, que es una expresión de la cantidad en relación con una unidad determinada, por ejemplo, se cuantifica la dimensión, el peso, volumen de objetos, intensidad de colores y sonidos. Los rasgos de la investigación cuantitativa permiten la enumeración y medición; es así que la medición se somete a criterios de confiabilidad y validez, se utilizan las matemáticas, busca reproducir numéricamente las relaciones entre objetos y fenómenos y se relaciona con diseños o investigaciones tradicionales, como son los experimentales o cuasiexperimentales.

Diferente a lo que se hace en el enfoque cualitativo, ya que, como señalan Hernández Sampieri et. al. (2014) se pueden desarrollar preguntas e hipótesis antes, durante o después de la recolección y análisis de datos, usualmente estas actividades sirven para descubrir cuáles son las preguntas de investigación más importantes para después perfeccionarlas, responderlas o generar nuevas interrogantes en el proceso de interpretación, de tal manera que la investigación se mueve entre los hechos y

cómo los interpreta el investigador. Las fases en este enfoque no suelen llevar un proceso lineal, sin embargo, se parte de una idea, que puede llevar a un planteamiento del problema, a la inmersión inicial en el campo, al diseño del estudio, a la definición de la muestra, a la recolección y análisis de datos, a la interpretación de resultados y a la elaboración del reporte, todas estas fases enmarcadas en la literatura existente.

En este enfoque la cualidad se revela por medio de las propiedades de un objeto o de un fenómeno, y se pueden observar las siguientes características: se interpretan cosas y fenómenos que no pueden ser captados plenamente por la estadística o matemáticas, se utiliza preferentemente la inferencia inductiva y el análisis diacrónico en los datos, así como los criterios de credibilidad, transferabilidad y confirmabilidad como formas de hacer creíbles y confiables los resultados, se utilizan múltiples fuentes, métodos e investigadores para estudiar un problema o tema, los cuales convergen en un punto central de estudio, se utiliza como técnica de recolección de datos a la observación y la entrevista abierta y no estandarizada, y centra su análisis en la descripción de fenómenos y cosas observadas (Cerda, 2002).

Otro enfoque que se desprende de los mencionados anteriormente es el enfoque mixto. Éste, representa un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implica la recolección, análisis de datos, integración y discusión de los resultados cuantitativos y cualitativos, al realizar metainferencias de la información recabada para lograr un mayor entendimiento del fenómeno estudiado (Hernández-Sampieri y Mendoza, 2018).

Al respecto, Cerda (2002) afirma que no es posible referirse a un enfoque cualitativo sin referirse a uno cuantitativo, ya que en realidad ambos son parte de una misma realidad, por ejemplo, en la vida social y natural no existe un solo objeto

que posea un aspecto únicamente cualitativo o cuantitativo, de tal manera que cada fenómeno representa una unidad de determinada calidad o cualidad que es caracterizada por la medida, una categoría filosófica que expresa la unidad orgánica de calidad y cantidad; así como a cada objeto cualitativamente específico, le son inherentes determinadas características cuantitativas, variables y móviles.

El educador especial deberá determinar, de acuerdo con el fenómeno que esté estudiando, si va a inclinarse por un enfoque mayoritariamente cuantitativo, cualitativo o mixto. Ya que, como se mencionó anteriormente, cualquier enfoque es válido para hacer ciencia, medir y describir objetos de investigación para dar explicación a fenómenos de la realidad.

Alcances de la investigación

En un estudio de tipo cuantitativo, es posible analizar qué alcance tendrá la investigación, éste depende de la estrategia de la investigación, del estado del conocimiento sobre el problema a investigar, así como la perspectiva que va a tener el estudio; Hernández Sampieri et. al. (2014) mencionan cuatro alcances: exploratorio, descriptivo, correlacional y explicativo, a continuación, se describirá cada uno de ellos y se hace la acotación que, en la práctica, cualquier investigación puede tener elementos de uno o varios de estos alcances.

En un estudio exploratorio, el objetivo consiste en examinar un tema que ha sido poco estudiado o es novedoso, esto es posible cuando la revisión de la literatura indica que hay líneas no investigadas o ideas vagas y relacionadas con el problema de estudio. Por ejemplo, cuando se desean analizar fenómenos como una nueva enfermedad, una catástrofe en una región, o el surgimiento de una nueva tecnología o método educativo.

En un estudio descriptivo, el objetivo consiste en describir fenómenos, situaciones, contextos y sucesos, cómo son y se manifiestan. Con ello, se busca especificar propiedades y características importantes de cualquier fenómeno que se analice, describir tendencias de un grupo o población. Es decir, se pretende medir o recoger información sobre los conceptos y variables a las que se refieren. Por ejemplo, especificar las propiedades y perfiles de personas, grupos, procesos, comunidades, objetos o fenómenos de estudio, como medir las dimensiones de la inclusión educativa en una institución y luego describir cada una.

En un estudio correlacional, el objetivo consiste en conocer la relación o grado de asociación que exista entre dos o más conceptos, categorías o variables en una muestra o contexto en particular, mediante un patrón predecible para un grupo o población. Para evaluar el grado de asociación entre dos o más variables, primero se mide cada una de éstas y después se cuantifican, analizan y establecen vinculaciones, dichas correlaciones se sustentan en hipótesis sometidas a prueba. Por ejemplo, si se desea analizar la asociación entre la motivación para aprender y el desempeño académico de los estudiantes, se debe medir la motivación para aprender y el desempeño académico de cada estudiante y después analizar si los estudiantes con mayor desempeño académico tienen mayor o menor motivación para aprender.

En un estudio explicativo, el objetivo consiste en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta o por qué se relacionan dos o más variables. Va más allá de la descripción de fenómenos o de establecer relaciones entre variables, está dirigido a responder por las causas de los eventos y fenómenos físicos o sociales. Por ejemplo, cuando se desea indagar sobre qué efectos tiene el establecimiento de un programa de lectoescritura novedoso en estudiantes con bajo desempeño lector. Cabe mencionar que, en un estudio explicativo, están implicados

los propósitos de los demás alcances y proporcionan un sentido de entendimiento del fenómeno a que hacen referencia.

Además, es posible que una investigación se caracterice por ser únicamente de un alcance, o tome elementos de los demás, o que inicie como exploratorio o descriptivo y llegue a ser correlacional y explicativo.

El método de investigación

Si queremos elaborar un platillo, recurrimos a una receta que nos da los pasos a seguir; en ese sentido un método es una receta para hacer algo, formulado de manera explícita, es una regla o conjunto de reglas para actuar siguiendo un orden y dirección a una meta (Bunge y Ardila, 2002). A su vez, el método es la manera ordenada y sistemática de hacer las cosas o determinada cosa, aquí el orden se refiere a la manera de estar ubicadas las cosas y los elementos que hacen parte de un todo, y lo sistemático se refiere a ese conjunto de reglas y procedimientos que contribuyen a alcanzar unos fines determinados; es la manera de alcanzar un objetivo y un determinado procedimiento para ordenar la actividad que se realiza (Cerda, 2002).

Hay diversos métodos Kerlinger y Lee (2002) describen cuatro, el primero es el método de la tenacidad, que se refiere a cuando la gente sostiene firmemente la verdad, la asumen como cierta por su apego a ella, a menudo la gente se aferra a sus creencias frente a hechos que indican lo contrario. Un segundo método es el de la autoridad o de creencias establecidas, se da cuando alguien o algo notable lo dice, por ejemplo: una institución, una persona prominente, un libro religioso, y las personas lo creen como cierto. El tercer método es el a priori, el cual se basa en el supuesto de que las proposiciones aceptadas por el apriorista son en sí mismas evidentes, ya que concuerdan con la razón. Y el cuarto método es el de la ciencia,

donde las creencias no tienen cabida, el cual para obtener el conocimiento se basa en la autocorrección.

Para hacer ciencia e investigación científica es necesario utilizar el método científico, al respecto Bunge (1983) detalla sus características más importantes: es fáctico, trasciende los hechos, se atiene a reglas metodológicas, se vale de la verificación empírica, es autocorrectivo y progresivo, sus formulaciones son de tipo general y es objetivo.

Asimismo, Bunge y Ardila (2002) detallan una secuencia ordenada de operaciones de conocimiento en el método científico:

- Identificar un problema.
- Formular un problema con claridad.
- Buscar información, métodos o instrumentos con probabilidad de resultar pertinentes al problema.
- Tratar de resolver el problema con ayuda de los medios recogidos.
- Inventar nuevas ideas (hipótesis, teorías o técnicas) si no fue posible resolver el problema, con el fin de producir nuevos datos empíricos, o proyectar nuevos experimentos o artefactos para resolver el problema.
- Obtener una solución del problema.
- Extraer consecuencias de la solución.
- Controlar la solución propuesta.
- Corregir la solución defectuosa al examinar todo el procedimiento o utilizar supuestos o métodos alternativos.
- Examinar el impacto de la solución sobre el cuerpo de conocimientos del marco de referencia y formular nuevos problemas a que da lugar.

El escrito de la investigación científica

El método científico, como ya se abordó anteriormente, se caracteriza por ser sistemático y llevar a un conocimiento probado y verdadero. Por ello, a continuación, se marcan algunas pautas para el escrito detallado que da cuenta de lo que se hace en una investigación.

Para identificar un problema de investigación es posible tener conocimientos empíricos relacionados al problema, o realizar una investigación documental extensa para entender que hay un problema que debe resolverse. Esto comúnmente se conoce como planteamiento del problema.

Al formular un problema se establecen posibles explicaciones, supuestos o hipótesis que deben ser descritas en una investigación y que van a guiar la redacción de objetivos y de preguntas de investigación. Esto es lo que se conoce como el apartado de supuestos o hipótesis; del objetivo general y específicos; y de la pregunta o preguntas de investigación.

Se debe buscar información de fuentes bibliográficas confiables que den un marco de referencia teórico, metodológico, empírico y conceptual del problema que se está abordando, para tener una mayor claridad de qué autores han hablado del problema y qué soluciones se han planteado o no del mismo, empleando citas textuales o paráfrasis del texto de los autores que se están revisando. Esto es lo que se conoce como elaboración de un marco teórico y un estado del arte o estado de la cuestión.

Una vez que se tiene cierta claridad conceptual, empírica y metodológica del problema a investigar, se pueden empezar a emplear métodos ya probados o diseñar uno para encontrar la solución. En este apartado se define a los participantes, se hace el diseño de investigación, se describe la situación a analizar, los materiales

requeridos, y los pasos a seguir. Esto es lo que se conoce como elaborar la metodología.

Luego de llevar a cabo la metodología, se obtienen datos que pueden ser analizados de manera estadística o descriptiva, y los resultados se tabulan, grafican y se discuten las conclusiones o soluciones halladas, en contraste con las hipótesis planteadas, para observar si fue resuelta la pregunta de investigación y si se cumplieron los objetivos propuestos inicialmente. Esto es lo que se conoce como el apartado de resultados y el apartado de conclusiones.

Finalmente, se detalla el escrito o informe final que da cuenta de cada momento de la investigación. Usualmente se aplica un formato preestablecido para el documento y la gestión de citas y referencias, como APA, Vancouver, Chicago, y Harvard. En Ciencias Sociales y en Educación Especial, el formato por elección es APA.

Diseño de investigación

En una investigación que lleva el método científico, es necesario sistematizar los pasos a seguir o el procedimiento, esto es el diseño de la investigación, que es el esbozo, esquema, prototipo, modelo o estructura que indica el conjunto de decisiones, pasos, fases y actividades por realizar en el curso de una investigación; como una estrategia a seguir por el investigador para alcanzar una adecuada solución al problema planteado; es decir, es un modelo de verificación que permite contrastar los hechos con la teoría, y su forma es la de una estrategia o plan general que determina las operaciones para hacerlo (Cerdeña, 2002).

De manera similar, Kerlinger y Lee (2002) describen al diseño de la investigación como el plan y estructura de la investigación, y se concibe de forma

específica para obtener respuestas a las preguntas de investigación, este plan es un esquema o programa general de la investigación que incluye un bosquejo de lo que el investigador hará.

Dichos autores mencionan dos propósitos básicos, el primero se refiere a proporcionar respuestas a las preguntas de investigación, y el segundo a controlar la varianza experimental, extraña y del error del problema de investigación particular en estudio.

Asimismo, para Cerda (2002) el diseño de la investigación implica dos momentos: el primero, como un plan sistemático o serie de instrucciones para realizar el proyecto de investigación, como un plano de construcción; y el segundo, implica dinero, recursos humanos y tiempo.

De tal manera que este plan de investigación se concibe de manera deliberada y específica, ejecutado para obtener evidencia empírica que apoye al problema de investigación, problema que puede ser expresado en forma de hipótesis a ser probadas empíricamente; además, el diseño de investigación debe ser elaborado para proporcionar respuestas confiables y válidas a las preguntas de investigación contenidas en las hipótesis, establece el marco de referencia para el estudio de las relaciones entre variables e indica qué tipo de análisis estadístico emplear (Kerlinger y Lee, 2002).

Los diseños de investigación en el enfoque cuantitativo

En el enfoque cuantitativo, existen diversos diseños, por ejemplo, Hernández Sampieri, et. al. (2014) mencionan la clasificación de diseño experimental y no experimental, a su vez el diseño experimental se clasifica en diseños preexperimentales, diseños cuasiexperimentales y diseños de experimentos puros.

Por su parte, el diseño no experimental tiene a su vez la división de diseños transversales y diseños longitudinales.

El diseño experimental tiene por característica fundamental que el investigador busca manipular tratamientos, estímulos, influencias o intervenciones, llamados variables independientes, para observar los efectos en otras variables, dependientes, en una situación de control. Así, se pretende establecer el posible efecto de una causa que se manipula.

Son tres los requisitos que se deben cumplir en un experimento:

1. Manipular intencionalmente una o más variables independientes. La variable independiente es la causa, la condición antecedente, y el efecto provocado por dicha causa se llama variable dependiente, lo consecuente. La variable independiente se puede manipular en dos o más grados, el primer nivel implica la presencia o ausencia de la variable independiente, y cada nivel o grado comprende un grupo en el experimento. En el nivel de un grado, presencia o ausencia, hay dos grupos: uno experimental, que es al que se le administra el tratamiento o estímulo experimental; y uno control, que es al que no se le administra el tratamiento o estímulo experimental. En el nivel de dos más grados se puede dosificar el tratamiento o estímulo experimental, así a cada grupo experimental se le administrará una cantidad o nivel del tratamiento y al grupo control no.
2. Medir el efecto que la variable independiente tiene en la variable dependiente, la medición debe ser adecuada, válida y confiable. Para ello existen diversas técnicas y estrategias, que pueden ser a partir de instrumentos de registro, diseñados para valorar el impacto de las variables.
3. Alcanzar el control o validez interna de la situación experimental. Cuando hay una o más variables independientes que hacer variar a las dependientes, se

debe asegurar que la variación de las últimas se debe a la manipulación de las primeras y no a otros factores o causas. Lograr el control en un experimento implica contener la influencia de otras variables extrañas en las variables dependientes, para conocer en realidad si las variables independientes que interesan tienen o no efecto en las dependientes.

En los diseños preexperimentales el grado de control es mínimo y suele ser útil para acercarse al problema de investigación en la realidad, y no hay aleatorización o emparejamiento de participantes o grupos. En esta categoría se encuentra el estudio de caso con una sola medición y el diseño de preprueba/posprueba con un solo grupo.

En el estudio de caso con una sola medición se administra un estímulo o tratamiento a un grupo y después se aplica una medición de una o más variables para observar cuál es el nivel del grupo en éstas; aquí no hay manipulación de la variable independiente, grupos de contraste, ni referencia previa sobre el nivel que tenía el grupo.

En el diseño de preprueba/posprueba con un solo grupo, se administra una prueba previa al estímulo o tratamiento experimental, después se le administra el tratamiento al grupo y luego se aplica una prueba posterior al estímulo; en este diseño no hay grupo de comparación.

En los diseños de experimentos puros es posible alcanzar el control y validez interna, ya que hay grupos de comparación y equivalencia de los grupos (ya sea por asignación al azar; o por emparejamiento, donde se iguala a los grupos en relación con alguna variable específica). En esta categoría se encuentran:

- Diseño con posprueba únicamente, y grupo de control: se incluyen dos grupos, uno recibe el tratamiento experimental y el otro no. Los sujetos se asignan de

manera aleatoria y cuando concluye la manipulación, a ambos grupos se les administra una medición sobre la variable dependiente en estudio.

- Diseño con preprueba-posprueba y grupo de control: se incorpora la administración de prepruebas a los grupos que componen el experimento. Los participantes se asignan al azar a los grupos y después se les aplica simultáneamente la preprueba, un grupo recibe el tratamiento experimental y otro no (el grupo control), finalmente, se les administra simultáneamente una posprueba.
- Diseño de cuatro grupos de Solomon: consiste en tener dos grupos experimentales y dos grupos control, únicamente a un grupo experimental y un grupo control se les administra la preprueba, y a los cuatro grupos se les aplica la posprueba. Todos los participantes se asignan de manera aleatoria.
- Diseños experimentales de series cronológicas múltiples: útiles para analizar efectos en el mediano o largo plazo. Se efectúan al paso del tiempo varias observaciones o mediciones sobre una o más variables, sea experimental o no, pueden tener dos o más grupos y los participantes se asignan al azar.
- Diseños factoriales: útiles para analizar el efecto que sobre las variables dependientes tiene la manipulación de más de una variable independiente. Se manipulan dos o más variables independientes e incluyen dos o más niveles o modalidades de presencia en cada una de las variables independientes. Todos los niveles o modalidades de cada variable independiente son tomados en combinación con todos los niveles o modalidades de las otras variables independientes.

Por su parte, en los diseños cuasiexperimentales está implicado el uso de grupos intactos, también manipulan deliberadamente, al menos, una variable independiente para observar su efecto sobre una o más variables dependientes.

Difieren de diseños de experimentos puros en el grado de seguridad que puedan tener sobre la equivalencia inicial de los grupos, ya que los sujetos no se asignan al azar a los grupos ni se emparejan, sino que los grupos ya están conformados antes del experimento.

En los diseños no experimentales, la investigación se realiza sin manipular deliberadamente las variables, son estudios en los que sólo se observan los fenómenos en su ambiente natural para analizarlos. Son situaciones ya existentes, no provocadas intencionalmente en la investigación por quien la realiza.

Dentro de los diseños no experimentales se encuentran los diseños transeccionales o transversales, y los longitudinales o evolutivos, los primeros recolectan datos en un solo momento, en un tiempo único, su propósito es describir variables y analizar su incidencia e interrelación en un momento dado, es como una fotografía de algo que sucede. Pueden abarcar varios grupos o subgrupos de personas, objetos o indicadores, comunidades, situaciones o eventos. Estos se dividen en:

1. Diseños transeccionales exploratorios: su propósito es conocer una variable o un conjunto de variables, una comunidad, un contexto, un evento, una situación. Es una exploración inicial en un momento específico, generalmente se aplican a problemas de investigación nuevos o poco conocidos.
2. Diseños transeccionales descriptivos: su objetivo es indagar la incidencia de las modalidades, categorías o niveles de una o más variables en una población y proporcionar su descripción.
3. Diseños transeccionales correlacionales-causales: describen la relación entre dos o más categorías, conceptos o variables en un momento determinado, se puede limitar a establecer relaciones entre variables sin precisar sentido de causalidad o analizar relaciones causales, ya sea en términos correlacionales

o en función de la relación causa-efecto. En este tipo de diseños, las causas y efectos ya ocurrieron en la realidad o suceden durante el desarrollo del estudio, y quien investiga los observa y reporta.

En cuanto a los diseños longitudinales o evolutivos, el investigador analiza cambios al paso del tiempo en determinadas categorías, conceptos, sucesos, variables, contextos, comunidades, o la relación entre éstas. En este tipo de diseños se recaban datos en diferentes puntos del tiempo, para realizar inferencias acerca de la evolución del problema de investigación o fenómeno, sus causas y efectos, estos puntos o periodos generalmente se especifican de antemano. Estos diseños se dividen en:

- Diseños de tendencia: se analizan cambios al paso del tiempo en categorías, conceptos, variables o sus relaciones de alguna población en general. La atención se centra en la población o universo. Se puede estudiar a toda la población, o tomar una muestra de ella cada vez que se observen o midan las variables o relaciones entre éstas, los participantes o casos de investigación no son los mismos, pero la población sí.
- Diseños de análisis evolutivo de grupos (cohorte): se examinan cambios a través del tiempo en subpoblaciones o grupos específicos, su atención son las cohortes o grupos de individuos vinculados de alguna manera o identificados por una característica común, por ejemplo, la edad, época o región geográfica. Estos diseños hacen un seguimiento de los grupos al paso del tiempo y por lo común, se extrae una muestra cada vez que se recolectan datos sobre el grupo o la subpoblación, más que incluir a toda la subpoblación.
- Diseños panel: en estos diseños son los mismos casos o participantes que son medidos u observados en todos los tiempos o momentos. Los individuos, y no sólo la muestra, población o subpoblación, son los mismos. Como ventaja,

además de conocer los cambios grupales, se conocen los cambios individuales.

Los diseños de investigación en el enfoque cualitativo

De acuerdo con Denzin y Lincoln (2001), la investigación cualitativa constituye una perspectiva interpretativa de la experiencia humana, se vale de la sensibilidad interpretativa, a través de la recolección de material empírico relacionado con los contextos donde interaccionan las personas. Son diversos los diseños de investigación que se pueden abordar en el enfoque cualitativo, su elección depende del planteamiento del problema, la pregunta de investigación y el trabajo de campo, por ejemplo, Hernández Sampieri, et. al. (2014) mencionan diversos diseños, entre los que destaca la teoría fundamentada, el diseño etnográfico, el diseño narrativo, el diseño fenomenológico, el diseño de investigación-acción e investigación-acción participativa, y el estudio de caso. Para fines de este capítulo solo se hace mención de los principales diseños en el enfoque cualitativo, ya que uno de los objetivos del presente libro es explicar a mayor profundidad los Principios de estadística para el educador especial, mismos que resaltan más en el enfoque cuantitativo y mixto.

Los diseños de investigación en el enfoque mixto

El enfoque mixto se puede clasificar en función de qué secuencia decide realizar el investigador en cuanto al muestreo, recolección, análisis e interpretación de datos; de tal manera se tienen diseños concurrentes, secuenciales, de conversión y de integración de procesos. Por ejemplo, en una ejecución o diseño concurrente, se aplican ambos métodos de manera simultánea: datos cuantitativos y cualitativos se

recolectan y analizan más o menos al mismo tiempo (Hernández Sampieri et. al., 2014).

De acuerdo con Onwuegbuzie y Johnson (2006) son cuatro las condiciones y características de los diseños concurrentes:

- Se recaban en paralelo y de forma separada los datos cuantitativos y cualitativos.
- El análisis de datos cualitativos y cuantitativos se ejecuta de manera separada, es decir, uno no es la base del otro.
- Los resultados de ambos análisis se consolidan cuando ambos conjuntos de datos se recolectaron y analizaron de manera separada.
- Luego de la recolección de datos e interpretación de resultados se establecen metainferencias que integran las conclusiones de ambos métodos y su conexión.

Por otro lado, en la ejecución o diseño secuencial primero se recolectan y analizan datos de los métodos cualitativos o cuantitativos, y luego se recaban y analizan datos del otro método (Hernández Sampieri et. al., 2014). Como característica distintiva, los datos recolectados y analizados en una fase de estudio (ya sea cualitativo o cuantitativo) se emplean para informar a la otra fase (cualitativa o cuantitativa), de este modo el análisis inicia antes que todos los datos sean recabados (Onwuegbuzie y Johnson, 2006).

En cuanto al diseño de conversión, en éste se transforman los datos para su análisis, Hernández Sampieri et. al. (2014) explican que esto sucede cuando un tipo de datos es convertido en otro, es decir, se cuantifican datos cualitativos, o se cualifican datos cuantitativos; luego se analizan ambos conjuntos de datos bajo la lógica del método cualitativo o cuantitativo.

Finalmente, en el diseño de integración de procesos, ambas aproximaciones, cualitativa y cuantitativa, se entremezclan desde el inicio hasta el final en la mayoría de las etapas y se oscila entre el pensamiento inductivo y deductivo. Hernández Sampieri y Mendoza (2018) mencionan las siguientes características:

- Se recolectan datos cuantitativos y cualitativos, de manera simultánea o en diferentes secuencias.
- Se elaboran análisis cuantitativos y cualitativos sobre los datos de ambos tipos durante todo el proceso.
- Se pueden involucrar otros diseños en el mismo estudio, por ejemplo, los experimentos.
- Los resultados finales se reportan hasta el final.
- El proceso es iterativo.
- Son diseños útiles para abordar problemas complejos.
- Los resultados se pueden generalizar y se puede desarrollar al mismo tiempo teoría emergente, probar hipótesis, explorar, etc.

Hasta aquí se han abordado de manera general los tres enfoques de la investigación para aproximarse al campo de la educación especial; en cada uno, se han presentado de manera sucinta sus principales características, estos aspectos no son totalitarios, y como este libro tiene la intención de discutir con mayor profundidad los principios estadísticos para el educador especial, a continuación se presentan a manera de ejemplo los hallazgos de diversas investigaciones con un enfoque cuantitativo y que hacen uso de la estadística.

Ejemplos de estudios en educación especial y su relación con la estadística

Para ejemplificar como es el uso de la estadística en estudios que involucran temas educativos y de educación especial, se iniciará reportando el estudio realizado por Sánchez-Escobedo et al. (2020), el cual tuvo como objetivo describir la aceptación social de alumnos con Alto Rendimiento Académico (ARA), el término refiere a alumnos que sobresalen por su esfuerzo, dedicación, con inteligencia típicamente normal y reconocidos por los maestros como buenos estudiantes. Los participantes fueron estudiantes de bachillerato, conformándose una muestra de 248 alumnos (49.6%) de la ciudad de Obregón, Sonora y 252 (50.4%) de la ciudad de Mérida, Yucatán. Los autores se valen de tablas para ejemplificar de mejor manera la población que participó en el estudio, de esta manera se pueden aportar más información, como por ejemplo el género, el grado escolar. A continuación, se muestra la tabla 1 con dicha información.

Tabla 1. Alumnos participantes del estudio.

	Sonora		Yucatán		Total	%
	Mujeres	Hombres	Mujeres	Hombres		
Primero	62	61	17	68	214	43
Segundo	33	40	52	31	156	31
Tercero	28	18	19	65	130	26

Fuente: Elaboración propia (Sánchez-Escobedo et al. 2020).

Las tablas permiten observar datos interesantes, por ejemplo se observa que participaron más alumnos de primer semestre, en Mérida fueron menos mujeres las

que participaron en primer y tercer grado, mientras que en Sonora se observa mayor equilibrio en la participación de ambos generos.

Como en cada reporte de investigación, se describe el instrumento para recolectar la información y su procesamiento de análisis, en este caso correspondió a una Escala de Aceptación Social, que valora tres dimensiones: aceptación en actividades académicas, sociales y de amistad, el cual se compone de ítems, con formato de respuesta Likert.

En investigación cuantitativa siempre se reporta la confiabilidad del instrumento, usualmente se utiliza coeficiente de Alfa de Cronbach, el cual entre más cercano este a .1 es más alta su confiabilidad, lo anterior significa que la aplicación del instrumento siempre mostrará resultados consistentes y coherentes con los que se pregunta. En el caso de la escala aplicada, se obtuvo un Alfa de Cronbach de .91.

Posterior de la aplicación de los instrumentos, se hizo la captura de las respuestas en un programa especializado para realizar análisis estadísticos, el cual se nombra en español como Paquete Estadístico para las Ciencias Sociales (SPSS), este programa tiene diferentes versiones, en el caso de este estudio correspondió a la versión 23.

De los análisis realizados, se hicieron ejercicios descriptivos, que consistió en la sumatoria de las respuestas y extraer los puntajes de medias y desviaciones estándar; con respecto a la definición de media son resúmenes de los conjuntos de las medidas a partir de las cuales se calculan, indican qué conjuntos “son parecidos” en promedio, pero también son comparados para probar relaciones, y la desviación estándar, consiste en analizar las puntuaciones ordinarias e identificar el nivel de dispersión del conjunto de dichas puntuaciones, entre menos dispersión mayor

homogeneidad de respuesta del grupo que se estudia, entre mayor dispersión menor heterogeneidad de las respuestas, en la tabla 2 se observan los datos anteriormente descritos con la variable de actividades sociales y amistad.

Tabla 2. Medias y las Desviaciones estándar

	N	M	DE
Aceptación social	500	66.86	15.20
Actividades escolares	500	24.68	4.50
Actividades sociales	500	21.97	6.07
Amistad	500	20.20	5.24

Fuente: Elaboración propia (Sánchez-Escobedo et al. 2020)

Estos datos son importantes por que permiten hacer inferencias con respecto a las respuestas que tuvieron los informantes, por ejemplo, en la dimensión de aceptación social, las actividades escolares tuvieron la media más alta de 24.68, seguido de actividades sociales por 21.97 y finalmente amistad con 20.20, lo que indica que hay aceptación social hacia los alumnos, pero es más común en las actividades escolares, que en las de amistad, mientras que la desviación estándar, es decir en la dispersión de los datos, se observa que es mayor en actividades sociales y de amistad, lo es que es congruente con las medias. Los autores llegan a la conclusión de que los estudiantes poseen una percepción social positiva respecto a los alumnos con ARA, así mismo no se evidencia el costo social del alto rendimiento académico o rechazo a estos alumnos por parte de sus pares. De las comparaciones por grupos según sexo, las mujeres reportaron mayor promedio de calificaciones y una aceptación social mayor, respecto a los alumnos con ARA hombres.

Otra investigación para comentar en este capítulo es la realizada por Ordaz y Acle (2021), la cual tuvo como objetivo conocer el perfil de desempeño matemático de los aprendizajes clave en alumnos de los tres primeros grados de educación primaria a través de un sistema de rúbricas como instrumento de evaluación. Los participantes fueron 444 alumnos de cinco diferentes escuelas primarias ubicadas en la zona oriente de la delegación Iztapalapa de la ciudad de México. La edad de los participantes fue de 6 a 9 años con una media de 7.39 años, 44.6 por ciento niños y 55.4 por ciento niñas. Para la obtención de los datos se utilizó el instrumento de Evaluación del Logro Matemático (ELMA), diseñado por una de las autoras, dicho instrumento está conformado por a) instrucciones para el aplicador, b) ficha de identificación, c) protocolo de aplicación para 1º, 2º, y 3º de primaria, d) fichas de trabajo y, por último, e) protocolo de evaluación. El protocolo consta de 22 ítems para primero y 25 para segundo y tercer grado. El sistema de rúbricas de calificación comprende cuatro descriptores consolidado (3), en proceso (2), inicial (1) e insuficiente (0) y está compuesto por dos áreas generales: la primera por habilidades pre académicas, referidas a las habilidades fundamentales que le permitirán al niño/niña asimilar los conocimientos matemáticos, están compuestos por direccionalidad, seriación, clasificación y conservación; y, la segunda, por aprendizajes curriculares clave, que se basa en el Plan y Programas de Estudios (SEP, 2011; 2017), específicamente en el eje de sentido numérico y pensamiento algebraico.

Este estudio da la posibilidad de analizar dos aspectos, el primero de ellos es la validación del instrumento, diversas investigaciones cuantitativas se enfocan en realizar este tipo de procedimientos, validar instrumentos que pueden ser aplicados a diversas poblaciones, por ejemplo algún tipo de test que mide autoestima y que fue diseñado en Europa, se valida ahí, y luego se hace una validación en México, pues

las formas de organizar las oraciones o algunas palabras puede tener un significado distinto.

Retomando el ejemplo que se está analizando de Ordaz y Acle, (2021), las autoras hacen una validez de contenido, para ello recurren a ocho jueces expertos, cuatro de ellos profesores de los primeros grados de primaria, que tienen estudios de nivel licenciatura hasta doctorado, así como dos docentes de la maestría en educación especial de la UNAM, y dos doctoras en psicología del área de Educación y Desarrollo Humano de la misma institución. Se aplicó la prueba W de Kendall con respecto a cuatro criterios, claridad, coherencia, relevancia y suficiencia de la cantidad de los reactivos de cada dimensión, lo que les permite concluir que el instrumento presenta una sintaxis y semántica adecuada, así como una relación lógica con cada dimensión estudiada, adicionalmente los reactivos fueron relevantes y su número fue conveniente para la medición de cada dimensión, a continuación, se muestra la tabla con los valores.

Tabla 3. Valores del coeficiente de W de Kendall

Criterio	W de Kendall	Significancia
Claridad	.934	.0000
Coherencia	.910	.0000
Relevancia	.843	.0000
Suficiencia	.875	.0008

Fuente: Elaboración propia (Ordaz y Acle, 2021).

La prueba W de Kendall, fue desarrollada en 1948, busca el coeficiente de concordancia, el cual consiste en que grupos de evaluadores asigna una calificación

de acuerdo a los rangos previamente establecidos, sino hubiera asociación alguna entre dos de los grupos evaluadores, y se calculara un coeficiente de correlación del orden de los rangos entre los rangos, éste debería ser cercano a cero; y si existe acuerdo, el coeficiente debería ser significan diferente de cero, como es el caso, podemos ver que los puntajes están alejados del cero y cercanos al uno, lo que indica que existe acuerdo entre los evaluadores que el instrumento es adecuados. Además, se aplicó la prueba Alpha de Cronbach, donde los puntajes fueron los siguientes .820 para la prueba de primer grado, .761 para la prueba de segundo y .722 en la prueba de tercer grado; que como se ha explicado anteriormente entre más cercano este a uno, más alta es la confiabilidad.

En continuidad con el estudio, después de aplicar el instrumento a los alumnos de primero, segundo y tercer grado, se concentró la información en el software SPSS y se realizaron dos análisis a) análisis descriptivo del desempeño de los alumnos de cada uno de los indicadores y b) pruebas no paramétricas Kruskal-Wallis, ésta prueba consiste en analizar el ordenamiento de los rangos, se pueden considerar un factor o de dos factores, o análisis de medidas repetidas, esta fórmula permite evaluar la significancia de las diferencias y U de Mann Whithney, analiza las igualdades de las dos distribuciones, a continuación se muestra en la tabla 4 los resultados del estudio.

Tabla 4. La construcción del número. Frecuencias y porcentajes del nivel de desempeño.

Nivel de desempeño	Grado escolar					
	1°		2°		3°	
	N	%	N	%	N	%
Insuficiente	2	1.44	1	0.60	4	0.92
Inicial	8	5.76	15	8.93	34	7.57
En proceso	26	18.71	57	33.93	127	25.25
Consolidado	103	74.18	95	56.55	279	63.26

Fuente: Elaboración propia (Ordaz y Acle, 2021)

En dicha tabla se observa que, en la construcción del número, que refiere a realizar representaciones no convencionales y convencionales de los objetos e identificar adecuadamente cantidades de acuerdo al grado escolar, el 74% de los niños de primer grado tienen un nivel de consolidados, mientras que este se reduce en segundo grado al pasar a un 56.55%, y el porcentaje se mantiene con poco cambio en tercer grado con un 59.12%.

En cuanto a la prueba Kruskal-Wallis no se encontraron diferencias para el sistema decimal numérico, algoritmo de suma, algoritmo de multiplicación, problemas de multiplicación, así como para problemas de división ($p > .05$). Sin embargo, si hubo diferencias para construcción del número, sucesiones, calculo mental, algoritmo de resta y problema de aditivos ($p < .05$). Esto se explica porque cuando “p” es mayor ($>$) a 0.05, no hay diferencias en los resultados y cuando “p” es menor ($<$) a 0.05 si hay diferencias en los resultados, la $p > .05$ significa la probabilidad de error, porque al multiplicar el 0.05 por cien es 5%, esa probabilidad de error cuando es mayor es que se tiene una mayor probabilidad para equivocarse, sí, se afirma que son diferentes, por tanto, entre menor sea a 0.05 menor probabilidad hay de equivocarse.

Con respecto a los resultados aplicando la prueba U Mann Whitney, se pueden observar los desempeños en la tabla 5.

Tabla 5. Resultados de la prueba U Mann Whitney, medida de tendencia central

Resultados de U de Mann Whitney			
Aprendizaje Clave	Medias		
	3°	2°	1°
Número	5.00	4.92	5.30
Sucesiones	1.79	1.97	1.94
Cálculo mental	4.99	7.90	8.27
Algoritmo resta	2.09	1.91	1.73
Problemas aditivos	7.06	6.98	4.73

Fuente: Elaboración propia (Ordaz y Acle, 2021)

Las autoras indican que el desempeño de los alumnos de primer año fue superior respecto al de segundo en la construcción del número, mejorando los puntajes para segundo grado en la resolución de problemas aditivos, mientras que no hubo diferencias para sucesiones, cálculo mental ni para el algoritmo de la resta.

Tabla 6. Resultados de la prueba U Mann Whitney, probabilidades de error

Resultados de U de Mann Whitney			
Aprendizaje Clave	Comparación por grados		
	1° con 2°	1° con 3°	2° con 3°
Número	P<0.01	P<0.01	-.96
Sucesiones	-.64	-1.41	P<0.01
Cálculo mental	-.15	P<0.01	P<0.01
Algoritmo resta	-.15	P<0.01	P<0.01
Problemas aditivos	P<0.01	P<0.01	-1.23

Fuente: Elaboración propia (Ordaz y Acle, 2021)

Como puede observarse en la tabla 6, los valores ahora se manejan en “p” menor ($<$) que .0.1, que aplica el mismo principio de $p < 0.05$, en este caso la probabilidad de equivocarse es menor al 1%.

En un análisis de comparar primero con tercero se encontraron diferencias en todos los desempeños excepto en sucesiones: a favor de los alumnos de primer año se ubicó el conocimiento del número y el cálculo mental; en tercero el desempeño fue superior en el algoritmo de la resta y la resolución de los problemas aditivos.

La conclusión a la que llegan las investigadoras es que se puede afirmar que el sistema de rúbricas es capaz de delimitar el dominio de los aprendizajes claves señalados por diferentes planes, programas y materiales de estudio, además de que las rúbricas como instrumento de evaluación, pueden utilizarse desde los primeros grados de educación básica.

La investigación cuantitativa, además de aplicar instrumentos que permiten aproximarse a valores sobre diversos temas, también posibilita comparar resultados previos a la aplicación de un experimento; a continuación, se presenta el ejemplo de una investigación cuasiexperimental efectuada por Uriel y Franco en el año 2020. El objetivo del estudio fue realizar una intervención terapéutica fonoaudiológica en lectura y escritura, acompañados por terapias alternativas (Equinoterapia). La población objeto del estudio fueron 34 personas adolescentes y adultos jóvenes (entre los 11 y 42 años) que tienen un diagnóstico de discapacidad cognitiva (intelectual) y asisten a una institución educativa especial de la ciudad de Santander, Colombia. Para realizar esta investigación se solicitaron los permisos correspondientes a las autoridades educativas de la institución, así como a padres de familia y los propios participantes.

El instrumento utilizado para evaluar la lectura y la escritura se fundamentó en el material para la investigación clínica del lenguaje de Espeleta (1996), el cual se aplicó antes y después de la intervención. Mientras que la intervención se realizó de manera individual con cada participante se efectuaron 20 sesiones fonoaudiológicas que consistieron en escritura espontánea de orientación temática, con apoyos sensorio-perceptuales, escritura automática al dictado, lectura automática de voz alta, lectura comprensiva en voz alta y oraciones desordenadas, así como aspectos cognitivos comunicativos, complementando lo anterior se realizaron sesiones de equinoterapia que se realizaba dos veces por semana de treinta minutos cada sesión, el cual posibilita mejorar la postura corporal y el control del tronco.

Para el análisis de la información se utilizó el paquete estadístico en Stata v13 que permitió realizar pruebas de tendencia central, y para realizar análisis bivariado se utilizó la prueba de McNemar.

De los resultados iniciales de la lectoescritura (pre-intervención) de los 34 estudiantes evaluados, 28 presentan retardo de la lectoescritura por secuela del retardo de patogenia anártrico-alálico correspondiente al 82% de la población y seis presentan retardo de la lectoescritura por secuela del retardo de patogenia audiógeno – alálico, que corresponde al 18%.

Posterior a la aplicación de la intervención donde solo concluyeron 29 de los participantes, se hicieron las pruebas estadísticas, a continuación, se muestra en la tabla 7 los principales resultados.

Tabla 7. Habilidades de lectura y escritura.

Variables	Antes de la intervención (n=34)			Después de la intervención (n=29)		
	No realiza (%)	Primer nivel (%)	Segundo nivel (%)	No realiza (%)	Primer nivel (%)	Segundo nivel (%)
Escritura por copia	0	76.47	23.53	0	62.07	37.93
Escritura espontánea con orientación temática	41.18	38.24	20.59	13.79	48.28	37.93
Escritura espontánea con orientación	35.29	44.12	20.59	13.79	25.59	58.62
Escritura espontánea con apoyo sensorceptual	41.17	41.18	17.65	13.79	48.28	37.93
Lectura automática en voz alta	29.41	44.12	26.47	27.58	44.83	27.59
Lectura comprensiva en voz alta	29.41	50	20.59	27.59	51.72	20.69
Lectura comprensiva de oraciones desordenadas	26.47	58.82	14.71	25	60.71	14.29

Fuente: Elaboración propia (Uriel y Franco en el año 2020)

En dicha tabla se pueden observar los cambios en los puntajes, por ejemplo, el aspecto más notorio corresponde a la escritura espontánea con orientación, hay un cambio entre la primera prueba y la segunda, posterior a la aplicación de la intervención de 38.03 puntos, y de manera general se observa un aumento en el porcentaje de individuos en todos los tipos de lectura en voz alta.

Mientras que para establecer si los resultados obtenidos en los diferentes tipos de escritura y lectura automática en voz alta presentan significancia estadística; se aplicó el test de McNemar para comparación de proporciones en grupos emparejados, a continuación, se muestra la tabla con dicho análisis.

Tabla 8. Análisis Bivariado de las habilidades

Variables	Pre- Intervención	Post- Intervención	p*
Lectura automática en voz alta	26.47	27.59	0.2500
Escritura con copia	23.53	37.93	0.0625
Escritura espontánea con orientación temática	20.59	37.93	0.0625
Escritura espontánea con orientación	20.59	58.62	0.0010
Escritura espontánea con apoyo Sensoperceptual	17.65	37.93	0.0313
Escritura automática al dictado	23.54	34.48	0.2500

Fuente: Elaboración propia (Uriel y Franco en el año 2020)

En este ejercicio es posible notar que la escritura espontánea con orientación y la escritura espontánea con apoyo sensoperceptual, fueron las habilidades que presentaron cambios estadísticamente significativos ($p < .05$), aspecto que permite a los autores concluir que la equinoterapia en los individuos es efectiva para mejorar estas habilidades de lectoescritura pasando del nivel de lectura automática en el proceso de silabeo, a un nivel de silabeo con reintegración y en la escritura espontánea (redacción) de un nivel secundario de descripción enumerativa con esbozos de iniciación literaria.

Con este ejercicio estadístico es posible valorar cómo a partir de una intervención intencionada, controlada y monitoreada existen cambios comportamentales y de aprendizaje, por ello la importancia de realizar investigaciones de tipo experimental.

Conclusiones

El campo de la educación especial se alimenta de la investigación, pues en ella se plantean programas de intervención y conocimiento nuevos que buscan mejorar las prácticas docentes, también se aplican pruebas y test con el fin de tener acercamientos a la realidad que viven las familias, los docentes y los niños en diversos aspectos, como puede ser la resiliencia, autoestima, motivación, estrés, ansiedad, actitudes, percepciones, etc. O bien identificar perfiles de desempeño en ámbitos como la matemática, la comprensión lectora, las ciencias naturales, por mencionar algunos.

La innovación y la creatividad son elementos sustanciales para realizar investigaciones, pues el campo de la educación especial aun requiere que se profundicen sobre las diversas discapacidades, trastornos generales del neurodesarrollo, aptitudes sobresalientes, entre otras condiciones como los problemas de aprendizaje, de lenguaje, etc.

A lo largo de este capítulo se han expuesto los diversos enfoques y métodos con los cuales se pueden abordar los diversos fenómenos que circunscriben en la educación especial, de manera particular se ha puesto especial énfasis en mostrar cómo la investigación cuantitativa y el uso de la estadística tiene una aplicabilidad importante en educación especial, pues posibilitan la generalización y comprobación de nuevos métodos de atención a partir del trabajo con diseños no experimentales y/o experimentales.

Se le invita a los educadores especiales, que se encuentran activos en la profesión, así como aquellos que están en formación, para que se atrevan a soñar e investigar sobre aquellos temas que sean de su interés, y que pueden ser tan creativos e innovadores que revolucionen o incrementen los conocimientos actuales que se

tienen sobre la concepción, valoración y atención de las personas objeto de esta profesión.

Referencias bibliográficas

- Álvarez-Gayou Jurgenson, J.L. (2009). *Cómo hacer investigación cualitativa. Fundamentos y metodología*. Paidós.
- Ander-Egg, E. (1983). *Técnicas de investigación social*. Humanitas.
- Bunge, M. (1983). *Exploring the world*. Reidel.
- Bunge, M., y Ardila, R. (2002). *Filosofía de la Psicología*. Siglo XXI editores.
- Cerda, H. (2002). *Los elementos de la investigación. Como reconocerlos, diseñarlos y construirlos*. Ed. El Búho.
- Denzin, N.; y Lincoln, Y. (2011). *El campo de la investigación cualitativa. Manual de Investigación Cualitativa*. Vol. I. Ed. Gedisa.
- Hernández Sampieri, R., Fernández Collado, C., y Baptista Lucio, M. P. (2014). *Metodología de la investigación*. McGraw Hill.
- Hernández-Sampieri, R., y Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. McGraw Hill.
- Kerlinger, F.N. y Lee, B.H. (2002). *Investigación del comportamiento*. McGraw-Hill.
- Jurado de los Santos, P., y Sanahuja Gavaldá, J.M. (1997). La investigación en educación especial. Tendencias y orientaciones. *Educar 21*, 105-114.
- Ordaz, G., y Acle, G. (2021). Desempeño matemático. Evaluación por rúbricas en los primeros grados de educación básica. *Perfiles Educativos*, 43(173).
<https://doi.org/10.22201/iisue.24486167e.2021.173.59772>
- Onwuegbuzie, A.J., y Johnson, R.B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48-63.

Uriel, L.M. y Franco, S.G. (2020). Fonoaudiología y equinoterapia: efectividad para el desarrollo de lectura y escritura en personas con discapacidad cognitiva. *Revista Lasallista de Investigación*, 17(2), 27-40. <https://doi.org/10.22507/rli.v17n2a2>

Sánchez-Escobedo, P., Camelo, A.K., Vales, J.J., y Barrera-Hernández, L.F. (2020). Aceptación social de alumnos con alto rendimiento académico en México: Un estudio comparativo entre Yucatán y Sonora. *PSICUMEX*, 10(1), 92-102. <https://doi.org/10.36793/psicumex.v10i1.344>

Villoro, L. (2002). *Creer, saber, conocer*. Siglo XXI editores

¿PARA QUÉ CUANTIFICAR EN EDUCACIÓN ESPECIAL?

*Josué Antonio Camacho Candia [‡]
Myriam Nayeli Villafuerte Vega
María Isabel Solís Meza*

Resumen

En el presente capítulo se apoya la postura que señala que la educación especial debe ser vista como una disciplina científica multidimensional y multiparadigmática en la que convergen diversas profesiones y disciplinas, por lo tanto, son varios los conocimientos, herramientas, instrumentos y teorías que pueden sumarse para mejorar la práctica educativa en la educación especial. Entre los anteriores se enfatiza la importancia de la estadística, la cual puede ser vista en la educación especial como una herramienta fundamental para organizar e interpretar la información referente al desempeño, maduración y afectación de habilidades, competencias y capacidades tanto básicas como adaptativas, sociales y académicas de personas con discapacidad, con alteraciones del neurodesarrollo o con aptitudes sobresalientes. Lo anterior con la finalidad de fortalecer la práctica de la educación especial, haciendo más objetivas sus evaluaciones, interpretaciones y recomendaciones, lo cual mejorará la comunicación multidisciplinaria inherente en la educación especial. Señalando en las conclusiones, la necesidad de tener presente que la estadística debe basarse en teorías, conceptos y metodologías claras, coherentes y probadas, para evitar falacias interpretativas con apoyo estadístico.

Palabras clave: educación especial, estadística, inclusión, cuantificación, evaluación

[‡]Facultad de Ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala.
Carretera Tlaxcala-Puebla Km. 1.5. Tlaxcala, Tlax. jacamachoc_fcdh@uatx.mx

Introducción

La actividad profesional relacionada con la educación institucionalizada es organizada, evaluada y administrada por instancias oficiales que, tanto a nivel federal como estatal, establecen los objetivos educativos, los paradigmas a seguir, las herramientas a utilizar y la política educativa que determinará la organización escolar en distintos planteles educativos.

En el caso de la educación especial, las directrices oficiales también determinan su organización, modelos educativos, prácticas educativas, modelos pedagógicos y formas de evaluación. Sin embargo, es importante considerar que la planeación de la educación regular en general no está diseñada para población que presenta algunas condiciones particulares de discapacidad, alteraciones del neurodesarrollo o aptitudes sobresalientes, sino que está enfocada en personas con neurodesarrollo típico, así como con desarrollo físico, sensorial y lingüístico sin alteraciones. Las exigencias educativas y de salud para estas personas han hecho necesaria la creación de instituciones especializadas en el tratamiento y atención de algunas condiciones como los trastornos del neurodesarrollo, síndromes, discapacidades y trastornos que afectan el desarrollo motor, sensorial y cognitivo de un individuo, lo que hace necesario no solo la atención médica, sino también una atención educativa especializada e individualizada en centros o instituciones preparadas para ello y con personal altamente capacitado. En términos educativos, además, se hace necesaria y urgente una autonomía educativa plena, orientada al desarrollo individual, académico y social de las personas que sea acorde con sus necesidades particulares.

En general, la educación especial sigue siendo considerada como un servicio educativo destinado para los educandos que se enfrentan a Barreras para el Aprendizaje y la Participación (BAP), con mayor prioridad hacia los que tienen

alguna discapacidad, trastornos en el desarrollo, o aquellos con aptitudes sobresalientes. Siendo deber del educador especial atender a los alumnos de manera adecuada, con equidad social y según corresponda a sus propias condiciones y necesidades (Comisión de Política Gubernamental en Materia de Derechos Humanos, 2016).

El Artículo 7 de la Ley General de Educación en México menciona que se deben atender las capacidades, circunstancias, necesidades, estilos y ritmos de aprendizaje de los educandos, eliminando las distintas barreras para el aprendizaje y la participación, estableciendo la educación especial para todos los tipos, niveles, modalidades y opciones educativas. El Artículo 35 de la misma Ley señala que la educación especial buscará la equidad y la inclusión, la cual deberá estar disponible para todos los tipos, niveles, modalidades y opciones educativas establecidas en esta Ley y podrá impartirse educación con programas o contenidos particulares para ofrecer una oportuna atención (Ley General de Educación, 2019).

Si bien es cierto que la educación especial puede ser entendida en algunos casos como un servicio, tal como se leen en la mayor parte de los documentos oficiales, también es necesario que deba ser vista como una disciplina científica multidimensional y multiparadigmática en la que convergen diversas profesiones, ya que esto nos permitirá revisar las metodologías que emplea la educación especial para favorecer el aprendizaje cuando existen condiciones de desarrollo que demandan otras herramientas, espacios, y mucha mayor rigurosidad y profundidad teórica (Camacho et al., 2022; Acle, 2011).

Sin embargo, desde la creación de la Dirección General de Educación Especial (DGEE) en 1970 hasta nuestros días, la educación especial se ha visto influenciada por paradigmas que de forma temporal determinan la forma de comprender a la discapacidad, las alteraciones del neurodesarrollo y las aptitudes sobresalientes que

presenta la población que se atiende. Pero además, como parte de esa regulación oficial, la educación especial como servicio de apoyo se ha visto determinada por los parámetros generales que se aplican a la educación regular, considerando que su función de apoyo a ésta no permite que se le otorgue mayor autonomía en cuanto a su planeación académica, para que en las escuelas de educación especial se puedan establecer objetivos propios de acuerdo a las condiciones de diversidad que presentan sus estudiantes. De esta forma, se trasladó de la educación regular a la educación especial no solo los modelos educativos vigentes, sino también muchas prácticas educativas que en la misma educación regular resultan cuestionables. Un caso de lo anterior es la forma de planear y evaluar los avances o logros educativos.

Con respecto a lo anterior, podemos mencionar como ejemplo las complicaciones de la evaluación que se acentúan en la educación especial, debido a que en la educación regular la evaluación suele acompañarse de formas de valoración cualitativas que compensan las asignaciones numéricas que no suelen basarse en el desempeño, sino en la entrega de trabajos, asistencia y otros criterios arbitrarios que se refieren al dominio de temas abstractos que el estudiante puede aparentar dominar, mediante prácticas de memorización, repetición u otras similares sin demostrar su dominio o sin evaluar el desempeño. Sin embargo, en la educación especial, con mayor claridad es necesario que la evaluación se centre en el desempeño, en la ejecución o en la demostración directa de los aprendizajes adquiridos. Considerando que, dependiendo del grado de severidad o afectación motora, lingüística o cognitiva, las habilidades a desarrollar en cada individuo pueden ser de distinto tipo como básicas, adaptativas, sociales, académicas o culturales. Esta jerarquización de habilidades no suele abarcarse de forma explícita en la educación regular (debido a que los estudiantes no presentan alteraciones del desarrollo), donde a pesar de que pueden ser tomadas en cuenta, el énfasis se sitúa

en las habilidades académicas (Camacho, Cabrera et al., 2015; Camacho et al., 2022).

Por lo tanto, se hace necesaria la evaluación con mayor cuidado y objetividad en educación especial, cuando se busca mayor grado de precisión en el avance de los logros académicos, sociales o adaptativos que se requieran trabajar, esto permitirá que se vean con mayor claridad en dónde se encuentran las áreas que requieran mayor atención y también se permitirá que las metodologías puedan usarse y adecuarse con mayor probabilidad de éxito en otras personas o poblaciones con requerimientos similares.

Una forma de mejorar la evaluación y el registro de los logros y aprendizajes es mediante el uso de la estadística, la cual se puede ver como una herramienta de análisis matemático, que busca llegar a resultados imparciales mediante la estandarización de procedimientos de recopilación, ordenamiento y presentación de datos numéricos (Gil, 2003). Dentro del campo de la investigación y la práctica educativa, la estadística se considera un conjunto de métodos, técnicas y procedimientos para el manejo de datos, descripción, análisis e interpretación, que contribuyen al estudio científico de los problemas planteados en el ámbito de la educación y a la adquisición de conocimiento sobre las realidades educativas, a la toma de decisiones y a la mejora de la práctica desarrollada por los profesionales de la educación (Gil, 2003).

Así, tenemos que las intervenciones para regular el comportamiento de los estudiantes, así como aquellas orientadas al desarrollo académico pueden ser cuantificadas. Por ejemplo, se han empleado registros para evaluar la eficacia del principio de Premack para regular la conducta de niños de sección maternal (Homme et al., 1985), para evaluar enseñanza y retención de la lectura (Chumacero et al., 2010; Corey y Shamow, 1985), para evaluar el juego cooperativo en preescolar (Hart

et al., 1985), para el control de la agresión en niños de preescolar (Pinkston et al., 1985), para el estudio del lenguaje (Camacho et al., 2007; Camacho et al., 2008; Camacho, 2016; Gómez et al., 2012; Mejía y Camacho, 2007) para evaluar el aprovechamiento escolar de niños con Síndrome de Down (Dalton, et al., 1985), para la enseñanza de la escritura y la ortografía (Rayek y Nesselroad, 1985), para evaluar procedimientos correctivos para el desarrollo de la conducta verbal en niños con deficiencias del lenguaje (Sloane et al., 1985), para el tratamiento y atención de problemas de conducta en niños con Trastorno del Espectro Autista y otras alteraciones del neurodesarrollo (Camacho et al., 2015; Camacho et al., 2021; Wolf et al., 1972), para el establecimiento del control instruccional (Herrera y Martínez, 2015), para el estudio del desarrollo, la evaluación y los déficits de atención (Sánchez et al., 2015), para el estudio de los trastornos de aprendizaje mediante resonancia magnética funcional (Riveroll et al., 2015), entre muchos otros.

También se han propuesto programas de modificación de conducta orientados a la población con discapacidad y alteraciones del neurodesarrollo que se basan en el registro y cuantificación objetiva del desempeño, es decir, se suelen hacer registros sobre la emisión de alguna conducta antes y después de la aplicación de un plan de intervención para evaluar de una forma más clara los beneficios de la intervención realizada y el grado de impacto que puede tener en una población en particular (Camacho et al., 2021; Galindo et al., 1990; Ribes, 2002).

También se ha utilizado la cuantificación para el diseño de instrumentos y metodologías de trabajo que son aplicables a la educación especial desde distintos paradigmas o disciplinas, por ejemplo, para el diagnóstico y corrección de la actividad voluntaria en la edad preescolar y escolar que propusieron Salmina y Filimonova (2001), se utilizó el cálculo del coeficiente “t Student con una $p=0.01$ para verificar la diferencia significativa de los indicadores del diagnóstico del nivel

de desarrollo de la voluntariedad, antes y después de la impartición de algunas sesiones lúdicas, trabajo realizado entre los años 1991 y 1997 en la ciudad Sergiyev Posad de la provincia de Moscú, Rusia.

Si bien es necesario que la práctica del educador especial se apoye en técnicas, conceptos, modelos, paradigmas con evidencia científica, las cuales hayan sido obtenidas mediante la aplicación de procedimientos rigurosos, sistemáticos y objetivos, que permitan obtener y validar el conocimiento que será relevante en la instrumentación de los programas y actividades educativas especiales (Camacho et al., 2016; Camacho et al., 2015; Camacho et al., 2022; Cook et al., 2008), en el caso particular de México, esto no es así, ya que se ha reportado que existe una carencia de investigación sistemática con claros criterios de calidad y pertinencia que impacte al área de educación especial, en sus aspectos científicos y técnicos, y por ende, quienes elaboraban las leyes educativas no se basan en estos fundamentos teóricos y metodológicos sino en criterios políticos, administrativos, legislativos que desafortunadamente no incluyen los avances académicos y educativos con los que se cuentan a nivel mundial (Suro y Acle, 2014).

Entonces, la cuantificación y el uso de la estadística en educación especial nos permite mayor objetividad, mejor comparación de ejecuciones en evaluaciones iniciales, de seguimiento y finales, nos permite evitar interpretaciones personales sobre el desempeño del infante, nos ayuda a organizar y sistematizar las propuestas educativas y nos permite el diálogo profesional y científico con otras áreas o disciplinas. La falta de uso de la estadística en educación especial la puede hacer una profesión vulnerable, poco confiable y fácilmente reemplazable por carecer de práctica basada en evidencia (Acle- Tomasini, 2010).

Debemos tener claro que no hay interpretación profesional si se describen los logros, objetivos o evaluaciones con percepciones u opiniones subjetivas o

personales. Por ejemplo, en las distintas profesiones como medicina, ingeniería, odontología, generalmente se manejan distintos tipos de lenguajes, el lenguaje profesional-científico y el lenguaje técnico y coloquial u ordinario, el primero grupo de formas o tipos de lenguaje es para que los profesionales puedan dialogar entre sí justamente como expertos de sus disciplinas, con base en postulados y metodologías probadas y en uso, con términos que no les permitan interpretaciones distintas entre ellos y pueda haber un claro trabajo colaborativo que además a largo plazo, permita la mejora del trabajo que se realiza. Así se han ido descubriendo y replanteando técnicas, sustancias, procedimientos que mejoran la práctica profesional. El segundo grupo de tipos de lenguaje, el técnico y el coloquial u ordinario, permite que el profesional se comunique y se haga entender con el público en general, el usuario, paciente o educando, los familiares y padres de familia, pero no significa que con ese lenguaje el profesional pueda articular una metodología de trabajo formal y a disposición de otros colegas y que además permita el avance de su propia disciplina. Por lo que es necesario no confundir estos lenguajes para sistematizar la práctica educativa que se realiza en educación especial. Por ejemplo, en educación especial sería un error o una imprecisión profesional emitir resultados de evaluaciones o de logros académicos entre profesionales en términos tan coloquiales como decir que *“el alumno va mejorando”* o decir que *“ya se está portando mejor, o que el estudiante se autocontrola más”*, no hay problema si se emplean estos términos para comunicarse con los padres de familia, siempre y cuando el educador especial tenga detrás de ello, registros objetivos de las evaluaciones o del desempeño del estudiante que le permita comunicarse de una forma más profesional con otros colegas (Ribes, 2010).

Evaluaciones en educación especial: Uso de test y escalas estandarizadas

Los test y escalas estandarizadas componen una forma de observación y evaluación ordenada, cuya correcta aplicación permite al educador especial conocer el perfil cognitivo y madurativo del niño explorado, bajo condiciones de procedimientos estandarizados, con el objetivo de recabar información objetiva y fidedigna acerca de las necesidades específicas y globales del infante para la creación del plan de intervención (Mantilla, 2016).

Una de las principales críticas hacia el uso de las pruebas estandarizadas en el ámbito educativo, es que surge el riesgo de un reduccionismo instrumental y que el estudiante sea categorizado. Sin embargo, se ha descrito que la mayoría de los docentes en Latinoamérica no comprende los aspectos técnicos involucrados en el diseño de las evaluaciones nacionales e internacionales, ni los tipos de resultados que las mismas arrojan (Díaz y Ozuna, 2016; George, 2020). Razón por la cual, en el contexto educativo ha surgido la idea de que las pruebas estandarizadas son un “elemento agresivo que segrega socialmente”, obviamente, sin tomar en cuenta que las pruebas estandarizadas, se han implementado y mantenido por necesidades educativas y sociales de estandarización de contenidos, para servir como medio de verificación del aprendizaje logrado por los alumnos, para fortalecer las políticas educativas nacionales, así como para obtener indicadores que permitan comparar los sistemas educativos de varias regiones o incluso países, y desde luego, desde una visión economicista, para tener acceso a financiamiento internacional (Jiménez, 2016).

Por ejemplo, en un estudio realizado por Sánchez et al. (2010), se buscó identificar las competencias verbales en un grupo de 33 alumnos seleccionados al azar, de primer grado de una Escuela Primaria Oficial de Mérida, Yucatán; sin embargo, 3 de dichos alumnos fueron excluidos por causas de necesidades

educativas especiales (1 por sordera y 2 más por trastornos del desarrollo no especificados). A los 30 alumnos restantes se les aplicó una batería de pruebas en escalas verbales, la cual incluía el WISC-IV, SAGES-2:KA y PDNL. Una vez obtenidos los resultados, los investigadores lograron organizar a los estudiantes en 3 grupos, de acuerdo con su competencia verbal: 8 alumnos (26%) con alta competencia verbal, 13 (43%) con nivel medio y 9 (30%) con baja competencia verbal. De la misma manera se evidenció que 5 casos de 9 alumnos con baja competencia verbal no presentaron capacidad lectora atribuible y en los 4 restantes, existió la *sospecha* de que exista algún trastorno del aprendizaje. Los resultados contribuyeron para identificar que, a nivel nacional, existen demandas excesivas al maestro de primaria, quien es el que hace frente a una población heterogénea, con diversas necesidades educativas especiales y regulares, acompañado de diversas carencias, como lo son los recursos limitados, y la ausencia de procedimientos de detección y diagnóstico. Además de demostrar la inmensa necesidad de proveer servicios de educación especial a un número representativo de estudiantes de las escuelas públicas mexicanas, que requieren de diagnósticos precisos para atender su problemática (Sánchez, 2011). Este estudio es un ejemplo de la objetividad y precisión que brinda la aplicación de pruebas estandarizadas utilizadas para evidenciar la problemática existente en el ámbito educativo.

A nuestro parecer, el acercamiento de la educación especial a la estadística es claramente necesario, tanto para educadores especiales en formación, como aquellos que ya se encuentran en la práctica docente, no solo por el reconocimiento, consolidación, expansión y sistematización del trabajo que se hace en la propia disciplina, sino también porque repercute en beneficio del trabajo que se hace con los alumnos y con la población de educación especial en general.

Conclusiones y comentarios finales.

El movimiento actual de inclusión educativa que implica construir espacios de convivencia, comunicación y desarrollo económico, educativo, social y cultural, será posible solo si se consideran los distintos factores implicados, los cuales, además de la indispensable e ineludible necesidad social de respeto a las diferencias y a la diversidad, implica la consecución de logros de aprendizaje efectivos, objetivos, demostrables que favorezcan el desarrollo individual, académico y social de las personas con discapacidad que son sujetos de educación especial (puede haber personas con discapacidad que no requieran educación especial), de personas con alteraciones del neurodesarrollo y con aptitudes sobresalientes.

Otro aspecto muy relevante para el logro de la inclusión educativa es la necesidad de integración de profesionales con distinta formación, con distintos fundamentos pero con un objetivo común, ya que ésta ha sido una de las formas en las que la educación se ha transformado a lo largo de su existencia, no solo porque ha permitido que las distintas profesiones y disciplinas interactúen en favor de la educación, sino porque esta interacción también provoca el cambio, reorganización y reflexión de la labor educativa.

Así, la transformación, profesionalización y aplicación científica es parte inherente del proceso educativo e indispensable para lograr la inclusión educativa, desde luego sin caer en el *cientificismo*³ ni en un *materialismo científico*⁴ (promovido por algunas prácticas *pseudocientíficas* que establecen los organismos

³ Postura dogmática y reduccionista que afirma que la ciencia empírica es la única que puede validar el conocimiento, negando todo aquel que no provenga de una práctica científica.

⁴ Cualquier forma o expresión de materialismo que considere a las ciencias naturales como el paradigma básico para estudiar e interpretar los fenómenos o problemas de diversas disciplinas.

reguladores de la ciencia a nivel mundial), que en lugar de beneficiar a la formación educativa la termine distorsionando de forma inadvertida. Sin embargo, cuidando lo anteriormente señalado, debemos tener claro que las observaciones, medidas, evaluaciones, registros, aplicación de instrumentos y reportes psicopedagógicos pueden ser presentados de forma profesional en la medida que incluyan conceptos, metodologías, instrumentos y formas de medición adecuadas al desarrollo de los aprendizajes, habilidades, competencias, que se hayan trabajado con los estudiantes.

Es importante señalar que el uso de la estadística nos permite lo anterior, pero también debemos tener claro que la estadística únicamente representará los datos que el docente proporcione, la estadística en este sentido no es infalible, a pesar de ser una ciencia exacta, no representa a plenitud un fenómeno social o educativo, sino que representará los datos que se decida utilizar. En un ejemplo muy básico, podríamos decir que podemos obtener el promedio objetivo de un grupo o de un individuo en distintas ejecuciones, antes y después de la aplicación de algún programa, metodología o proceso de enseñanza, el promedio podrá mostrar diferencias en la ejecución al pasar de un 30% a un 85% de respuestas correctas del grupo o del individuo. Sin embargo, esto no significa que la persona o el grupo en general no puedan, en otra evaluación obtener un resultado diferente, o que a pesar de haber obtenido un porcentaje alto no pueda deteriorarse la ejecución o que incluso ese porcentaje alto no se deba de forma directa y exclusiva al programa en el que participó, sino que sea producto de otras circunstancias o variables que pudieron contribuir en la mejora de dicha ejecución. En otro ejemplo más práctico, tal vez hemos advertido que cada fin de sexenio –o a veces cada informe de gobierno- los presidentes reportan, incluso con apoyo estadístico, que han reducido la pobreza, sin embargo, socialmente dicha conclusión podría no ser totalmente compartida por el resto de la población, las razones para ello son diversas, pero aquí solo

mencionaremos lo que a la estadística le corresponde. En primer lugar la estadística no interpreta las definiciones ni los modelos empleados para la recogida de datos, en el ejemplo que nos ocupa, normalmente se utiliza un modelo multidimensional de la pobreza, lo que facilita que se asignen datos de registros arbitrarios con los cuales se identifica a la población en ese estado, así tenemos que, de acuerdo con reportes del CONEVAL en el 2019 el porcentaje de población en pobreza extrema pasó de 11% en 2008 a 7.4% en 2018, de hecho 24 entidades federativas tuvieron reducción importante de población en situación de pobreza extrema y los estados que presentaron la mayor reducción fueron Aguascalientes, Hidalgo y Tlaxcala. A pesar de que, para asegurar la continuidad en la medición de la pobreza se procuran mantener los criterios metodológicos de acuerdo con los “Lineamientos y criterios generales para la definición, identificación y medición de la pobreza”, en la práctica la subjetividad de dichos lineamientos permite adecuaciones que no salen de dichos criterios. Así que no debemos considerar que la disminución de la pobreza implica que la gente tiene más dinero o mejores condiciones económicas, ya que otros factores demográficos pueden estar afectando la medición y los reportes estadísticos que se presentan, como las defunciones, nacimientos, la implementación de seguros populares, la migración, entre muchas otras.

Los ejemplos anteriores buscan dar mayor claridad y subrayar el punto de que no es el mero uso de la estadística el que mejorará nuestras evaluaciones y mediciones, sino tenemos claros los criterios, teoría, metodología y variables que pueden afectar la medición o aplicación de una prueba, entonces la estadística solo nos mostrará el tratamiento de los datos recabados, pero no es responsable de que los criterios y variables no hayan sido incluidas o del grado de subjetividad de las mismas, cuidando también no caer en una “tendencia desarrollista” que pone el foco

de interés en lo económico y lo administrativo, dejando de lado el auténtico desarrollo educativo y social (Bello-Domínguez, 2011).

Por lo tanto, debemos tener mucha claridad y mucho cuidado en primer lugar en nuestras definiciones, conceptos, variables, metodología a emplear en un proceso o evaluación educativa y por otro en elegir el tipo de prueba o medición estadística más adecuado para ello, asunto que se desarrollará a lo largo de la presente obra.

Finalmente, replantearse la forma de entender, de practicar y de consolidar a la educación especial implica transformar la educación en general y hacer de la inclusión algo más concreto que una realidad deseada o una utopía sobreestimada. Debemos abandonar de una vez por todas el enfoque “esencialista” y “estático” de la educación que la mira como un proceso determinado por la naturaleza del individuo, así como por las características de su pensamiento producto de su convivencia social y cultural y considera que la única forma de educar es a través de un paradigma institucional cuya función esencial es transmitir, informar, memorizar o imitar acciones técnicas (Ribes, 2008). La educación es un proceso complejo pero intencional, planeado, con objetivos claros, dinámico, eminentemente social y que es ella la que transforma la sociedad a su vez que la sociedad es quién le retroalimenta sobre las necesidades de la educación. La educación especial requiere el abandono de esas posturas que se conforman con una educación informal que no desarrolla las habilidades y competencias necesarias para la inclusión educativa y social especialmente de quienes se encuentran en situación de discapacidad intelectual y trastornos del neurodesarrollo severos (Camacho et al., 2015).

En conclusión, algunas acciones específicas que deben acompañar la transformación de la educación especial y el uso de la estadística como herramienta de apoyo pueden situarse en:

1. Desarrollo de aprendizajes específicos de acuerdo con una lógica de desarrollo psicológico, social y biológico.
2. Diferenciar a las personas con desarrollo lingüístico de aquellos que no tendrán dicho desarrollo, ya que la educación de ambos es posible, pero con técnicas, metodologías y supuestos distintos (Camacho, 2013).
3. La especialización y actualización deben incidir directamente sobre las cuestiones teóricas y metodológicas de aplicación inmediata, es decir, deben ser un tipo de “asesorías” sobre los problemas teóricos y metodológicos que tengan los profesores de educación especial y sobre casos supuestos o conceptos abstractos que no pueden llevar a su ejercicio de forma inmediata.
4. Diferenciar y reconocer la necesidad de categorizar y jerarquizar el desarrollo de habilidades y competencias, especialmente en los casos identificados como severos en educación especial, ya que ello permitirá la mejor consecución y avances en la tarea educativa (Camacho et al., 2015; Camacho et al., 2016).
5. Re-pensar la escuela, re-organizar la escuela, re-inventar la escuela. Las escuelas no son instituciones inertes, o dependientes de criterios políticos, antes que todo, éstas se deben transformar ante las demandas educativas y sociales y deben ser ellas quienes propongan y cuestionen las directrices políticas que las regulan, es decir, la relación de la escuela con los organismos federales y estatales que la regulan deben tener una relación bidireccional, pero con una visión transformadora.
6. Re-plantearse la relación y colaboración con padres de familia, especialmente en educación especial, ya que ello implica adecuación de rutinas, de responsabilidades, de espacios físicos y de otros aspectos que se hacen necesarios para la consecución de aprendizajes en favor de la inclusión educativa (no simulación de la misma) y social efectiva y auténtica.

7. Identificar de forma clara a las personas con discapacidad y alteraciones del neurodesarrollo que podrán continuar una formación académica, para proporcionales herramientas útiles en este camino y enfatizar los conocimientos previos necesarios. Pero también identificar a quienes por la severidad de la condición y las limitaciones del lenguaje, no podrán continuar con dicho proceso académico para centrarse en el desarrollo de habilidades básicas, adaptativas y pre-sociales que les implique una mejora en su calidad de vida (Camacho, 2013; Camacho et al., 2022).

Es importante considerar los puntos anteriores como aspectos mínimos a considerar en la tarea educativa, para que de esta forma, el uso de la estadística llegue a colaborar a un fin educativo auténtico, de lo contrario, si se emplea la estadística con los vicios y problemas teóricos, metodológicos, didácticos y de evaluación que actualmente pueden identificarse en la educación especial, la estadística solo llegará a reforzar o dar mayor credibilidad a algo que no es claramente un logro educativo ni una metodología adecuada. Recordemos que se puedan validar estadísticamente falacias, supuestos y ambigüedades didácticas, podríamos decir, como un ejemplo exagerado, que estadísticamente podríamos encontrar correlación entre el desempeño en matemáticas y cantar una canción infantil en todas las escuelas al inicio y término de cada día de clases, y aunque estadísticamente se pueden trabajar dichos fenómenos una vez que se le asignan números, en realidad podríamos ver que no existe razón teórica o metodológica que pueda soportar dicha correlación. En suma, el uso de mejores herramientas en educación, como lo es la estadística, requiere la mejora de los procesos educativos en general, para no alimentar simulaciones o falsas relaciones de mejora educativa.

Referencias bibliográficas

- Acle, G. (2011). ¿Por qué la Educación Especial es Especial? *Integra2 Revista Electrónica de Educación Especial y Familia*, 2(2), 5-18.
https://www.fcdh.uatx.mx:80/media/integra2/numero_completo/numero_completo_Volumen_2_Numero_2_julio-diciembre_2011.pdf
- Acle-Tomasini, G. (2010). Prácticas basadas en evidencia científica: asignatura pendiente en educación especial. *Revista Mexicana de Psicología Educativa*, 1(1), 13-20.
https://www.researchgate.net/publication/259461561_Practicas_basadas_en_evidencia_cientifica_asignatura_pendiente_en_educacion_especial
- Bello-Domínguez, J. (2011). ¿Escuelas inclusivas en una sociedad excluyente? En J. Bello-Domínguez (Ed.). *Educación Inclusiva, una aproximación a la utopía* (pp. 21-44). Castellanos Editores.
- Camacho, J.A. (2013). La deuda pendiente de la educación especial. *Integra2 Revista Electrónica de Educación Especial y Familia*, 4(2), 25-33.
<https://integra2.fcdh.uatx.mx>
- Camacho, J.A. (2016). La atención básica y la importancia del lenguaje en la educación especial. En J.A. Camacho, C.B. Mendoza, A. Saldivar, y J. Guzmán, (Eds.), *Educación especial, de la atención básica a la inclusión educativa y social* (pp. 73-102). Castellanos Editores y Universidad Autónoma de Tlaxcala.
- Camacho, J.A., Almanza, M.L., y Romero, R.A. (Eds.). (2015). *Neurociencia y Educación Especial: Conceptos, procesos y principios básicos*. Universidad de Guadalajara-Universidad Autónoma de Tlaxcala.
- Camacho, J.A., Gómez, A.D., Meraz, E., Zepeta, E., y Cabrera, F. (Eds.). (2022). *Planteamientos conceptuales y de atención psicológica dirigidos a la educación especial*. Universidad Veracruzana-Universidad Autónoma de Tlaxcala.
<https://drive.google.com/file/d/1bi7xmlWPvzuTnyR-dkWeCsYvxnBOMirS/view>
- Camacho, J.A., Irigoyen J., Gómez, D., Jiménez M. y Acuña K. (2007). Adquisición y transferencia de modos lingüísticos en tareas de discriminación condicional sin

- retroalimentación reactiva. *Enseñanza e Investigación en Psicología*. 12 (1), 79-91. <http://www.redalyc.org/pdf/292/29212106.pdf>
- Camacho, J.A., Mejía, M., Rivera, A., Corona, I., Camacho, A., y Gómez, D. (2008). Variación del aprendizaje al emplear distintos modos del lenguaje en una interacción académica en la asignatura de biología. *Enseñanza e Investigación en Psicología*, 13(2), 373-385. <http://www.redalyc.org/pdf/292/29213211.pdf>
- Camacho, J.A., Mendoza, C.B., Saldivar, A. y Guzmán, J. (Eds.). (2016). *Educación especial, de la atención básica a la inclusión educativa y social*. Castellanos Editores y Universidad Autónoma de Tlaxcala.
- Camacho, J.A., Meraz, E., Cortés, A. y Cabrera, F. (2015). Análisis de la conducta y educación especial, una relación fracturada. En F. Cabrera, O. Zamora, H. Martínez, P. Covarrubias y V. Orduña (Eds.), *Estudios sobre comportamiento y aplicaciones Vol. IV*. (pp. 225-249). Universidad Nacional Autónoma de México-Universidad de Guadalajara.
- Camacho, J.A., Navarrete, L., Aguilar, F.J., Tecamachaltzi, M.B. y Cabrera F. (2021). Aprendizaje discriminativo y RDI como alternativa para la reducción de conductas estereotipadas motoras en el Trastorno del Espectro Autista. *Revista Electrónica de Psicología Iztacala*, 24(3), 898-923.
- Camacho. J.A., Rojas, M. y Cabrera, F. (2015). Control del estímulo y modificación de patrones de comportamiento en la educación especial (pp. 219-246). En J.A. Camacho, M.L. Almanza y R.A. Romero. (Eds.). *Neurociencia y Educación Especial; conceptos, procesos y principios básicos*. Universidad de Guadalajara-Universidad Autónoma de Tlaxcala.
- Chumacero, N., Camacho, J.A. y Irigoyen, J.J. (2010). La enseñanza inicial de lectura dentro del aula regular: niveles de desempeño funcional que se promueven en niños que presentan dificultades para su adquisición. *Integra2, Revista Electrónica de Educación Especial y Familia*, 1(1), 25-47. En Línea: <https://integra2.fcdh.uatx.mx>

- Comisión de Política Gubernamental en Materia de Derechos Humanos. (2016). *Glosario de términos sobre discapacidad*. http://www.semar.gob.mx/derechos_humanos/glosario_terminos_discapacidad.pdf
- Consejo Nacional de Evaluación de la Política de Desarrollo Social [CONEVAL]. (s. f.). *Pobreza en México*. <https://www.coneval.org.mx/Medicion/Paginas/PobrezaInicio.aspx>
- Cook, B.G., Tankersley, M., Cook, L., y Landrum, T. J. (2008). Evidence based practices in special education: Some practical considerations. *Intervention in School and Clinic*, 44(2), 69-75. <https://doi.org/10.1177/1053451208321452>
- Corey, J.R., y Shamow, J. (1985). Efectos del desvanecimiento en la adquisición y retención de la lectura oral. En S.W. Bijou, y E. Rayek. (Eds.), *Análisis conductual aplicado a la instrucción* (pp. 108- 115). Trillas.
- Dalton, A.J., Rubino, C.A., y Hislop, M.W. (1985). Algunos efectos de las recompensas con fichas, con respecto al aprovechamiento escolar, en niños con síndrome de Down. En S.W. Bijou, y E. Rayek. (Eds.), *Análisis conductual aplicado a la instrucción* (pp. 422-436). Trillas.
- Díaz, K.M., y Osuna, C. (2016). Las evaluaciones estandarizadas del aprendizaje y la mejora de la calidad educativa. *Temas de Educación*, 22(1), 131-146. <https://revistas.userena.cl/index.php/teeducacion/article/download/741/828/1935#:~:text=Se%20cre%C3%B3%20en%201999%20con,realiza%20evaluaciones%20externas%20e%20independientes>
- Galindo, E., Bernal, T., Hinojosa, G., Galguera, M. I., Taracena, E., y Padilla, F. (1990). *Modificación de conducta en la educación especial: Diagnóstico y programas* (2.^a ed.). Trillas.
- George, C.E. (2020). Pruebas estandarizadas y calidad de la educación en México, sexenio 2012-2018. *Revista Universidad y Sociedad*, 12(4), 418-425. <http://scielo.sld.cu/pdf/rus/v12n4/2218-3620-rus-12-04-418.pdf>
- Gil, F.J. (2003). La estadística en la investigación educativa. *Revista de Investigación Educativa*, 21(1), 231-248. <https://revistas.um.es/rie/article/view/99191/94791>

- Gómez, A.D., Camacho, J.A., Zepeta, E., Arias, S., y Pérez, M. (2012). Entrenamiento en los modos reactivos y su efecto en el modo activo escribir. *IPyE: Psicología y Educación*, 6(12), 30-45.
- Hart, B.M., Reynolds N.J., Baer, D.M., Brawley, E.R., y Harris, F.R. (1985). Los efectos del reforzamiento social contingente y del no contingente sobre el juego cooperativo de una niña en edad preescolar. En S.W. Bijou, y E. Rayek. (Eds.), *Análisis conductual aplicado a la instrucción* (pp. 154- 161). Trillas.
- Herrera, D., y Martínez., H. (2015). Factores que intervienen en el establecimiento del control instruccional. En J.A. Camacho, M.L. Almanza, y R.A. Romero. (Eds.) *Neurociencia y Educación Especial; Conceptos, procesos y principios básicos* (pp. 247- 264). Universidad de Guadalajara- Universidad Autónoma de Tlaxcala.
- Homme, L.E., deBaca, P.C., Devine, J. V., teinhorst, R., y Rickert, E.J. (1985). El empleo del principio de Premack en el control de la conducta de niños de sección maternal. En S.W. Bijou, y E. Rayek. (Eds.) *Análisis conductual aplicado a la instrucción* (pp. 70- 71). Trillas.
- Jiménez, J. A. (2016). El Papel de la Evaluación a Gran Escala como Política de Rendición de Cuentas en el Sistema Educativo Mexicano. *Revista Iberoamericana de Evaluación Educativa*, 9(1), 109-126. <https://revistas.uam.es/riee/article/view/4016>
- Ley General de Educación. (2019, 30 de septiembre). Cámara de Diputados del H. Congreso de la Unión. Diario Oficial de la Federación (s/n). <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGE.pdf>
- Mantilla, J.P. (2016). *Detección, diagnóstico e intervención temprana en niños menores de tres años de alto riesgo biológico y ambiental pertenecientes a la comunidad indígena de Salasaca provincia de Tungurahua*. [Tesis doctoral, Universidad Complutense de Madrid]. <https://eprints.ucm.es/id/eprint/40397/1/T38102.pdf>
- Mejía, M., y Camacho, J.A. (2007). Variación de aprendizaje al emplear distintos modos del lenguaje en una interacción académica. *Enseñanza e Investigación en Psicología*. 12(2), 277-289. <http://www.redalyc.org/pdf/292/29212204.pdf>

- Pinkston, E.M., Reese, N.M., LeBlanc, J.M., y Baer, D.M. (1985). El control independiente de la agresión en un niño de edad preescolar y la interacción con sus compañeros, mediante la atención contingente del maestro. En S.W. Bijou, y E. Rayek. (Eds.) *Análisis conductual aplicado a la instrucción* (pp. 189- 200). Trillas.
- Rayek, E. y Nesselroad, E. (1985). Aplicación de los principios conductuales a la enseñanza de la escritura, la ortografía y la composición. En S.W. Bijou, y E. Rayek. (Eds.) *Análisis conductual aplicado a la instrucción* (pp. 502- 516). Trillas.
- Ribes, E. (2002). *Técnicas De Modificación De Conducta: Su aplicación al retardo en el desarrollo* (2.^a ed.). Trillas.
- Ribes, E. (2008). Educación básica, desarrollo psicológico y planeación de consecuencias. *Revista Mexicana de Psicología*, 25(2), 193-207.
- Ribes, E. (2010). Lenguaje ordinario y lenguaje técnico: un proyecto de currículo universitario para la psicología. *Revista Mexicana de Psicología*, 27(1), 55-64.
- Riveroll, R., Pasaye, E., y Carvalho, M.C. (2015). Estudio de los trastornos en el aprendizaje mediante resonancia magnética funcional. En J.A. Camacho, M.L. Almanza, y R.A. Romero. (Eds.) *Neurociencia y Educación Especial; Conceptos, procesos y principios básicos* (pp. 321- 346). Universidad de Guadalajara- Universidad Autónoma de Tlaxcala.
- Salmina, N.G., y Filimonova, O. G. (2001). *Diagnóstico y corrección de la actividad voluntaria en la edad preescolar y escolar* (1.^a ed.). Universidad Autónoma de Tlaxcala.
- Sánchez, A. (2011). Comentarios a: ¿Por qué la Educación Especial es Especial? *Integra2, Revista Electrónica de Educación Especial y Familia*, 2(2), 19-56. https://www.fcdh.uatx.mx:80/media/integra2/numero_completo/numero_completo_Volumen_2_Numero_2_julio-diciembre_2011.pdf
- Sánchez, P., Gómez, M., y Guzmán, S. (2010). Competencias verbales de alumnos en un aula típica de una primaria pública de Yucatán. *Educación y Ciencia, Cuarta Época*. 2(37), 133-144.

<https://www.researchgate.net/publication/277214576> Competencias verbales de alumnos en un aula típica de una primaria pública de Yucatan

- Sánchez, S. M., Reyes, A., y Sánchez, J. (2015). Atención: desarrollo, evaluación y déficits. En J. A. Camacho, M. L. Almanza, y R. A. Romero. (Eds.) *Neurociencia y Educación Especial; Conceptos, procesos y principios básicos* (pp. 321- 346). Universidad de Guadalajara- Universidad Autónoma de Tlaxcala.
- Sloane, H.N., Johnston, M.K., y Harris, F.R. (1985). Procedimientos correctivos para enseñar conducta verbal a niños pequeños con defectos o deficiencias de lenguaje. En S.W. Bijou, y E. Rayek. (Eds.) *Análisis conductual aplicado a la instrucción* (pp. 559- 581). Trillas.
- Suro, J., y Acle, G. (2014). Investigación en Educación Especial (2002-2011): logros y desafíos. *Diálogos sobre educación. Temas actuales en investigación educativa*, 5(8), 1-6. <http://dialogossobreeducacion.cucsh.udg.mx/index.php/DSE/article/view/316>
- Wolf, M.M., Risley, T., y Mees, H. (1972). Aplicación de los procedimientos de condicionamiento operante a los problemas de conducta de un niño autista. En R. Ulrich, T. Stachnik, y J. Mabry. (Eds.) *Control de la conducta Humana* (pp. 332- 342). Trillas.

EL MUESTREO EN INVESTIGACIÓN RELACIONADA CON LA EDUCACIÓN ESPECIAL

Francisco Javier Aguilar Guevara [‡]

Resumen

En el presente capítulo se reconoce a la Educación Especial como una interdisciplina que, en tanto tal, se dedica a la resolución de la problemática de relevancia social que tiene que ver con la provisión de educación a personas con discapacidad, trastorno y/o aptitudes sobresalientes. Se señala que la forma en la que la Educación Especial cubre su encargo social es al amparo del conocimiento desarrollado en disciplinas o multidisciplinas científicas. En este tenor, se sostiene que la labor científica del Educador Especial puede darse únicamente adscribiéndose de forma absoluta a una de esas disciplinas o multidisciplinas científicas de las cuales se nutre. Dada la diversidad de éstas, el abanico metodológico disponible para el Educador Especial es muy amplio. Es por ello que en un capítulo que verse acerca de los aspectos a considerar en la selección de la muestra o de los participantes, una de las pocas formas de abarcarles es atendiendo a la distinción entre metodología cuantitativa y cualitativa. Por tanto, en el capítulo se pretende cubrir los aspectos esenciales en relación con el muestreo, entendido éste como el proceso de selección e incorporación de sujetos o participantes a estudios de corte cuantitativo; así mismo se tratan aspectos asociados a la selección e inclusión de participantes en estudios cualitativos.

Palabras clave: Muestreo, Selección de participantes, Educación Especial

[‡] Facultad de Ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Carretera Tlaxcala-Puebla Km. 1.5. Tlaxcala, Tlax. fjaguilarg_fcdh@uatx.mx

Introducción

De acuerdo con autores como Ribes (2005; 2011) y Sánchez-Sosa (2008), se pueden identificar dos grandes tipos de disciplinas. Por un lado se encuentran las disciplinas científicas, cuyas prácticas se concentran en la generación de conocimiento y por tanto requieren de un objeto de estudio propio que les distinga de otras disciplinas científicas. Por otro lado se encuentran las disciplinas profesionales, o mejor dicho, interdisciplinas que dirigen sus esfuerzos a la resolución de problemáticas específicas, entendidas como encargos sociales. En consecuencia, a las disciplinas científicas no les es posible la resolución de problemáticas sociales específicas a partir del conocimiento que generan, en tanto no tienen un encargo social particular que se relacione de forma directa con su objeto de estudio, y a las interdisciplinas profesionales no les es posible generar conocimiento propio de su disciplina, debido a que no cuentan con un objeto de estudio constitutivo.

Únicamente con fines de clarificación del argumento anterior, es necesario mencionar que una de las principales razones que posibilitan que las ciencias puedan generar conocimiento, pero no resolver problemas de carácter social, es el lenguaje técnico que emplean. Para que el análisis empírico-conceptual de un fenómeno sea posible, es necesario el uso de un lenguaje que abstraiga relaciones, propiedades y dimensiones de entidades y acontecimientos. Es este lenguaje abstracto el que permite segmentar cosas del mundo, y analizarlas, y el que hace imposible encontrar eso abstracto como una situación que requiere atención por parte de un conjunto de personas para su convivencia social.

Un ejemplo claro es la Velocidad promedio. La fórmula de la Velocidad promedio es:

$$V_{av} = \frac{X_f - X_i}{t_f - t_i}$$

En la que X_f es la posición final, X_i es la posición inicial, t_f es el tiempo final y t_i es el tiempo inicial. Como se puede observar, todo en esta fórmula son abstracciones y relaciones específicas de las mismas, que aplican a cualquier conjunto de valores específicos posibles. Así, nadie comenta con su mejor amiga o con sus hermanos que tiene un problema de t_f o incluso de V_{av} que le quita el sueño. Tampoco hay alguien que se preocupe por saber cuál es el profesional que puede generar modificaciones en X_i . Sin embargo, seguro son conceptos cuya relación bien puede ser tema de conversación diaria en el Instituto Nacional de Aeronáutica Civil de Paraguay. Por otro lado, también se puede encontrar profesionales que cotidianamente requieran incluir valores de velocidad específicos, es decir no abstractos, entre muchas otras cosas que tienen que ver con disciplinas distintas a la física, para la solución de problemas sociales. Tal es el caso del Jefe de la Unidad Departamental de Programación de la Secretaría de Movilidad del Gobierno de la Ciudad de México, a quien le interesará que se reduzcan los accidentes viales. Por supuesto éste último tampoco pretende hacer investigación a partir de la modificación ordenada de parámetros de la Velocidad promedio o de parámetros asociados a la misma.

Bajo esta lógica, la Educación Especial es una interdisciplina profesional, que tiene por encargo social servir educativamente a niñas y niños quienes por su discapacidad, trastorno o aptitudes sobresalientes se enfrentan con mayor frecuencia o en mayor medida con Barreras para el Aprendizaje y la Participación (BAP) impuestas por su medio social (Ver Aguilar y Camacho, en prensa). Justamente por ser una interdisciplina, las y los educadores especiales, echan mano del conocimiento generado desde diversas disciplinas o multidisciplinas científicas para poder resolver los diversos casos específicos de problemas educativos que se les

presentan. Incluso, dependiendo del caso, el tipo de conocimiento disciplinar empleado, será distinto.

Actualmente, las y los educadores especiales atienden a poblaciones muy diversas, que lógicamente tienen características, situaciones y realidades altamente variopintas. Es por ello que no se requiere de exactamente el mismo conjunto de saberes cuando se atiende a un estudiante con algún trastorno del neurodesarrollo, que cuando se atiende a un estudiante con discapacidad, auditiva por ejemplo, o a uno con aptitudes sobresalientes; aunque cierto conjunto de habilidades a desplegar sí que se comparten en la atención de todos los casos. Así mismo, se debe considerar el repertorio de habilidades con el que cada uno de estos estudiantes cuenta, el nivel de apoyo profesional y familiar que ha tenido, su contexto familiar, entre otras muchas variables, para decidir cuáles aportaciones científicas son las que se han de emplear o se han de tener en consideración.

Esto no requiere que las y los educadores especiales deban ser especialistas en todas las disciplinas científicas de las que se valen. Una tarea de ese tipo es impracticable. Lo que las y los educadores especiales requieren en la mayoría de los casos son conocimientos desprofesionalizados. Como lo indica Ribes (2005), la desprofesionalización del conocimiento científico para su uso eficaz por parte de profesionistas implica dos etapas. La primera implica adecuar el conocimiento científico a lenguajes interfase con los que el primero pueda ser vinculado a referentes lingüísticos y criterios de uso propios, en este caso de las y los educadores especiales. El segundo paso consiste en la reorganización del conocimiento científico abstracto a formas sintéticas, concretas y específicas de aplicación.

En este punto y desde esta perspectiva, probablemente la pregunta sea: ¿Entonces las y los educadores especiales están vetados del quehacer científico? La respuesta es no, desde dos perspectivas mínimamente. Una de ellas tiene que ver con

el concepto de investigación traslacional y la otra tiene que ver con un enfoque que ha cobrado trascendencia en las últimas décadas, que es el de la intervención basada en evidencia.

En lo que tiene que ver con la investigación traslacional, Aguilar, et al., (2020) mencionan que no solo se da el proceso más conocido, y del que se ha venido hablando hasta el momento en el presente documento, en el que el conocimiento científico se puede aplicar a situaciones específicas y así solucionar algunas problemáticas que tienen sentido en contextos sociales determinados; también se suele dar el proceso en el que en los escenarios aplicados surgen necesidades o retos para los cuales el conocimiento científico disponible muestra ser erróneo o insuficiente y requiere nuevas formas de relacionar las variables identificadas o requiere la identificación de nuevas variables que se tornan relevantes o esenciales. Este segundo tipo de proceso es en el que las y los educadores especiales pueden proceder de forma ordenada y metodológicamente pertinente y generar conocimiento que sea útil en su misma interdisciplina o que permita la búsqueda y solución de nuevas preguntas de investigación científico-disciplinar. Cabe mencionar que si lo que ocurre es la búsqueda y solución de nuevas preguntas de investigación en el marco de una disciplina científica, las y los mismos educadores especiales lo pueden llegar a hacer pero totalmente inmersos en los aspectos filosóficos, teóricos y metodológicos de la disciplina científica en cuestión.

En lo que toca a la intervención basada en evidencia, Cook, et al., (2009), Cook y Cothren (2011) y Odom, et al., (2005), mencionan que en Educación Especial se ha adoptado recientemente (desde hace 10 o 15 años, aproximadamente) una nueva lógica o forma de intervención que surgió en la disciplina de la Medicina, la Intervención Basada en Evidencia. Esta forma de intervención se caracteriza por ser opuesta a la intervención por sentido común o a la intervención basada en

costumbres o tradiciones gremiales, en tanto la evidencia a la que hace referencia es evidencia científica. En esta línea, los autores marcan una serie de requisitos que deben cumplir los estudios que se identifican como evidencia. Uno de los requisitos principales es el tipo de diseños empleados, que va de la mano con el requisito de control de variables. Los autores mencionados recomiendan los diseños de grupo, los diseños de sujeto único y los diseños de ensayos aleatorizados, por el control de variables independientes que se logra al amparo de dichos diseños de investigación. Adicionalmente mencionan que los estudios deben contar con la operacionalización de los procesos, la replicabilidad de procedimientos, especificación de espacios o lugares en los que tales procedimientos se han de implementar, la especificación de la población a la que van dirigidos, resultados esperados y las bases teórico-conceptuales.

La parte más importante de esta perspectiva de intervención, en el contexto de este escrito, es que, para lograr contar con la evidencia necesaria para realizar la Intervención Basada en Evidencia, es necesario probar las técnicas y procedimientos generados en el marco de disciplinas científicas, no solo en escenarios de alto control experimental y casi nula validez ecológica, sino también en escenarios naturales con menor control, pero mayor validez ecológica. Es en este segundo tipo de estudios en el que las y los educadores especiales pueden realizar contribuciones científicas de alta trascendencia.

A partir de lo anterior se puede recuperar una serie de puntos o elementos relevantes para los propósitos de este escrito. Uno es que la relación de colaboración y de desarrollo mutuo entre las disciplinas científicas y las disciplinas profesionales o interdisciplinas es por demás constante. Otro punto es que, independientemente de que el contacto entre estos dos tipos de disciplinas sea ininterrumpido, cada uno de los tipos de disciplina conserva sus prácticas y propósitos. De ello se genera un punto

adicional: siempre que las y los educadores especiales entran en contacto con la investigación, lo hacen con la intención de proveer un mejor servicio a sus usuarios y lo hacen en el marco de criterios teórico-metodológicos de alguna disciplina científica particular. Dicho de otra forma, *la Educación Especial no cuenta, ni contará, con criterios teóricos ni metodológicos de investigación que le son propios.*

Si se parte, entonces, de que la Educación Especial por ser una interdisciplina que resuelve problemas relevantes socialmente, no cuenta con procesos ni criterios metodológicos exclusivos para generar conocimiento, cuando se habla de cualquier proceso científico-metodológico para generar conocimiento relacionado con la Educación Especial, como el muestreo, se debe atender a los criterios de las disciplinas científicas con las que la Educación Especial tiene interacción. Por supuesto, no con todas al mismo tiempo. Dependiendo del conocimiento que se pretenda generar, se atenderá a los criterios de una disciplina científica u otra.

Cabe mencionar que la cantidad de disciplinas con las que la Educación Especial se relaciona es muy amplia. Camacho, et al., (2015) mencionan varias disciplinas, multidisciplinas e interdisciplinas, unas científicas y otras no, que de una forma u otra, y en grados distintos, coadyuvan en las tareas de las y los educadores especiales, a saber: la psicología, la neuropsicología, la medicina, la terapia física, la biología genética, la neurología, la pedagogía, distintas posturas teóricas psicológicas, la filosofía, el derecho y la antropología. Estos mismos autores subrayan que la lista que aportan no es exhaustiva. Sin embargo, lo más relevante para este capítulo, cuyo tema principal es *muestreo*, es que no todas las disciplinas científicas mencionadas por los autores tienen una metodología compartida.

Aun cuando se pueden mencionar múltiples diferencias y en distintos grados entre las metodologías empleadas por las distintas ciencias con las que se relaciona la Educación Especial, para buscar conseguir la mayor claridad posible en la

exposición de ideas, en este capítulo la distinción que se va a recuperar es aquella entre metodología cuantitativa y cualitativa. Para ello, el primer argumento, que es básico, es que la noción misma de muestreo, que surge del concepto de *muestra*, parte de una lógica de investigación cuantitativa y por tanto solamente cubre un segmento del espectro de ciencias relacionadas con la educación especial. Tal como lo mencionan Hernández, et al., (2014) una muestra es “un subgrupo del universo o población del cual se recolectan los datos y que debe ser representativo de ésta” (p.173). La necesidad de que la muestra sea representativa del universo, se debe a que se va a buscar que los datos obtenidos en dicha muestra puedan ser generalizados a todo el universo. Este criterio de científicidad es un rasgo que únicamente distingue a las ciencias cuya metodología es cuantitativa. Por ello en los siguientes apartados se hablará, por un lado, del muestreo en metodología cuantitativa y, por otro lado, de selección e invitación de participantes en metodología cualitativa.

Muestreo en metodología cuantitativa

En cualquier situación de la vida ordinaria una pregunta puede ser respondida. Sin embargo, no siempre las respuestas que siguen a una pregunta son verdaderas o suficientemente específicas. La metodología es un conjunto de técnicas o pasos a seguir para desarrollar investigación, que promueve que las preguntas disciplinares que se hacen los investigadores, puedan ser respondidas de forma tal que las respuestas tengan sentido, que sean verdaderas al amparo de un conjunto de postulados y premisas y, que en la medida de lo posible, tengan un buen nivel de especificidad.

El muestreo o selección de la muestra es uno de los pasos metodológicos esenciales en cualquier investigación y, por lo tanto, es uno al que se le debe poner

especial atención. Es solamente mediante un muestreo adecuado que se puede responder cabalmente a las preguntas de investigación que se establezcan, aunque no sea el único elemento importante para lograrlo. En otros capítulos de este mismo libro se detallan otros procesos o elementos metodológicos que también son ineludibles para la cabal respuesta a las preguntas de investigación planteadas.

Como se mencionó anteriormente la muestra es un segmento representativo de una población total que comparte una serie de características relevantes para la investigación en cuestión. Por lo tanto, para lograr un muestreo adecuado, lo primero que se ha de conseguir es identificar cuál es la población con la que se puede obtener respuesta a la(s) pregunta(s) de investigación planteada(s). Dependiendo del tipo de estudio, que con frecuencia está en función de la disciplina a la que éste se adscribe, la población puede estar conformada por personas, animales, incluso plantas o piedras, o bien por un conjunto de datos que se obtienen de empresas, grupos, encuestas, entre otras muchas opciones. Necesariamente, la idoneidad de cierta población para una investigación excluye o cancela como opciones poblacionales a muchos o a todos los demás elementos existentes en el entorno.

Por ejemplo, Camacho, et al., (2021) reconocieron la existencia de una discrepancia teórica en términos de los orígenes y, en consecuencia, de los posibles métodos para la reducción o eliminación de conductas estereotipadas en personas con Trastorno del Espectro Autista (TEA). Por ello, la pregunta de investigación de estos autores fue si una forma específica de intentar reducirlas o eliminarlas, las técnicas de modificación de conducta, que han sido probadas y confirmadas para la modificación de repertorios conductuales en diversos contextos y con poblaciones variadas, eran útiles para la modificación de la conducta estereotipada de personas con TEA en términos de reducción o eliminación.

En este caso, la población idónea para llegar a responder la pregunta se compone de personas con TEA, por lo que el resto de la humanidad que no tenga TEA queda automáticamente excluida como parte de la población adecuada para el estudio. Por otro lado, el tipo de pregunta de investigación hace pensar que no hay otros estudios que hayan intentado responder previamente la misma pregunta, por lo que sería un ejercicio fútil tratar de responderla en términos de la proporción de casos en los que se ha demostrado que dichas técnicas son útiles para la reducción de conductas estereotipadas en contraste con la proporción de aquellos casos en los que se ha demostrado lo contrario. Es posible que se pudiera tener una respuesta cercana a la deseada si se usara como población a sujetos animales que hayan sido alterados para ser modelos animales del TEA, sin embargo, en este caso y para esta pregunta, habiendo personas disponibles con TEA, y no habiendo posibilidad de daño físico o psicológico para los participantes, no hay necesidad ni razón para buscar una respuesta aproximada si se puede tener la respuesta directa a la pregunta de investigación planteada. Habrá otras preguntas relacionadas con el TEA que, por cuestiones éticas u otros criterios solo puedan ser respondidas con modelos animales, pero este no es el caso.

Con el fin de no llegar a conclusiones erróneas por exceso de generalización de datos, es necesario acotar lo más posible la población a la cual se pueden extender los datos obtenidos a partir de la experimentación con cierto segmento muestral. Las poblaciones se suelen acotar por medio de la explicitación de algunos de sus atributos como: raza, edad, nivel de estudios, género, nivel socio-económico, país o región de residencia, entre otras. Evidentemente, no todos los atributos son relevantes para acotar poblaciones, y los atributos relevantes varían en función de la pregunta de investigación que se esté planteando. Por ejemplo, en el caso del estudio de Camacho, et al., (2021) el tipo de estereotipia que se presenta o el nivel de

profundidad del trastorno son variables que cobran mucha mayor relevancia que el género o la nacionalidad, que llegan a ser altamente relevantes en otro tipo de estudios.

Una vez que se tiene identificada la población con la que se puede dar respuesta legítima a la pregunta de investigación, lo que sigue es llevar a cabo un procedimiento adecuado de selección del segmento poblacional al que llamamos muestra. En la mayoría de los estudios científicos se hace necesario seleccionar una muestra porque sería impracticable o imposible hacer el estudio con todos los individuos que conforman la población. En algunos casos, los números de la población son descomunales, por lo que ocuparía toda una vida, y muchos millones de pesos o de dólares, hacer un solo estudio. En otros casos, los elementos de la población están desperdigados a lo largo y ancho del planeta. Si Camacho, et al., (2021) hubieran decidido evaluar su pregunta de investigación incluyendo a todas las personas con TEA del mundo, seguramente al estudio le faltaría mucho para concluir y para ser publicado. Es por ello que poder seleccionar una muestra es algo muy útil, pero el criterio que se debe lograr o cumplir para considerar que el procedimiento de muestreo fue adecuado es que la muestra obtenida sea representativa de la población. Es decir, que, como ya se había mencionado, los datos obtenidos sean aplicables o generalizables al resto de la población. Para lograrlo existen diferentes formas que se abordan a continuación, y que se agrupan en una de dos grandes categorías: muestreo probabilístico y muestreo no probabilístico (García, 2009; Hernández, et al., 2014).

La principal característica del *muestreo probabilístico* es que todos los elementos de la población seleccionada para responder a la pregunta tienen la misma posibilidad de resultar elegidos para formar parte de la muestra del estudio. Por el contrario, en el *muestreo no probabilístico* la muestra se elige atendiendo a

valoraciones que, de inicio, evitan que cualquier miembro de la población pueda ser seleccionado para ser parte del estudio. Cabe mencionar que ninguno de los dos tipos de muestreo es mejor que el otro; tampoco la elección de un tipo de muestreo le confiere mayor o menor calidad a una investigación.

Los criterios a los que los investigadores deben atender para optar por uno de los dos tipos de muestreo son múltiples⁶. Uno de ellos es el alcance del estudio, es decir, si éste es un estudio exploratorio^{NP}, descriptivo^{NP}, correlacional^P o explicativo^P. Otro criterio es el diseño. Cualquier característica del diseño decantará al investigador por un tipo de muestreo; desde el nivel de control de variables (no experimental^{NP}, preexperimental^{NP}, cuasiexperimental^P, experimental^P) hasta la ventana temporal que se debe abarcar en función de la pregunta a responder (longitudinal^{NP} o transeccional^P). Incluso hay criterios para seleccionar el tipo de muestreo que no son dictados por criterios metodológicos sino que atienden a cuestiones operativas o fácticas tan simples como la lejanía del lugar de residencia de algunos elementos de la población, o el desconocimiento de los lugares en los que se encuentran ciertos miembros de la población. Otro criterio operativo que está asociado al anterior es la holgura económica para realizar el estudio, que en la mayoría de las ocasiones tiene que ver con la obtención de financiamiento y con la magnitud del mismo.

Dada la multiplicidad de criterios que se han de considerar para la elección del tipo de muestreo a realizar, es imposible explicitar combinaciones específicas de los criterios mencionados en el párrafo anterior que lleven decididamente a la elección de un muestreo probabilístico o no probabilístico. Sin embargo, está

En el siguiente conjunto de opciones se pondrá una P a los criterios que *decantan* hacia un muestreo probabilístico, y una NP a los criterios que *decantan* hacia un muestreo no probabilístico. NO son criterios directos hacia un tipo de muestreo u otro.

disponible el ejemplo que ponen Hernández, et al., (2014), que no considera todos los criterios disponibles, en el que mencionan lo siguiente: “Las muestras probabilísticas son esenciales en los diseños de investigación transeccionales, tanto descriptivos como correlacionales-causales (las encuestas de opinión o sondeos, por ejemplo), donde se pretende hacer estimaciones de variables en la población” (p.177).

Un paso común para ambos tipos de muestreo es el cálculo del tamaño de la muestra. Ese paso es el que va arrojar cuál es el número mínimo que debe tener la muestra para ser representativa. Obviamente, para calcular el tamaño de la muestra se necesita conocer el tamaño de la población. Así mismo se necesita establecer el grado de error que estamos dispuestos a aceptar en nuestros cálculos, es decir, qué tanto estamos dispuestos a equivocarnos en relación con que nuestra muestra sea representativa o no lo sea. A ese error, se le llama error estándar y su valor suele ser de 1% o de 5%, que son valores muy bajos. Otro dato requerido, que se tiene automáticamente cuando se establece el error estándar, es el nivel de confianza. Este valor se tiene en automático teniendo el valor del error estándar, debido a que es el valor complementario que permite llegar al 100%. Por ejemplo, si el error estándar es de 5%, el nivel de confianza es 95%. Si el grado de error es de 1%, el nivel de confianza es de 99%. Otro dato necesario para calcular el tamaño muestral, y el último, según algunos autores (e.g., Hernández, et al., 2014), es la probabilidad de ocurrencia del fenómeno o porcentaje estimado de la muestra. Este valor casi siempre es 50%, aunque se pueden tener valores resultantes de muestreos previos que darían una cifra más exacta.

En este momento es necesario aclarar dos puntos. El primero es que el cálculo del tamaño de la muestra se puede hacer de forma manual y, para ello existen fórmulas establecidas, o se puede hacer por medio de programas que facilitan el

cálculo. Lógicamente esos programas van a requerir ciertos datos para generar el cálculo. Esos datos son los que se mencionan en el párrafo anterior. El segundo punto a aclarar es que algunos autores consideran que, para un cálculo adecuado del tamaño muestral se requieren más datos, como el número de grupos por estudiar, el tipo de estudio, si la hipótesis es de una o dos colas, etc. (Ver, Rendón-Macías y Villasís-Keever, 2017).

Una vez que se tiene el tamaño de la muestra y se ha determinado si el muestreo será probabilístico o no probabilístico, lo que sigue es elegir el procedimiento para la selección de la muestra. De nuevo, la selección del procedimiento estará en función de la pregunta en cuestión. En los próximos párrafos quedarán claras la diferencias que hay entre los procedimientos y cómo es que cada uno probabiliza la obtención de una respuesta adecuada a la pregunta planteada. Los procedimientos más comunes para el muestreo probabilístico son el muestreo aleatorio o al azar, el muestreo estratificado y el muestreo por racimos o conglomerados (Ver, Tecla Jiménez y Garza Ramos, 1993).

Con frecuencia el muestreo aleatorio o al azar se caracteriza como la selección de papeles que están dentro de una bolsa o de bolitas de plástico que están en una canasta o en una caja, casi siempre de un material traslúcido. Obviamente, y en correspondencia con lo que se ha señalado, en la bolsa, en la canasta o en la caja, deben estar todos los elementos o individuos que componen la población, para que todos tengan la misma posibilidad de ser elegidos. Sin embargo, como lo señala García (2009), esas formas mecánicas de muestreo aleatorio pueden mostrar fallas, tal vez desdeñables, dado que la probabilidad de que todos los casos sean elegidos no sea exactamente la misma. Por ejemplo, que algunos de los papeles, por estática, se peguen a la bolsa. Será menos probable que esos papeles pegados sean elegidos. Otro ejemplo es que no se revuelvan suficientemente las bolitas de la canasta y quien

saca dichas bolitas de plástico siempre saque de las que están en la superficie. Así, las bolitas de abajo no tendrán la misma probabilidad de ser elegidas. Por ello, García recomienda las tablas de números aleatorios que generan algunos programas matemáticos, para el muestreo aleatorio o al azar.

Un ejemplo hipotético en el que el muestreo al azar o aleatorio es adecuado, es el siguiente: Se quiere saber si los Centros de Atención Múltiple (CAM) de Veracruz, están desarrollando habilidades adaptativas entre sus estudiantes con Discapacidad Intelectual, sin importar la severidad. Las habilidades adaptativas entendidas como aquellas que permiten autosuficiencia en alimentación, vestido y aseo. Por tanto, mediante el conocimiento que se tiene del tamaño de la población de estudiantes veracruzanos con Discapacidad Intelectual en CAM, se asigna un número a cada uno de los integrantes de dicha población y se calcula el tamaño de la muestra. Pensemos que el tamaño muestral es de 123 niños. Con base en una tabla de números aleatorios (Ver Figura 1), se elegirá a los 123 niños cuyo número asignado se corresponda con los números que arroje la tabla en función de la dirección en que se seleccionen los números de la misma. Por ejemplo, la selección puede ser vertical, horizontal o diagonal. De esta manera, se identificará a las niñas y niños con los cuales se comenzará a hacer las evaluaciones respectivas para la identificación de sus habilidades adaptativas.

Figura1. Ejemplo de tabla de números aleatorios

TABLE OF RANDOM DIGITS

3	76520	13586	34673	54876	80959
5	64894	74296	24805	24037	20636
3	19645	09303	23209	02560	15953
9	09376	70715	38311	31165	88676
0	80157	36147	64032	36653	98951
7	34072	76850	36697	36170	65813
5	45571	82406	35303	42614	86799
2	02051	65692	68665	74818	73053
5	05325	47048	90553	57548	28468
3	03529	64778	35808	34282	60935
7	14905	68607	22109	40558	60970
1	39808	27732	50725	68248	29405
4	06288	98083	13746	70078	18475
0	86507	58401	36766	67951	90364

Fuente: Tomada de Quevedo (2012)

En el muestreo estratificado, se procede de una forma semejante a la del muestreo aleatorio. Sin embargo, este muestreo es útil, incluso necesario, cuando la pregunta que se pretende responder implica nichos o estratos de la población. Un ejemplo hipotético en el que el muestreo estratificado es el adecuado es el siguiente: Se pretende evaluar si las habilidades sociales, entendidas como aquellas que se apegan a reglas e instrucciones, de niños con Síndrome de Down de las distintas escuelas de Saltillo, adoptan cualidades diferenciadas en función del grado escolar en nivel Básico Primaria. Para ello, primero se van a definir los estratos en los que se va a incorporar a toda la población. En este caso, los estratos serán: primero, segundo, tercero, cuarto, quinto y sexto de primaria. Estos estratos conformarán grupos. Posteriormente, se hace el análisis que permite identificar el tamaño de la muestra representativa. En seguida, se realiza un muestreo que, en la forma, es idéntico al muestreo aleatorio ya descrito, pero ahora en cada uno de los estratos. Por tanto, se asignarán los números a cada niño, por estrato, no en general, y se hará la selección con la tabla de números aleatorios por estrato. De tal suerte que se

asegura que la muestra queda compuesta por miembros de todos los estratos y ello permite hacer la comparación entre ellos, como se planteó inicialmente. Cabe aclarar que en este muestreo no se busca y mucho menos se asegura que el número real de individuos o miembros en cada uno de los diferentes estratos sea igual al número de los demás estratos. El número de individuos que componen la muestra, por estrato, será proporcional al número de individuos que componen la población por estrato.

El muestreo por racimos o conglomerados puede quedar muy claro si se mantiene el nombre de *por racimos*. Si uno ve una vid, uno podrá observar que, cuando da sus frutos, los frutos están agrupados en racimos de uvas. Al ser todos ellos racimos de la misma vid, se puede presumir que todas las uvas de todos los racimos tendrán una composición semejante. Es decir, si se quiere tener una muestra de las uvas de esa vid, lo mismo da tomar las uvas de un racimo que de otro. Este tipo de muestreo, parte de la lógica de que, para tener mi muestra, no requiero tomar pocas uvas de todos los racimos. Simplemente tomo los racimos necesarios para conformar mi muestra de uvas de esa vid. Si la población, miembros con características semejantes, está agrupada en ciertos lugares, es mucho más ágil elegir aleatoriamente cierto número de lugares que me permitan tener mi muestra representativa, en comparación con la estrategia de conformar mi muestra a partir de pocos integrantes de todos los lugares. El muestreo por racimos permite conformar una muestra probabilística y llevar a cabo toda la investigación con menos recursos energéticos, económicos y temporales.

En este caso, para poner un ejemplo en el que el muestreo por racimos se puede aplicar, vale la pena regresar al ejemplo de los niños veracruzanos con Discapacidad Intelectual. Recordemos que el tamaño de la muestra era de 123 niños. Sabiendo esto, los números no los asigno a los niños con las características mencionadas de Veracruz, sino a las escuelas en las que están los niños. Como el

número de niños por escuela es distinto, con la tabla de números aleatorios se elegirá la primera escuela y se verá cuántos niños tiene con Discapacidad Intelectual. Pensemos que la primera tiene 5. Faltan 118. Entonces se elige el siguiente número de la tabla, que corresponde a otra escuela. En esa tienen 10. Faltan 108. Y así se repite el procedimiento hasta completar la muestra.

Así como existen tres procedimientos, que son los más comunes para el muestreo probabilístico, cuando se recurre al muestreo no probabilístico, también llamado *intencional* (García, 2009) o de *muestras dirigidas* (Hernández, et al., 2014), igualmente se dispone de tres procedimientos empleados con mayor frecuencia. Éstos son el de sujetos tipo, el de expertos, y el de por cuotas. En este momento, antes de describir los procedimientos enunciados, es imprescindible señalar que los muestreos no probabilísticos son los ideales para responder ciertos tipos de preguntas, en las que la generalización de los datos no sea el criterio medular, sino el control cuidadoso y riguroso en las características muestrales.

El muestreo por sujetos tipo se refiere a la selección de la muestra a partir de características muy concretas, habitualmente muy restringidas, de los individuos que la conforman. Para clarificar este tipo de muestreo, un ejemplo hipotético útil es en el que hay algunas escuelas de Nuevo León, en las que no se sigue el modelo propuesto por la Secretaría de Educación Pública para atender a la población con Aptitudes Sobresalientes, que es el más empleado en todos los estados de la República, sino que se ha optado por instaurar el uso de centros de interés. Los centros de interés son una estrategia de enriquecimiento que apoyan el desarrollo y la especialización en áreas muy concretas de desempeño. Se pretende evaluar si los estudiantes con Aptitudes Sobresalientes que están bajo este modelo de atención, desarrollan exponencialmente sus aptitudes sobresalientes, y establecer la correlación entre el desarrollo de las aptitudes y el decremento de problemas de

conducta bajo este modelo. De por sí, la cantidad de estudiantes con Aptitudes Sobresalientes, por escuela, suelen ser pocos. Si se añade el criterio de que solo se evalúe a los que están bajo el modelo descrito la población es todavía más reducida. Por ello, aunque no siempre es así, en el caso de los muestreos por sujetos tipo, la población y la muestra llegan a ser muy semejantes o iguales en términos numéricos.

El muestreo de expertos se emplea cuando la información que se requiere para responder a la pregunta debe producirse a partir de lo que digan o hagan personas con mucha experiencia o muchos conocimientos en una temática particular. Es común que, para ciertos fenómenos o temáticas, los expertos no abundan. Pero incluso hay ciertas preguntas que acotan la muestra todavía más. Por ejemplo, si la pregunta que se desea responder tiene que ver con las políticas públicas desarrolladas por el gobierno del estado de Apurímac, en la hermana República del Perú, para la atención de niños con Trastorno del Espectro Autista, automáticamente quedan excluidos todos los expertos en políticas públicas de otros lugares, incluyendo a los expertos en la misma temática de los otros 23 estados del Perú.

El muestreo por cuotas es un procedimiento en el que se puede observar de forma muy clara la norma de selección de una muestra representativa de la población sin que se tenga que cumplir el criterio de que todos los individuos tengan la misma probabilidad de ser elegidos. En el muestreo por cuotas, se establece un valor numérico, que comúnmente es un porcentaje de la población, que es igual o superior al valor que asegura la representatividad. A partir de ello, se selecciona al número de integrantes de la muestra, bajo el criterio único de cumplir con el número establecido. Así, si se puede cumplir la cuota, por ejemplo, con los miembros de la población que se encuentran más cerca no se está contraviniendo ningún criterio, norma o requisito. De esta manera, se cumple el criterio de representatividad poblacional y el criterio de la cuota establecida, pero, en el ejemplo, todos aquellos

miembros de la población que se encuentren lejos no tuvieron la oportunidad de ser elegidos para formar parte de la muestra.

Por ejemplo, pensemos que un equipo de investigadores que trabajan en la alcaldía Iztapalapa, quieren evaluar el tipo de apoyo educativo que dan los padres de familia de niños ciegos con bajos recursos económicos, que residen en la Ciudad de México. Proceden a establecer el tamaño poblacional y el tamaño muestral. Y resulta que la muestra es de 1500 niños, pero la cuota se establece en 1700 niños, siguiendo algún criterio metodológico adicional. Analizando la distribución muestral en la Ciudad de México, los investigadores se dan cuenta que, dado el tamaño de la alcaldía, la cuota la cubren con puras familias de Iztapalapa y deciden hacerlo así. Estos investigadores cumplieron con todos los requisitos del muestreo por cuotas y su muestra es metodológicamente pulcra.

Como se puede observar, desde la lógica de la investigación cuantitativa, hay una serie de criterios que se deben cumplir para favorecer la obtención de una respuesta lógica, congruente y coherente en relación con la pregunta que se planteó. Además, el cumplimiento de ciertos criterios permitirá alcanzar algunos de los estándares más altos de científicidad como lo son la generalización y la replicación de datos. En lo que refiere a la selección de la muestra, todos esos criterios se cumplen por medio de procesos y estrategias establecidos que, en cuanto se cumplen, permiten iniciar el estudio y llevar a cabo de forma rigurosa todos los pasos de un procedimiento bien planteado y apegado a un diseño que cumple otro conjunto de criterios metodológicos y teóricos. Se puede decir que, desde la lógica cuantitativa, todo lo que tiene que ver con muestreo se concibe como un conjunto de preparativos muy relevantes para llevar a cabo el estudio. A continuación se verá que desde la metodología cualitativa, en general, la lógica es distinta.

Selección e invitación de participantes en metodología cualitativa

Desde la perspectiva y la lógica de la metodología cualitativa, la noción de *selección de la muestra* o *muestreo*, así como la noción misma de *muestra* encuentran dificultades para ser incorporadas en el proceso de investigación de forma plena, por múltiples razones. Sin embargo, esta afirmación de ninguna manera pretende generar la idea de que en investigación cualitativa no existe un proceso con el cual se elige o se selecciona a individuos o miembros de un grupo a los que se les invitará a participar en la investigación bajo el papel de participantes. Este proceso existe y, al igual que como se mostró con la metodología cuantitativa, desde esta perspectiva metodológica se tienen procedimientos y criterios que se han de seguir para responder congruente y coherentemente a las preguntas planteadas.

Previo a la exposición de los procedimientos y criterios mencionados, vale la pena comentar un aspecto esencial de la lógica cualitativa en lo que toca al conocimiento de *las realidades*. La razón principal para afirmar que desde la postura filosófico-metodológica que se está desarrollando, la noción de selección de la muestra no es plena, se debe a la noción de *co-construcción* del conocimiento del que el(los) investigador(es) o la(s) investigadora(s) forman parte.

En cierto sentido, en investigación cualitativa, no se pueden elegir las características de todos los participantes puesto que, por un lado, como lo mencionan Denzin y Lincoln (2015) “los observadores ahora actúan como participantes que colaboran en las acciones de la situaciones de investigación” (p. 47); y por otro lado, estos mismos autores mencionan que “las características personales del entrevistador, como su raza, su clase, su grupo étnico y su género influyen en este método” (p.46). Es decir, en la investigación cualitativa, la generación de conocimiento se suele entender como un proceso de co-construcción en el que el

mismo investigador y sus características influyen de forma importante. Vasiliachis (2006) lo enuncia de la siguiente forma:

Quien realice investigación cualitativa [...] Debe tener presente que sus valores, perspectivas, creencias, deseos, expectativas influyen en la percepción y en la construcción de la realidad que estudia, y que la experiencia vivida es también una experiencia corporeizada, siendo la propia investigadora o el propio investigador una fuente de datos (p. 36)

Dicho de otra manera, el investigador no puede elegir si se invita a sí mismo y tampoco puede apartarse o des-elegirse del proceso de co-construcción que subyace a la investigación cualitativa.

Sin descuidar la afirmación anterior, es relevante contemplar que, tal como lo afirma Martínez-Salgado (2012), las unidades de observación, que naturalmente están íntimamente ligadas a los agentes de información, son muy variadas e incluyen: “personas, familias, grupos, instituciones, áreas geográficas o culturales, períodos, otras unidades de tiempo, programas, documentos, combinaciones de varias de las anteriores, entre otras” (p. 614).

Esta misma autora, junto con muchos otros (e.g., Ander-Egg, 2011; Gundermann Kröll, 2013; Mallimaci y Giménez Béliveau, 2006; Stake, 1999), hacen explícito que desde la metodología cualitativa, dependiendo del diseño y de los objetivos de investigación que se pretenden alcanzar, las unidades de observación pueden ser muy pequeñas, incluso pueden ser igual a 1, o bien pueden llegar a ser muy amplias. Lo que es cierto para toda la investigación cualitativa es que cada unidad es seleccionada de forma intencional por la posibilidad que tiene de ofrecer información. Por estas condiciones de selección, se llega a asemejar al muestreo por sujetos tipo o al muestreo de expertos de la metodología cuantitativa, sin la necesidad

de cumplir un criterio de representatividad poblacional. Es por ello que a la selección cualitativa se le ha conocido como selectiva, de juicio o intencional.

Si algo debe quedar claro para quienes se incorporan a la aventura de la investigación cualitativa es que los criterios para la selección de participantes, aunque existen, no son tan definidos o tan secuenciados como en la metodología cuantitativa. Debido a la lógica cualitativa, que entraña la subjetividad de y entre casos, los criterios son flexibles y los límites difusos. Por tanto, si alguien espera encontrar un texto en el que se le presente paso 1, paso 2, paso 3 o criterio 1, criterio 2, criterio 3 para la selección de participantes y para otros tantos aspectos metodológicos cualitativos, lo más seguro es que termine frustrado. Sin embargo, dicho lo anterior, algunos autores han realizado esfuerzos por aportar líneas, ni rectas ni rígidas, de sistematicidad a los diferentes procesos que conforman la metodología cualitativa. A continuación se presentan algunos de esos esfuerzos, sin la pretensión de la enunciación exhaustiva, por considerarla imposible.

Un ejemplo de sistematización dirigido a la selección de participantes es el aportado por Toro Jaramillo y Parra Ramírez, (2010), quienes distinguen a los participantes como *claves* o *informantes*, *especiales* y *representativos*, dependiendo del tipo de información que pueden aportar. Según estos autores, los participantes clave o informantes son aquellos que aportan información acerca del contexto o bien que facilitan o apoyan la participación de entrevistados que aporten información directa sobre las preguntas que se pretenden responder o sobre los objetivos que se quieren conseguir. En contraste, los participantes especiales se distinguen por aportar información directa respecto del objetivo o de la pregunta central de la investigación, y por pertenecer a una élite de personas que ocupan posiciones únicas dentro de grupos o instituciones. En el caso de la Educación Especial, podrían ser Secretarios o Sub-secretarios de Educación, Presidentes de Organismos, entre otros.

En cambio, los participantes representativos aportan información directamente relacionada con las preguntas u objetivos de investigación pero lo que los hace especiales o relevantes no es una posición de élite al interior de una agrupación particular, sino el lugar, función o papel que ocupan dentro de su contexto cultural, social e histórico.

En un sentido clasificatorio semejante, y si se quiere alternativo, al de Toro Jaramillo y Parra Ramírez, en el cual el criterio para la categorización funcional de los participantes aparece más apegado al papel histórico, Mallimaci y Giménez Béliveau (2006) plantean la distinción entre participantes como: *gran persona, marginal y naturales o normales*. De forma intuitiva, aquellos considerados como gran persona son quienes, en función del objetivo de investigación planteado, tuvieron un papel decisivo en el desarrollo de hechos históricos. Tal sería el caso de Adolfo Hitler en el genocidio judío de la Segunda Guerra Mundial o, en un caso más cercano a la Educación Especial, la Ex secretaria de Educación Delfina Gómez Álvarez en la creación de la Nueva Escuela Mexicana. En contraste con el concepto de gran persona, el concepto de participante marginal, de alguna manera es contraintuitivo. Este concepto hace referencia a participantes que están entre los límites de dos o más realidades sociales o culturales. La selección de participantes marginales, suele estar justificada por la necesidad de conocer aquellos aspectos, significados y apreciaciones poco comunes o poco consensuados. Las personas con discapacidad con su bagaje de vivencias, luchas, impulsos y significados en escenarios áulicos regulares suelen ser excelentes ejemplos de este tipo de participación. En contraste con la noción de marginal, los participantes naturales o normales son aquellos totalmente situados en las mayorías. Con estos participantes, el reto es separarse de las generalizaciones groseras y las interpretaciones comunes.

Desde una perspectiva distinta a las dos anteriores, Teddlie y Yu (2007), aportan una categorización de métodos para la selección de participantes, basada en propósitos. Tal sistema categorial consta de cuatro propósitos o subconjuntos de métodos cuyos nombres, en su inmensa mayoría, son indicativos del tipo de casos que seleccionan. Tal vez por ello, en muy pocas ocasiones definen o desarrollan el tipo de casos que pretenden reclutar.

El primero de ellos engloba a aquellos métodos cuyo propósito implica alguna comparación con otras situaciones o casos; ya sea para encontrar similitudes o diferencias. En este primer subconjunto incluyen el de selección de casos típicos, la selección de casos extremos, la selección de intensidad, que incorpora casos que manifiestan con ímpetu y contundencia el fenómeno bajo estudio, la de máxima variación, la homogénea y la de casos reputados. El subconjunto dos incluye a los tipos de selección que se centran en el caso mismo o en el conjunto de casos elegidos desde el inicio. En este subconjunto incluyen la selección de caso crítico, en el cual hay ciertas dimensiones que lo hacen crítico y que han de ser identificadas *a priori*, la selección de caso revelador, la selección de casos políticamente importantes y la de colección completa, que intenta incluir todos los casos o individuos que cumplen con cierto criterio que los hacen miembros de la colección. El tercer grupo es uno muy particular, en el que la selección de participantes es secuencial o gradual. Este tipo de selección puede deberse a una de dos razones. Una de ellas es la construcción de una teoría y por lo tanto, conforme se van obteniendo los datos se va decidiendo si el número de participantes ya fue suficiente o hacen falta más para la conformación de la teoría. Dicho de otra forma, en este tipo de investigaciones, el establecimiento del número de participantes no es algo que se establece al inicio o previo al estudio, sino que se decide sobre la marcha. La segunda razón es semejante, pero no obedece a la construcción de una teoría sino a la explicación de un

fenómeno. Dependiendo de los datos que se obtienen, se valora si existen rasgos o elementos sin clarificar. Entonces, algunos de los casos de este subconjunto son la selección teórica o basada en la teoría, la selección de casos confirmatorios o contradictorios, la selección de oportunidad o emergente y la de bola de nieve o de cadena en la que un participante sugiere a un(a) conocido(a) para fungir también como tal. El cuarto subconjunto incluye a los tipos de selecciones que combinan los tres subconjuntos anteriores.

Como lo señala Martínez-Salgado (2012), a diferencia de la metodología cuantitativa en la que la noción de representatividad implica criterios claros y certeros para determinar el número de elementos que conforman la muestra, en la investigación cualitativa no hay, ni se pretende tener, criterios así de claros para determinar el número mínimo o máximo de participantes.

Tal vez las selecciones englobadas por Teddlie y Yu (2007) en el tercer subconjunto, el de selección secuencial o gradual, sean las que cuentan con un criterio relativamente claro, aunque solo se pueda determinar conforme se van obteniendo los datos. Dicho criterio es el de saturación. En diversos textos se establece la distinción entre saturación de datos y saturación teórica. La saturación de datos es definida por Álvarez-Gayou (2003), como: “[el] momento en el que durante la obtención de la información, ésta empieza a ser igual, repetitiva o similar” (p. 33). Por supuesto, el criterio entonces es que cuando se llega al punto de saturación de datos, es el momento en el que no se requieren nuevos o más participantes.

En relación con la saturación teórica, el mismo autor refiere:

Cuando dejan de emerger nuevos datos, nuevas propiedades, nuevas categorías o nuevas relaciones se considera que se ha llegado a la saturación

teórica. Una vez que los datos no ofrecen diferencias ni distinciones conceptuales importantes, las categorías se han saturado. La saturación teórica se refiere a conceptos, no a datos, y a un momento en el que ya no se requiere mayor conceptualización de los datos (p.95).

Como se puede observar es un criterio semejante al de saturación de datos, que muestra el límite en el que ya no es necesario reunir más información y, por lo tanto, no hay que recurrir a más experiencias o a más participantes. Ambos tipos de saturación, que no dejan de ser criterios parcialmente subjetivos, en tanto los conceptos de similitud, semejanza y novedad no son precisos ni exactos, arrojan tanto el número mínimo como el número máximo de participantes, debido a que, en el caso de los dos tipos de saturación, el número mínimo y el número máximo, son la misma cosa.

Naturalmente, entre los teóricos de la metodología cualitativa está vigente el dilema sobre qué tan generalizado debe ser el uso de la saturación, o las saturaciones, como recurso para designar el número de participantes o incluso como criterio de validez de los datos y las investigaciones. Dicho de otra forma, ¿la saturación puede ser usada de forma a-teórica y al margen de los diferentes diseños de investigación cualitativa? Se insiste y se subraya que es un dilema y, por tanto, no hay una respuesta establecida. Lo cierto es que hay una serie de argumentos interesantes en contra de la saturación como regla de oro en investigación cualitativa (Ver, Braun y Clarke, 2019), mismos que no se recuperarán en este escrito, por considerar que escapan al objetivo del mismo.

Un ejemplo más en el que se ofrecen criterios para determinar el número más conveniente de participantes por estudio, Braun y Clarke (2019), proponen atender al criterio subjetivo de apegarse a los criterios establecidos por aquellos quienes suelen evaluar las investigaciones, por ejemplo: los editores de las revistas, los

evaluadores de proyectos, los comités de tesis, entre otros. Necesariamente, este criterio solo se puede poner en práctica a través de la comparación con estudios anteriores similares que hayan sido evaluadas por los agentes evaluadores mencionados.

A manera de conclusión de este apartado, parece correcto concluir que aquellas respuestas definitivas a las preguntas relacionadas con el número adecuado de participantes para los estudios deben seguir siendo buscadas como parte de un proceso natural de mejoramiento de la ciencia.

Síntesis a manera de conclusión

Si se acepta que la Educación Especial es una interdisciplina profesional, se ha de aceptar también que ésta no cuenta con criterios metodológicos de investigación propios. Sin embargo, se vislumbran opciones para que las y los educadores especiales puedan aportar al *corpus* científico de forma legítima, a saber: la investigación traslacional y la investigación basada en evidencia. En ambos casos, las y los educadores especiales deberán apegarse a los criterios metodológicos de las disciplinas científicas que aportan conocimiento de esta naturaleza a la Educación Especial para la solución de su encargo social.

Debido a las características disciplinares de la Educación Especial, las disciplinas científicas de las que se vale son lo suficientemente diversas para requerir aglutinarse bajo un criterio metodológico definido que les dé orden y criterios generalizados en el aspecto metodológico. El criterio metodológico para la conjunción de las diversas disciplinas fue el de la distinción dicotómica cuantitativo/cualitativo.

El presente escrito trata acerca de las directrices metodológicas enfocadas en la consecución de un conjunto de integrantes que aportarán la información necesaria para dar respuesta a las preguntas científicas justificadas disciplinar y teóricamente que se lleguen a plantear. Al margen de los detalles procedimentales para lograr dicho objetivo, en el presente escrito, se pretendió subrayar que dicho proceso, en las disciplinas que se suscriben a la metodología cuantitativa en contraste con las que se inscriben en la metodología cualitativa, es disímil de inicio a fin. Tal diferencia descansa en los aspectos epistemológicos que le dan sentido a ambas metodologías para la producción científica de conocimiento. Dado lo anterior, la claridad que deben tener las y los Educadores Especiales acerca de los objetos de estudio que aportan especificidad disciplinar a las ciencias que alimentan a la Educación Especial, es un componente crucial.

Referencias bibliográficas

- Aguilar, F., y Camacho, J. (en prensa). Los olvidados. En: V. Pacheco (Coord.) *Efectos e implicaciones de la pandemia en el trabajo académico*. Facultad de Estudios Superiores Iztacala-UNAM.
- Aguilar, F., Camacho, J., y Lima, N. (2020). Traslational research: The ontological relation between adjunctive and stereotyped behavior case. *Journal of behavior, Health and Social Issues*, 12(1), 25-31.
- Álvarez-Gayou, J. (2003). *Cómo hacer investigación cualitativa: Fundamentos y metodología*. Paidós Educador.
- Ander-Egg, E. (2011). *Aprender a investigar: Nociones básicas para la investigación social*. Editorial Brujas.
- Braun, V., y Clarke, V. (2019). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, 1-16.

- Camacho, J., Almanza, M., y Romero, R. (2015). Presentación. En: J. Camacho, M. Almanza y R. Romero (Coords.) *Neurociencia y educación especial: Conceptos, procesos y principios básicos*. (pp. 17-22). Universidad de Guadalajara y Universidad Autónoma de Tlaxcala.
- Camacho, J., Navarrete, L., Aguilar, F., Tecamachaltzi, M., y Cabrera, F. (2021). Aprendizaje Discriminativo y RDI como alternativa para la reducción de conducta estereotipadas motoras en el Trastorno del Espectro Autista. *Revista Electrónica de Psicología Iztacala*, 24(3), 898-923.
- Cook, B., y Cothren, S. (2011). Unraveling Evidence-Based Practices in Special Education. *The Journal of Special Education*, 47(2), 71-82.
- Cook, B., Tankersley, M., y Landrum, T. (2009). Determining Evidence-Based Practices in Special Education. *Council for exceptional children*, 75(3), 365-383.
- Denzin, N., y Lincoln, Y. (2015). *Métodos de recolección y análisis de datos*. Gedisa.
- García, B. (2009). *Manual de métodos de investigación para las ciencias sociales: Un enfoque de enseñanza basado en proyectos*. Manual Moderno.
- Gundermann Kröll, H. (2013). El método de los estudios de caso. En: M. Tarrés (Coord.) *Observar, escuchar y comprender: Sobre la tradición cualitativa en la investigación social*. (pp. 231- 263). Flacso.
- Hernández, R., Fernández, C., y Baptista, P. (2014). *Metodología de la investigación*. Mc Graw Hill.
- Mallimaci, F., y Giménez Béliveau, V. (2006). Historia de vida y métodos biográficos. En: I. Vasiliachis (Coord.). *Estrategias de investigación cualitativa*. (pp. 175- 212). Gedisa.
- Martínez-Salgado, C. (2012). El muestreo en investigación cualitativa. Principios básicos y algunas controversias. *Ciência & Saúde Coletiva*, 17(3), 613-619.
- Odom, S., Brantlinger, E., Gersten, R., Horner, R., Thompson, B., y Harris, K. (2005). Research in Special Education: Scientific Methods and Evidence-Based Practices. *Council for exceptional children*, 71(2), 137-148.
- Quevedo, L. (2012). Estudio sobre la simulación de eventos discretos. *Tesis no publicada*. <https://doi.org/10.13140/RG.2.2.33228.92809>

- Rendón-Macías, M. y Villasís-Keever, M. (2017). El protocolo de investigación V: el cálculo del tamaño de muestra. *Revista Alergia México*, 64(2), 220-227.
- Ribes, E. (2005). Reflexiones sobre la Eficacia Profesional del Psicólogo. *Revista Mexicana de Psicología*, 22(1), 5-14.
- Ribes, E. (2011). La psicología: Cuál, Cómo y Para qué. *Revista Mexicana de Psicología*, 28(1), 85-92.
- Sánchez-Sosa, J. (2008). Competencias científicas y profesionales: Cimientos metodológicos y de integración en las ciencias del comportamiento. En: C. Carpio (coord.) *Competencias profesionales y científicas del psicólogo: Investigación, experiencias y propuestas*. FES, Iztacala.
- Stake, R. (1999). *Investigación con estudio de casos*. Morata.
- Tecla Jimenez, A., y Garza Ramos, A. (1993). *Teoría, Métodos y Técnicas en la Investigación Social*. Ediciones Taller Abierto.
- Teddlie, C., y Yu, F. (2007). Mixed Methods Sampling: A Typology with Examples. *Journal of Mixed Methods Research*, 1(1), 77-100.
- Toro Jaramillo, I., y Parra Ramírez, R. (2010). *Fundamentos epistemológicos de la investigación y la metodología de la investigación. Cualitativa/cuantitativa*. Fondo Editorial Universidad EAFIT.
- Vasiliachis, I. (2006). La investigación cualitativa. En: I. Vasiliachis (Coord.). *Estrategias de investigación cualitativa*. (pp. 23- 64). Gedisa.

INTRODUCCIÓN A LA ESTADÍSTICA

*Susana Angelica Castro Chavira [‡]
M. del Consuelo Pedraza Aguilar*

Resumen

La educación está dirigida a incrementar las capacidades de los alumnos en la resolución de problemas y el aprovechamiento del conocimiento para lograr beneficios individuales y sociales a corto y largo plazo. La estadística permite recabar y transformar los datos obtenidos en el aula para su interpretación, comunicación y uso para la toma de decisiones en el ámbito de la educación especial. El modelo estadístico enfocado en la solución de problemas PPDAC (Problema – Plan – Datos – Análisis – Conclusión y Comunicación) propone una serie de cinco etapas para el uso de datos considerando su contexto, el objetivo de su análisis y la audiencia a la cual están dirigidos. Con frecuencia, la estadística se presenta como un conjunto de herramientas de análisis; sin embargo, como se muestra en la presentación del modelo para la solución de problemas PPDAC, el análisis es sólo la cuarta etapa de una serie de cinco que constituye al método estadístico. El modelo estadístico PPDAC puede ser aplicado a problemas como la exploración de los factores que intervienen en la adquisición de la lectoescritura en niños nacidos prematuramente con diversos factores de riesgo prenatales y perinatales para daño cerebral, independientemente del nivel de sofisticación matemática que se utilice. Se puede abordar problemas interesantes y sustantivos sin recurrir a herramientas de análisis complejas, conjuntos de datos enormes o grandes recursos computacionales. Finalmente, la visualización y comunicación de datos mediante un flujo narrativo es clave para transformarlos en información que pueda ser útil a educadores, directores, escuelas e instituciones.

Palabras Clave: Enfoque en solución de problemas, método estadístico, visualización de datos..

[‡] Unidad de Investigación en Neurodesarrollo “Dr. Augusto Fernández Guardiola”, Instituto de Neurobiología, Universidad Nacional Autónoma de México. Blv. Villas del Mesón 76230, Juriquilla, Querétaro castrochavirasa@inb.unam.mx

Introducción

El quehacer educativo está dirigido a incrementar las capacidades de los alumnos para la resolución de problemas y para el aprovechamiento del conocimiento que les permita lograr beneficios individuales y sociales a corto y largo plazo.

A partir de este objetivo, la enseñanza requiere del docente la capacidad para analizar con rapidez y cuidado las cambiantes interacciones sociales e instruccionales y responder ante diversas situaciones. La enseñanza exitosa requiere que los docentes tengan una comprensión profunda del desarrollo humano, de la teoría del aprendizaje y de la instrucción, de la motivación y de la teoría del manejo del salón de clases, así como conocimiento sobre la manera de evaluar el aprendizaje del alumno (Good y Brophy, 1996).

En particular, la Educación Especial está dirigida a ofrecer atención con equidad a alumnos con discapacidad o bien con aptitudes sobresalientes y de acuerdo con sus condiciones, necesidades, intereses y potencialidades. En México, el sistema educativo brinda atención educativa a alumnos que enfrentan barreras para el aprendizaje y la participación, asociadas a discapacidad y aptitudes sobresalientes, con el fin de promover su ingreso, permanencia y egreso de los niveles de educación inicial, preescolar, primaria, secundaria y formación laboral (Romero Contreras y García Cedillo, 2013).

Aunque cada alumno es distinto, existen patrones de aprendizaje comunes, así como estrategias de enseñanza específicas para las habilidades cognitivas que se requiere facilitar. Para que el proceso de enseñanza-aprendizaje sea efectivo, la evaluación, como herramienta de retroalimentación, permite adecuar las estrategias de enseñanza y mejorar las propias herramientas de evaluación. En la educación especial, resulta particularmente relevante basar el manejo y uso de

estrategias de enseñanza en la información obtenida a partir de evaluaciones integrales que proporcionen datos del desarrollo de los individuos en diversas áreas (Álvarez Ramos, 2007). La ciencia estadística permite recabar y transformar la información en forma de datos numéricos para su interpretación, comunicación y uso para la toma de decisiones en el ámbito de la educación especial. Más aún, ser capaz de visualizar datos y contar historias con ellos es clave para transformarlos en información que pueda ser usada para guiar la toma de decisiones de manera óptima (Nussbaumer Knafllic, 2015).

La visión común de la estadística como una “caja de herramientas” básica enfrenta grandes retos en la actualidad. Nos encontramos en la era de la ciencia de datos en la que grupos de datos grandes y complejos son extraídos de fuentes rutinarias como detectores de velocidad, publicaciones en redes sociales y compras por internet y son utilizados como base para innovaciones tecnológicas como la optimización de rutas de transporte, publicidad dirigida o sistemas de sugerencia de compra. Los grandes datos o *big data* se caracterizan por cuatro Vs: Volumen de datos enorme, Variedad de fuentes como son imágenes, cuentas de redes sociales o transacciones, Velocidad de obtención alta y Veracidad incierta debido a su obtención rutinaria (Spiegelhalter, 2019).

En la actualidad, el entrenamiento estadístico es visto como sólo un componente necesario de convertirse en un científico de datos, junto con habilidades de manejo de datos, programación y desarrollo de algoritmos, así como el conocimiento propio de la materia de estudio. Lejos de liberarnos de la necesidad de habilidades estadísticas, mayor cantidad de datos y el incremento en el número y complejidad de estudios científicos hace aún más difícil extraer conclusiones correctas. Más datos significa que necesitamos estar aún más conscientes del valor real de la evidencia (Spiegelhalter, 2019).

El conocimiento de datos describe no sólo la habilidad de realizar análisis estadísticos de problemas reales, sino también la comprensión y capacidad crítica de cualquier conclusión obtenida, por otros, con base en la estadística. Además, implica la capacidad para entender los principios detrás del aprendizaje a partir de datos, llevar a cabo análisis de datos básicos y discernir la calidad de las conclusiones obtenidas a partir de esos datos.

La ciencia de datos es el estudio y aplicación de técnicas para derivar discernimientos a partir de datos, incluyendo la construcción de algoritmos para predicción. La ciencia estadística tradicional forma parte de la ciencia de datos, la cual también incluye un fuerte componente de codificación y manejo de datos (Spiegelhalter, 2019).

El quehacer de la estadística

La estadística es la disciplina que se ocupa de 1) la obtención, organización, resumen y análisis de datos y 2) la obtención de inferencias a partir de un volumen de datos cuando se examina sólo una parte de estos (Daniel, 2006). Para usar la ciencia estadística en la generación de conocimiento, nuestras experiencias diarias deben ser transformadas en datos, lo que implica categorizar y designar eventos, registrar mediciones, analizar los resultados y comunicar las conclusiones (Spiegelhalter, 2019).

La estadística es siempre, hasta cierto punto, construida en base a juicios y es evidente que la complejidad completa del salón de clases no puede ser codificada sin ambigüedad en una hoja de cálculo u otro software. El uso de datos es útil para comprender las situaciones que se presentan en el aula y hacer mejores juicios. En ocasiones, una simple búsqueda de patrones mediante un trabajo

exploratorio e iterativo, sin matemáticas ni teoría, una especie de “estadística forense”, conduce a preguntas interesantes (Spiegelhalter, 2019).

Existen dos tipos de datos, los obtenidos del muestreo y los obtenidos mediante cálculos matemáticos y transformaciones numéricas (Daniel, 2006). Sin embargo, los datos como fuente de conocimiento tienen dos principales limitaciones. Primero, son casi siempre una medida imperfecta de nuestro verdadero objeto de interés. Segundo, lo que sea que decidamos medir diferirá entre un lugar y otro, de una persona a otra, de un tiempo a otro y el problema es extraer discernimientos significativos de toda esta aparentemente aleatoria variabilidad. Variabilidad se refiere a las inevitables diferencias que ocurren entre medidas u observaciones, algunas de las cuales podrían ser explicadas por factores conocidos y el resto atribuido a ruido / error aleatorio (Spiegelhalter, 2019).

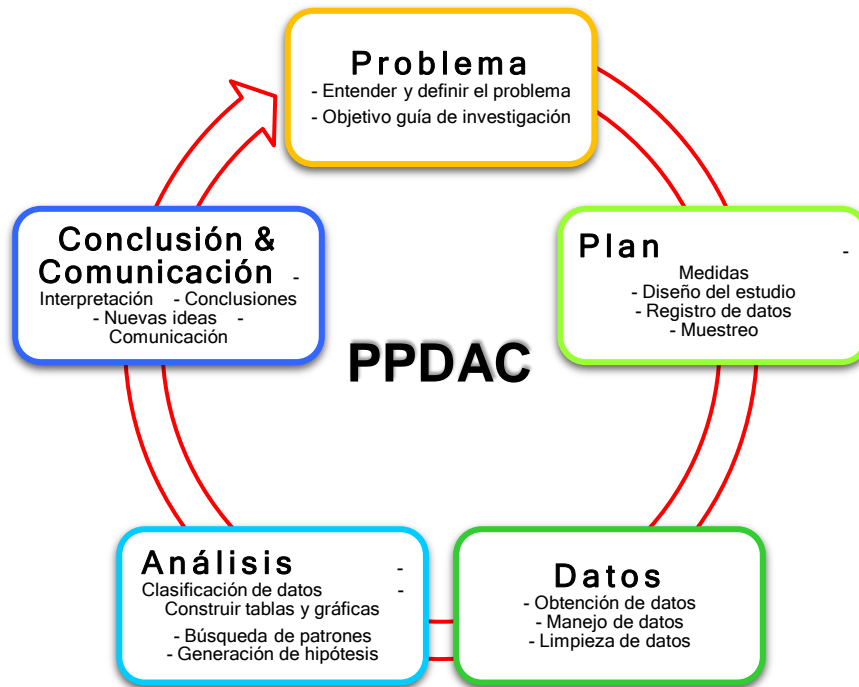
Medición es la asignación de números a objetos o eventos de acuerdo con un conjunto de reglas. La gran variedad de conjuntos de reglas utilizados para medir ha dado lugar a diversas escalas de medición. Dado que medir y analizar todas las ocurrencias de todas las conductas en todos los individuos resulta impráctico, una mejor aproximación consiste en medir una muestra. Una muestra puede definirse simplemente como una parte de una población. Por ejemplo, una población puede componerse de los pesos de todos los niños inscritos en el sistema de educación primaria nacional y si se escoge sólo una fracción de los niños, entonces se tiene únicamente una parte de la población, es decir, una muestra (Daniel, 2006).

Análisis Estadístico Enfocado en la Solución de Problemas

La ciencia estadística tiene una larga y exitosa historia, pero ahora está cambiando a la luz de una mayor disponibilidad de datos desde un enfoque en los métodos matemáticos a uno basado en un ciclo completo de solución de problemas.

Las necesidades de la ciencia de datos y el conocimiento de datos demandan una aproximación más dirigida al problema en la que la aplicación de herramientas estadísticas específicas es vista como sólo un componente de un ciclo completo de investigación. El modelo estadístico consiste en un ciclo de cinco etapas: Problema, Plan, Datos, Análisis y Conclusión y Comunicación (PPDAC; Figura 1); y provee un marco de trabajo adecuado para datos de gran tamaño y complejidad (Spiegelhalter, 2019; MacKay y Oldford, 2000; Wild y Pfannkuch, 1999).

Figura 1. Modelo estadístico PPDAC orientado a la solución de problemas compuesto por el ciclo: Problema – Plan – Datos – Análisis – Conclusión y Comunicación.



Fuente: Modificado de Spiegelhalter, 2019.

Para ejemplificar el uso del modelo estadístico PPDAC en la educación especial, describiremos su aplicación en el análisis estadístico de la adquisición del idioma español escrito como lengua materna en una muestra de niños participantes en el protocolo de investigación de la Unidad de Investigación en Neurodesarrollo (UIND) del Instituto de Neurobiología de la Universidad Nacional Autónoma de México. Este grupo de 64 niños ingresó a la UIND antes de los dos meses de edad presentando factores de riesgo prenatales y perinatales para daño cerebral y recibieron un abordaje multidisciplinario hasta los ocho años por las áreas de neuropediatría, trabajo social, terapia de neurohabilitación, terapia de rehabilitación, nutrición, psicología, neurofisiología, lenguaje y aprendizaje. Los datos sobre la adquisición del

español escrito presentados en este trabajo fueron obtenidos por el área de Aprendizaje.

La adquisición de la escritura en español consta de siete niveles de interpretación fundamentales en los que los niños formulan hipótesis, las ponen a prueba y las reformulan en caso de que resulten insuficientes para interpretar escrituras que el medio les presenta o entren en contradicción con otras hipótesis que ellos mismos hayan construido (Tabla 1; Ferreiro y Teberosky, 1979; Gómez Palacio Muñoz, 1982; Gómez Palacio Muñoz et al., 1996).

Tabla1. Niveles de adquisición de la lengua escrita

Nivel	Hipótesis	Descripción
1		Trazos verticales y horizontales
2		Identificación de su nombre
3		Dibujo de la familia
4	Presilábica	1 Representaciones gráficas primitivas
		2 Escrituras unigráficas
		3 Escrituras sin control de cantidad
		4 Escrituras fijas
		5 Secuencia de repeticiones fijas con cantidad variable
		6 Cantidad constante con repertorio fijo parcial
		7 Cantidad variable con repertorio fijo parcial
		8 Cantidad constante con repertorio variable
		9 Cantidad y repertorio variables
5	Silábica	Una grafía por sílaba
6	Silábica/Alfabética	Coexistencia de hipótesis silábica y alfabética
7	Alfabética	Una grafía por letra del abecedario

Cada etapa del modelo estadístico posee sus propios elementos a comprender y resolver. Una etapa lleva al siguiente y depende de las etapas previas. Por ejemplo, un Plan que no está dirigido a resolver el Problema posee poco valor. Por el contrario, un Plan bien diseñado, puede simplificar el Análisis.

Esta estructura para el método estadístico es útil en dos maneras: primero, provee un modelo para usar activamente datos empíricos y, en segundo lugar, para revisar estudios previos de manera crítica (Mackay y Oldford, 2000).

Problema

Una primera etapa es hacer claras las preguntas, supuestos y expectativas y, sobre todo, interpretar las respuestas que recibimos, ya sea que analicemos estilos de aprendizaje o la efectividad de alguna técnica de enseñanza (Spiegelhalter, 2019). De acuerdo con el filósofo griego de la antigüedad Sócrates, la mitad de la respuesta radica en comprender la pregunta (Gaarder, 2002). El propósito de la etapa Problema en el modelo estadístico PPDAC es proveer una declaración clara del objetivo del estudio. Una estructura bien definida y terminología clara ayudan a traducir el problema contextual de forma que puede guiar el diseño e implementación de las etapas subsecuentes (MacKay y Oldford, 2000).

Unidades y población de estudio. La población de estudio es el total de unidades a estudiar. Para algunos estudios, podría ser más fácil definir las unidades o el total en términos del proceso que los genera. En tales casos, podría ser más conveniente referirse al proceso de estudio más que a la población de estudio. En este estudio de adquisición de la escritura, la población de estudio está constituida por la adquisición de la escritura del idioma español como lengua materna de cada niño. Por tanto, el nivel de adquisición en el que centramos nuestro análisis es el Alfabético. Estamos interesados particularmente en la población mexicana, sin embargo, no se cuenta con datos poblacionales sobre la edad de adquisición de la lectoescritura, aunque se sabe que durante el primer año de educación primaria, es decir a partir de los seis años de edad, se enseña la escritura del español. Datos

reportados en la Estadística Nacional de la Dinámica Demográfica 2018 llevada a cabo por el Instituto Nacional de Geografía, Estadística e Informática (INEGI), indican que, en México, 10.8 millones de niños tiene entre 5 y 9 años de edad. Este rango de edad contiene al rango de 5.02 a 7.69 años de edad de los niños pertenecientes a este estudio de adquisición del español escrito (INEGI, 2020). En cuanto a la educación especial proporcionada en Centros de Atención Múltiple (CAM) y Unidades de Servicios de Apoyo a la Educación Regular (USAER), durante el ciclo escolar 2020-2021, se atendió a un total de 465,272 niños en los niveles preescolar y primaria (Mejoredu, 2022).

Variables. Las variables son características de cada unidad en el proceso y pueden tomar valores numéricos o categóricos. Los valores de variables típicamente difieren de una unidad a otra.

La variable de interés primaria, a la cual llamamos variable respuesta, es el tiempo que toma la adquisición de la escritura de la lengua española en el nivel Alfabético. Existen muchas otras variables a las que podemos llamar variables explicativas ligadas a cada unidad. Por ejemplo, las semanas de gestación.

Atributos de la población. Los atributos de la población son resúmenes que describen las características de la población. Formalmente, un atributo es una función aplicada a toda la población y determinada a través de los valores de la variable a unidades individuales. Los atributos pueden ser numéricos o gráficos. En este estudio, el promedio y la desviación estándar de la edad a la cual los niños presentan el nivel Alfabético, 6.42 ± 0.62 años, constituyen atributos de la población. También gráficas de la edad a la que cada niño logra el nivel Alfabético son atributos de la población (como presentaremos más adelante).

Carácter del problema. El carácter define la naturaleza básica del problema y es causativa, predictiva o descriptiva.

Un problema con un carácter causativo investiga la naturaleza de una relación causativa entre una variable explicativa y una variable de respuesta. Por tanto, un cambio en el valor de la variable explicativa (sin variación en las otras variables explicativas) para todas las unidades de la población resulta en un cambio en el valor de un atributo de interés.

Un problema tiene un carácter predictivo si el propósito es predecir los valores de variables en una o más unidades en la población objetivo.

Un problema tiene un carácter descriptivo si el propósito es estimar o describir uno o más atributos de la población.

En este estudio de adquisición de la escritura, el problema es de carácter descriptivo puesto que se busca estimar el tiempo que toma a los niños adquirir la escritura.

Es importante decidir el carácter en la etapa Problema debido a los requerimientos especiales que puede imponer al Plan (Mackay y Oldford, 2000).

Plan

El propósito de esta etapa es desarrollar un plan para la obtención y análisis de los datos. Ignorar la necesidad de un Plan cuidadoso es una idea tentadora, sin embargo, sólo pueden obtenerse conclusiones confiables de un estudio que ha sido apropiadamente diseñado (Spiegelhalter, 2019).

Se recomienda separar la planeación en varias subetapas, algunas de las cuales se traslaparán inevitablemente. En un uso activo del modelo PPDAC,

podría requerirse cierta iteración dentro de la etapa y entre etapas antes de que un plan satisfactorio sea desarrollado.

Determinación de las unidades y población de estudio. La población de estudio es el total de unidades de estudio para las que los valores de las variables de interés podrían ser determinadas. La diferencia entre los atributos de interés en la población de estudio y los correspondientes atributos en la población blanco es llamada error del estudio. Esta es una simple evaluación cuantitativa para atributos numéricos, pero puede ser complicada para los atributos gráficos. Una correcta estimación del error requiere conocimiento del contexto.

El modelo estadístico PPDAC asegura la consideración de la relevancia de la población de estudio para la población blanco forzando a los investigadores a abordar directamente el error del estudio. Criterios más allá del error del estudio como el costo, conveniencia y consideraciones éticas son también importantes al determinar la población de estudio.

En el caso de nuestro estudio sobre adquisición de la lengua española escrita, se cuenta con las edades de adquisición del nivel Alfabético de lectoescritura de 53 niños nacidos prematuramente participantes en el proyecto de la UIND.

Selección de las variables de respuesta a medir. Las variables de respuesta, correspondientes tanto como sea posible con aquellas usadas para definir los atributos de interés en la población blanco, deben estar claramente definidas.

En este estudio, la variable de respuesta es la edad de adquisición del nivel Alfabético de la lectoescritura.

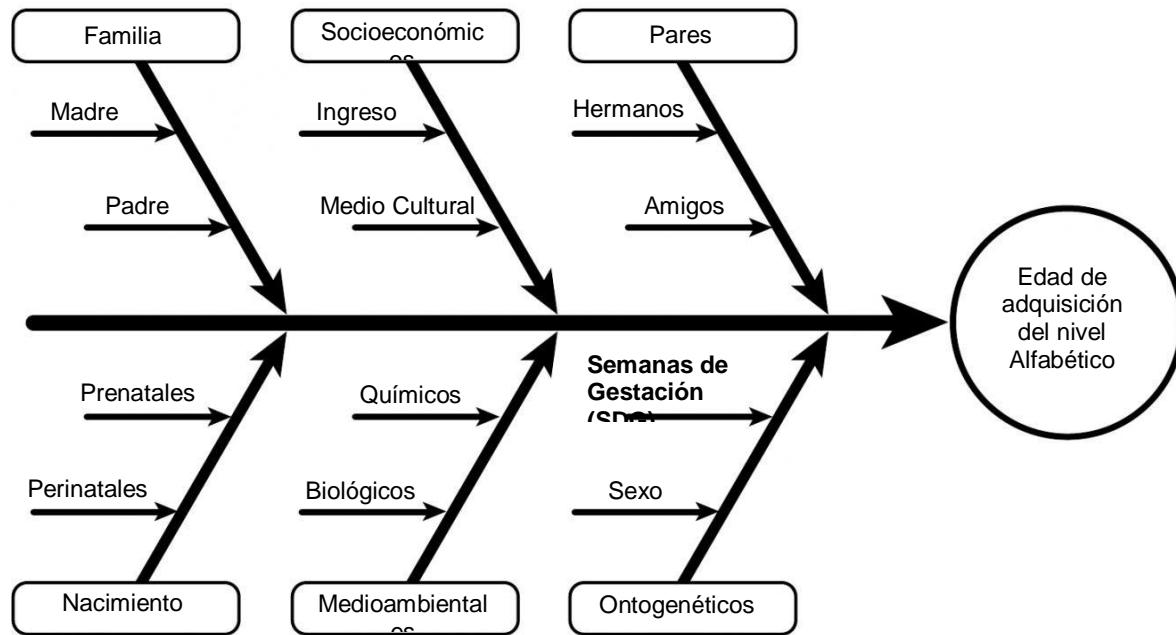
Abordaje de las variables explicativas. En este punto es útil listar todas las posibles variables explicativas que podrían explicar la variación en la respuesta y

organizarlas. Una organización útil es el diagrama de pescado (ver la Figura 2 para el estudio de adquisición de la escritura).

Es importante decidir cómo se abordará las variables explicativas durante la etapa de planeación. Existen tres opciones. La primera consiste en que una variable explicativa puede ser fijada o restringida a un rango de valores de manera que se restringe la población de estudio. La segunda consiste en que una vez que una unidad está en una muestra, el valor de una variable explicativa puede ser elegido deliberadamente o medido para su posterior uso en el análisis. Finalmente, la variable explicativa puede ser completamente ignorada. Este tercer curso de acción es tomado si se sabe por anticipado que la variable explicativa es despreciable (por ejemplo, no explica la variación en las variables de respuesta) o por ignorancia, no reconociendo la presencia o importancia de la variable.

Debido a que nuestro objetivo es describir la asociación entre la edad de adquisición de la lectoescritura y las semanas de gestación de cada niño, decidimos utilizar el número de semanas de gestación como variable explicativa.

Figura 2. Diagrama de pescado del estudio de adquisición de la escritura que considera diversos factores involucrados en la edad de adquisición del nivel alfabético de escritura. Las semanas de gestación (SDG), variable explicativa de este estudio, se encuentra marcada en negritas.



La diferencia básica entre Plan experimental y observacional es sobresaliente en esta etapa. En un Plan experimental, los valores de las variables explicativas correspondientes a factores de interés son determinadas por el experimentador y asignadas a unidades en la muestra. El diseño experimental tradicional provee detalles de la tarea. En un Plan observacional, las variables explicativas no son manipuladas deliberadamente, excepto quizá al restringir la población de estudio o el protocolo de muestreo. Sus valores medidos son usados en el análisis.

Este estudio de adquisición de lectoescritura es observacional ya que no se manipula el número de semanas de gestación.

Los procesos de medición. Un elemento clave del Plan es decidir cómo medir la respuesta seleccionada y las variables explicativas en las unidades de la muestra. Para determinar el valor de cualquier variable en una unidad, llamamos

a los dispositivos de medición, métodos e individuos involucrados el proceso de medición. Una vez que un proceso de medición es especificado, es importante entender sus propiedades. Llamamos error de medición a la diferencia entre el valor de la variable determinada por el proceso de medición y el valor “real”. El error de medición se propaga al Análisis y, por tanto, a la Conclusión y Comunicación.

En muchas aplicaciones, se lleva a cabo un ciclo PPDAC separado más pequeño para investigar los atributos del proceso de medición dentro del estudio completo. Las propiedades del proceso de medición son definidas en términos de medir repetidamente la misma unidad de estudio. Dos conceptos relevantes son el sesgo de medición, un atributo del proceso de medición blanco que describe el error de medición sistemático, y la variabilidad de medición, un atributo del proceso de medición blanco que describe el cambio en el error de medición de una determinación a otra.

La medición de la edad de adquisición de la lectoescritura en nuestro estudio correspondió con las citas programadas a los niños a las que efectivamente asistieron por lo que pudo existir un margen de error de semanas o meses. El cálculo de la edad se llevó a cabo usando fechas de nacimiento y evaluación exactas.

El protocolo de muestreo. Es el procedimiento usado para seleccionar unidades de la población de estudio para medirlas. La meta del protocolo de muestreo es seleccionar unidades que sean representativas de la población de estudio con respecto a el(los) atributo(s) de interés. El protocolo de muestreo aborda cómo y cuándo se selecciona a las unidades y cuántas unidades se selecciona.

Llamamos error de muestra a la diferencia entre el atributo de interés en la población de estudio y el correspondiente atributo en la muestra. Como sucede en los procesos de medición, puede haber sesgo y variabilidad asociados con el protocolo de muestreo. Estos son propiedades del protocolo y no de alguna muestra de unidades particular. Como sucede con el proceso de medición, el sesgo y la variabilidad de muestreo son definidos en términos de las propiedades del error de la muestra cuando el protocolo de muestreo se aplica repetidamente a la población de estudio. Estas replicas son siempre hipotéticas, lo que significa que el sesgo y variabilidad de muestreo pueden ser descritos sólo a través de un modelo del protocolo de muestreo. En el uso activo del modelo PPDAC, se utiliza modelos matemáticos para el protocolo de muestreo potencial para abordar situaciones como la determinación del tamaño de la muestra.

La muestra se recabó por conveniencia en niños que asisten a la UIND con factores de riesgo para daño cerebral prenatales y perinatales, siendo la prematuridad el más común. En Querétaro la población de niños entre 5 y 8 años de edad al 2020 fue de 158,622 (INEGI, 2020), por lo que considerando que el 10% de los niños nace prematuramente en Querétaro (Quadratín, 2022), la muestra arroja menos de 14% de error con un 95% de confianza.

El protocolo de obtención de datos. El protocolo de obtención de datos es el procedimiento para ejecutar los pasos antes descritos del Plan para obtener y registrar los datos. Aborda situaciones administrativas y de gestión como quién hace qué y cuándo. También incluye un plan para monitorear los datos a medida que se obtienen para asegurar la calidad.

Este paso incluye la consideración del ingreso de datos, estructura de archivos, software para análisis entre otros, especialmente para Planes en los que se acumulará una gran cantidad de datos (Mackay y Oldford, 2000).

Para el registro de datos, se conservan documentos (dibujos y escritos de los niños con frases y narraciones) y se lleva a cabo el vaciado de la información en una base de datos del Área de Aprendizaje de la UIND posterior a cada cita mediante una lista de cotejo que incluye los hitos y habilidades relevantes para la adquisición de la lectoescritura en cada nivel.

Datos

El propósito de la etapa Datos es ejecutar el Plan y asegurar la calidad de los datos como preparación para su Análisis. Obtener buenos datos requiere habilidades de organización y programación, cada vez más importantes en la ciencia de datos, particularmente para datos que podrían necesitar mucho trabajo de limpieza antes de que estén listos para ser analizados. La frase “datos encontrados” comunica claramente que la extracción de datos fue complicada, como si los datos hubieran sido encontrados al azar; esto último sucede si los datos no fueron adecuadamente registrados y organizados desde su obtención (Spiegelhalter, 2019).

Hacia el final de la etapa Plan, debe tenerse alguna noción de valores claramente discordantes para las variables. Monitorear los valores registrados de los datos a medida que se obtienen durante la ejecución del Plan es importante para asegurar su calidad y hacer cambios a los procedimientos según se requiera.

En nuestro estudio, al momento de registro de fechas y datos se corrobora con los documentos generados por los niños, en los cuales también se registra lo reportado verbalmente por el niño (lectura de su propia escritura), así como fecha y determinación de nivel de adquisición de la lectoescritura.

Evaluación de los datos. Aquí, se evalúa la consistencia interna de los datos como un todo, con la intención de asegurar la calidad de los datos para análisis

subsecuentes. Los datos son examinados en busca de patrones y rasgos inesperados. Una vez que los patrones han sido identificados, existen tres posibles decisiones: ignorarlos, rehacer las etapas Plan y Datos, o más probablemente trasladar la información para que sea manejada en la etapa Análisis.

El registro en la base de datos del Área de Aprendizaje de la UIND ha permitido un registro ordenado de los datos para cada niño con el que es posible realizar una inspección visual clara de la plausibilidad de los datos; además, el cálculo de las edades mediante fórmulas permite una comparación de las fechas y la visualización de las diferencias individuales que cada niño presenta en cada uno de los niveles del proceso de adquisición de la lectoescritura.

Almacenamiento de datos para análisis subsecuentes. En la actualidad, en ocasiones se requiere considerar mucho el tipo de medio y la definición y organización de la estructura de datos usada para almacenarlos (Mackay y Oldford, 2000). En el caso de nuestro estudio, los datos no requieren de gran espacio digital por su registro. Sin embargo, la digitalización de los documentos sí requiere mayor espacio en disco.

Análisis

El propósito de la etapa Análisis es usar los datos obtenidos y la información del Plan para abordar las preguntas formuladas en el paso Problema. La forma y formalidad del Análisis depende de varios factores, los cuales son los siguientes: la complejidad del Problema y el Plan, la habilidad del analista, la cantidad de variabilidad inducida por el Plan, y el público al que está dirigido el estudio. Se propone la siguiente división general de la etapa:

- Construcción de un modelo para Plan y Datos

- Ajuste y evaluación del modelo
- Uso del modelo final para abordar el Problema

Un primer paso, anterior a la construcción del modelo, es una visualización útil de los datos obtenidos (Spiegelhalter, 2019). Un modelo estadístico describe el comportamiento de las variables de respuesta medidas para las unidades incluidas en la muestra si ejecutamos repetidamente el paso Datos de acuerdo con el Plan. El modelo refleja propiedades de la población de estudio, el protocolo de muestreo y el sistema de medición utilizados. El modelo también incluye el efecto de las variables explicativas medidas sobre la variable de respuesta.

Una vez que un modelo inicial es postulado, herramientas de evaluación de ajuste y modelo pueden ser usadas para sugerir refinamientos al modelo. Este proceso iterativo continúa hasta que el modelo es consistente con la estructura interna de los datos obtenidos y la información conocida sobre el protocolo de muestreo y los sistemas de medición. El modelo final es usado para estimar atributos de interés en la población de estudio y evaluar la incertidumbre debida a los errores de muestreo y medición.

Ya que la cantidad de datos a analizar puede ser grande, es recomendable considerar la construcción de uno o más algoritmos que permita(n) repetir el Análisis sólo con los refinamientos sugeridos. Un algoritmo es una regla o fórmula que toma variables de entrada y produce un resultado, tal como una predicción, una clasificación o una probabilidad con base en la relación entrada/salida requerida en la etapa Problema (Mackay y Oldford, 2000; Cormen et al., 2009).

El modelo estadístico empleado para el análisis del proceso de adquisición de la lectoescritura hasta el momento únicamente implica análisis de correlación y regresión entre las edades de adquisición del nivel Alfabético de lectoescritura

y las semanas de gestación de cada niño. Para la elaboración de un modelo predictivo más completo se requiere incorporar más datos médicos y del neurodesarrollo.

Conclusión y Comunicación

El propósito de la etapa Conclusión y Comunicación es reportar los resultados del estudio en el idioma del Problema. Debe usarse resúmenes numéricos concisos y gráficas de presentación para aclarar la discusión. Además, la Conclusión y Comunicación provee una oportunidad de discutir las fortalezas y debilidades del Plan, Datos y Análisis especialmente respecto a posibles errores que puedan haber surgido. La clasificación de los errores provee una estructura para esta parte de la discusión (Mackay y Oldford, 2000).

La clave para una buena ciencia estadística es extraer conclusiones apropiadas que reconozcan en su totalidad las limitaciones en la evidencia y las comuniquen claramente. Toda conclusión generalmente desencadena más preguntas (Spiegelhalter, 2019).

El análisis de la variable de respuesta y la variable explicatoria no arrojó resultados significativos en cuanto a su valor predictivo y la correlación es baja. Por tanto, se requerirá que en futuros estudios se incorporen más variables explicatorias al modelo como el sexo, el nivel educativo de la madre, el número de sesiones de terapia administradas y la presencia de lesiones cerebrales en los niños posteriores al parto.

Contando una historia con los datos

Las gráficas deben ser elegidas con cuidado y consciencia de su impacto. La comunicación visual mediante datos requiere seis importantes pasos (una guía paso a paso es proporcionada en Nussbaumer Knaflic, 2015):

1. Entender el contexto situacional. Implica a la audiencia, el mecanismo de comunicación y el tono deseado. Crear una comprensión robusta del contexto situacional reduce las repeticiones en la elaboración del material visual.
2. Elegir un arreglo visual efectivo. Requiere una visión amplia de los tipos de gráficos existentes y el enfoque que cada uno permite comunicar. Es más adecuado evitar gráficos visualmente atractivos pero que no comunican la información adecuadamente, como los gráficos en tercera dimensión. }
3. Eliminar el desorden. Cada elemento que se añade a una página o pantalla implica una carga cognitiva en la audiencia. Por este motivo es necesario considerar los principios de la Gestalt de la percepción visual y aplicarlos a la presentación de tablas y gráficas. También se requiere considerar la alineación, el uso estratégico de espacios en blanco y el contraste.
4. Enforzar la atención de la audiencia donde se requiere. Es necesario considerar los mecanismos de percepción visual y memoria que enmarcan atributos preatencionales como tamaño, color y posición en la página. Estos atributos crean una jerarquía visual de los componentes y permiten dirigir la atención de la audiencia hacia la información que se desea comunicar.
5. Pensar como un diseñador. “La forma sigue a la función” es un adagio del diseño de productos que tiene una clara aplicación a la comunicación con datos. Lo que queremos comunicar (función) nos permite crear una presentación visual (forma) que logre dicha comunicación.

6. Contar una historia. Las historias resuenan y permanecen en la memoria en formas que los datos por sí mismos no lo hacen. Una historia debe tener un inicio, desarrollo y final claros. Estrategias como repetición y flujo narrativo permiten una comunicación clara.

A continuación, se presentan figuras que ilustran la distribución de las edades de adquisición del nivel Alfabético de la muestra y la relación entre la edad de adquisición y las semanas de gestación (Figuras 3 y 4).

Figura 3. Histograma del número de casos por rango de edad de adquisición del nivel Alfabético de la lectoescritura en la muestra.

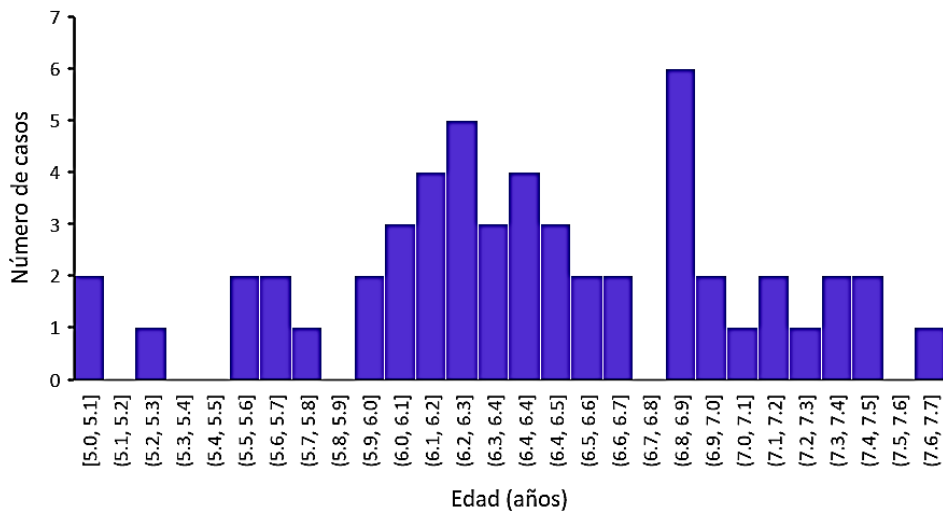
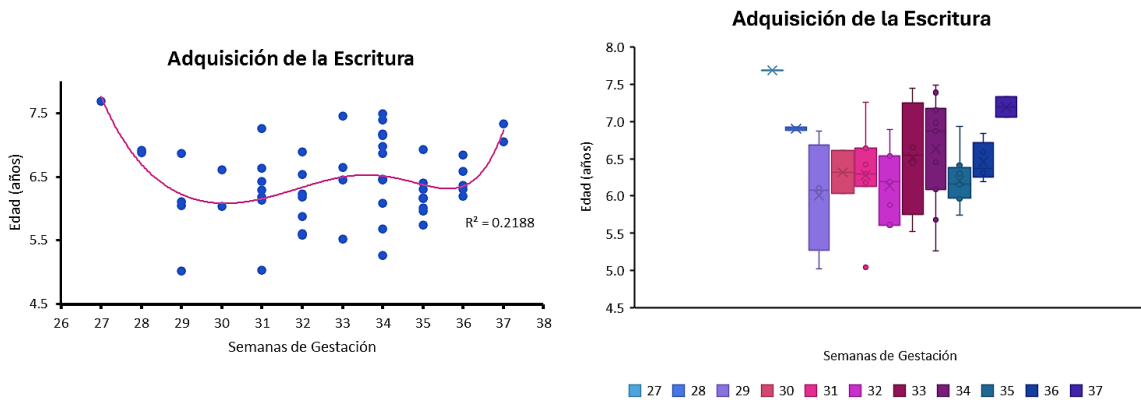


Figura 4. Gráfico de dispersión (izquierda) y gráfico de caja (derecha) de las edades de adquisición del nivel Alfabético de la lectoescritura respecto a las semanas de gestación de los niños. La línea para el análisis de regresión corresponde al ajuste máximo, con una relación polinómica de orden 6.



Conclusiones

Con frecuencia, la estadística se presenta como un conjunto de herramientas de análisis. Sin embargo, como se muestra en la presentación del modelo para la solución de problemas PPDAC, el análisis es sólo la cuarta etapa de una serie de cinco que constituye al método estadístico. Las primeras tres etapas previas al análisis son críticas para que las conclusiones obtenidas a partir de los datos sean adecuadas. El modelo estadístico PPDAC puede ser aplicado a la educación especial independientemente del nivel de sofisticación matemática que se utilice. Se puede abordar problemas interesantes y sustantivos sin recurrir a herramientas de análisis complejas, conjuntos de datos enormes o grandes recursos computacionales. Lo que se requiere es un contexto amplio que describa detalles de la estructura de los datos y el tema de estudio (MacKay y Oldford, 2000). Este marco de trabajo permite considerar los elementos importantes para la obtención de información a partir de datos obtenidos en el aula, como es ilustrado mediante el ejemplo de la adquisición de la lectoescritura del español como lengua materna.

El uso de los resultados de pruebas del desarrollo y evaluaciones de la enseñanza, además de permitir la retroalimentación a alumnos y cuidadores

primarios, se extiende a la obtención de datos con los que es posible obtener información sobre la efectividad de los métodos y contenidos de enseñanza, su uso y la pertinencia de las propias evaluaciones. Un uso informado por los datos obtenidos a partir de la estadística permite una adecuación continua específica para las habilidades y conocimientos que se requiere facilitar en los niños atendidos en la educación especial.

Referencias bibliográficas

- Álvarez Ramos, R. (2007). Evaluación de diagnóstico para determinar el nivel de aprendizaje de lengua escrita y las matemáticas. Primero y segundo grado. Resultados prácticos y planteamiento de problemas. En: *Avances, propuestas y soluciones*. 2ª Reunión Regional Norte, Centro América y Caribe de Evaluación Educativa.
- Cormen, T.H.; Leieserson, C.E.; Rivest, R.L.; y Stein, C. (2009). *Introduction to algorithms*. 3rd Ed. Cambridge: MIT Press.
- Daniel, W.W. (2006). *Bioestadística: Base para el análisis de las ciencias de la salud*. 4ta Ed. México: Limusa Willey.
- Ferreiro E., Gómez Palacio M., Guajardo E., Rodríguez B., Vega A. y Cantú L. (1979). *El niño preescolar y su comprensión del sistema de escritura*. O.E.A.
- Ferreiro E., y Teberosky A. (1979). *Los sistemas de escritura en el desarrollo del niño*. Siglo XXI.
- Gaarder, J. (2002). *El mundo de Sofía*. Grupo Editorial Patria.
- Gómez-Palacio Muñoz, M. (1982). *Propuesta para el aprendizaje de la lengua escrita*. Secretaría de Educación Pública.
- Gómez-Palacio, M., Villareal M.B., González, L.V., López Araiza, M.L., Jarillo, R. (1996). *El niño y sus primeros años en la escuela*. Secretaría de Educación Pública.
- Good, T.L. y Brophy, J. (1996). *Psicología educativa contemporánea*. 5ta Ed. McGraw-Hill.

- Instituto Nacional de Estadística y Geografía [INEGI]. (2020). *Estadísticas a propósito del día del niño*. Comunicado de prensa núm. 164/20. Págs. 1-2.
- MacKay, R.J., y Oldford, R.W. (2000). "Scientific Method, Statistical Method and the Speed of Light." *Statist. Sci.*, 15(3), 254 – 278.
- Mejoredu. (2022). *Datos sobre los servicios de educación especial en México*. Educación en Movimiento. Año 1, número 9. Págs. 10-12.
- Nussbaumer Knaflig, C. (2015). *Storytelling with data*. John Wiley & Sons.
- Quadratín. (2021). En Querétaro, uno de cada 10 nacimientos es prematuro. Nota periodística del 23 de noviembre de 2021. <https://queretaro.quadratin.com.mx/en-queretaro-uno-de-cada-10-nacimientos-es-prematuro/>
- Romero Contreras, S., y García Cedillo, I. (2013). Educación especial en México. Desafíos de la educación inclusiva. *Revista Latinoamericana de Educación Inclusiva*, 7(2), 77-91.
- Spiegelhalter D. (2019). *The art of statistics: How to learn from data*. BasicBooks.
- Wild, C.J. and Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *Internat. Statist. Rev.*, 67, 223–248.

Parte II:

ESTADÍSTICA DESCRIPTIVA



IMPLICACIÓN DE LAS MEDIDAS DE TENDENCIA CENTRAL Y DE DISPERSIÓN EN LA EDUCACIÓN ESPECIAL

Carolina Sotelo Tapia
Marisela Hernández González [‡]
Fabiola Alejandra Iribe Burgos
Abril Zagnitte Gómez Méndez

Resumen

En el área de ciencias de la salud, cuando se realiza una investigación científica es muy importante el uso de métodos estadísticos para analizar e interpretar resultados. Para fines de este capítulo, nos referimos a la estadística descriptiva como una herramienta encargada de desarrollar un conjunto de técnicas y mediciones cuantitativas cuyo objetivo es presentar y resumir las características de los datos observados, mediante la ordenación y el análisis, que posteriormente se representan gráficamente.

Se ha descrito que la estadística descriptiva se conforma de dos grandes medidas estadísticas, de tendencia central y de dispersión, cada una compuesta de sus respectivas mediciones. En la primera, para describir las características de los datos, se suele observar que el valor de las puntuaciones se enfoca en el centro de la distribución, y que en consecuencia representan un centro en el que se encuentran ubicados el conjunto de datos. Dentro de las medidas de tendencia central, se conocen principalmente tres: la moda, la mediana y la media.

[‡] Instituto de Neurociencias, Centro Universitario de Ciencias Biológica y Agropecuarias, Universidad de Guadalajara. Francisco de Quevedo 180, Arcos Vallarta, 44130 Guadalajara, Jal. marisela.hgonzalez@academicos.udg.mx

Por otro lado, se entiende a las medidas de dispersión como la representación de la variabilidad de los puntos alrededor de uno central. Y, se han descrito cuatro medidas: la desviación media, la varianza, la desviación y el error estándar.

De acuerdo con lo anterior, en el capítulo se resalta el impacto que tiene el uso de la estadística descriptiva en los reportes de proyectos y en el desarrollo de nuevos proyectos de investigación sobre la educación especial.

Palabras Clave: Estadística descriptiva, Medidas de tendencia central y Medidas de dispersión.

Introducción

La estadística permite interpretar los resultados que, en un principio, podrían parecer desordenados y sin un sentido aparente. Es una herramienta que nos permite recolectar, organizar, analizar, interpretar, y representar gráficamente una serie de datos. Ésta, mediante una serie de cálculos realizados, proporciona información sobre las variables estudiadas de manera ordenada y clara (Runyon y Haber, 1992; Montero-Lorenzo, 2007).

Como parte de la estadística, se desglosa la estadística descriptiva, la cual, se ha mencionado en la literatura científica, se encarga de desarrollar un conjunto de técnicas y mediciones cuantitativas cuyo objetivo es presentar las características de los datos observados de manera resumida, que se realiza mediante la ordenación de los datos en tablas o tabulaciones, y que posteriormente se representan gráficamente (Runyon y Haber, 1992; Kirk, 2008).

Parecería que la estadística descriptiva solo es usada en ciencias exactas debido a la cantidad de números y, por consecuencia, fórmulas y cálculos que

se utilizan en ellas; sin embargo, el uso de esta herramienta también es ampliamente utilizada en algunas otras ciencias (García-Cedillo, y Romero-Contreras, 2019). Centrando esta explicación en ciencias de la salud, podemos ver que cuando se quiere esclarecer el conocimiento de un fenómeno, la estadística permite explicar lo que pasa con los casos individuales y su comportamiento sobre las características compartidas con el resto de los datos. Es decir, aquellas características que comparten cada integrante del grupo, y que puede explicarse desde una unidad de variabilidad entre ellos (Fernández, Sánchez, Córdoba, y Largo, 2002).

Entonces, repasando un poco lo antes mencionado, debido a que la estadística trabaja solo con un conjunto de datos u observaciones, es siempre necesario resumirlos, para así, extraer de manera más accesible la información, y así, lograr extraer los puntos más significativos de los datos analizados (Parra, 1995). De ahí, la importancia de la estadística descriptiva y su aporte en el manejo de la información que nos proveerá conclusiones a partir del conjunto de datos que, de otra manera, serían confusos y con poca relevancia aparente (Bologna, 2013). Un ejemplo del uso de la estadística en ciencias de la salud puede ser una investigación de educación especial en donde se realizó una prueba de atención selectiva a un grupo de niños, cuyos datos que se obtuvieron de la observación directa no fueron suficientes para conocer de manera global el fenómeno de estudio, de modo que, el uso de la estadística descriptiva sería de gran ayuda para lograrlo.

De acuerdo con lo anterior, en el presente capítulo se abordarán en orden didáctico los temas de las medidas de tendencia central, de dispersión y de posición, y se definirá brevemente cada una. Así mismo, se presentará el cálculo de cada una, así como ejemplos simples del área de ciencias de la salud.

Ciertamente, este capítulo puede fungir como una guía de estudio para el educador, estudiante, y personas interesadas en el tema, ya que se describe paso a paso el cálculo de las fórmulas de cada una de las medidas estadísticas que se presentan. Todo con la finalidad de esclarecer el entendimiento y comprensión del uso de la estadística descriptiva con fines prácticos.

Medidas de tendencia central

Cuando se quiere iniciar con el análisis de un conjunto de datos en una investigación científica podemos empezar conociendo los datos obtenidos con medidas estadísticas básicas que nos ayudarán a resumir lo que pasa de un solo dato en un conjunto de datos. Estas se llaman medidas de tendencia central (MTC) o también llamadas de resumen, las cuales, se han establecido como indicadores que tienden a sintetizar o describir las características de una muestra y de una población, ubicando datos dentro de una distribución (Cebrián, 2009).

Como ya se mencionó, en la distribución muestral se suelen distribuir los valores para identificar las puntuaciones que tienden a ubicarse en el centro de la distribución, y que por consecuencia se les denominan medidas de tendencia central, y que representan un punto en el que se encuentran situados el conjunto de datos (Ricardi, 2011). Sin embargo, para referirse más detalladamente a esta centralización de datos de la que se ha estado hablando anteriormente, muchos estadísticos han preferido usar términos más precisos. Dentro de los cuales, se han mencionado principalmente tres medidas, como son: la moda, la mediana y la media (Runyon y Haber, 1992; Kirk, 2008).

Moda

La moda es el valor o categoría cuantitativa/cualitativa que se presenta con mayor frecuencia. Una limitación particular de la moda es que, podría “no existir”, ya que ningún valor asignado podría repetirse. Cuando una distribución tiene dos puntuaciones con la misma frecuencia máxima no puede designarse una moda, por lo que se suele reportar que la distribución es bimodal, derivado de las dos frecuencias máximas (Runyon y Haber, 1992).

Dentro de las ventajas que tiene la moda, es que es una medida descriptiva que es fácil determinar por mera observación, cuando el conjunto de datos es pequeño (Cebrián, 2009). Entonces, cuando la distribución de frecuencias no es agrupada, se analiza más por inspección visual que por cálculo como se muestra en el siguiente ejemplo.

Enseguida se muestra un ejemplo del uso de la moda en un caso de niños con TDAH:

Se ha descrito que la hiperactividad es una de las alteraciones más frecuentes en niños, y que es uno de los síntomas más comunes de Trastorno del déficit de atención e hiperactividad (TDAH). Una de las conductas que se han utilizado como criterio para identificarla en niños es la dificultad para mantener la atención durante ciertos periodos (Ávila y Polaino-Lorente, 2002; Candia, Sepúlveda, y Orozco, 2015). Lo que podría estar afectando el desempeño escolar de los niños. A continuación, se muestra una serie de datos sin intervalos, de calificaciones generales del ciclo escolar de 10 niños con TDAH en donde se identifica la moda:

5, 6, 6, 8, 7, 6, 5, 7, 8, 6

Lo que se observa en el ejemplo es que la calificación que más se repite en estos niños es el 6. Lo que podría indicar que, aunque los niños con TDAH no reprobaron, están muy cerca de hacerlo. Esta descripción podría servirnos como base para posteriores investigaciones acerca de este trastorno.

En contraste, cuando la distribución de frecuencias es agrupada y en intervalos, la moda se designa como el punto medio del intervalo con la mayor frecuencia (Kirk, 2008).

Y la fórmula de la moda para datos agrupados en intervalos se realiza de la siguiente manera:

$$M_o = L_i + \frac{f_i - f_{i-1}}{f_i - f_{i-1} + f_i - f_{i+1}} \cdot A_i$$

En donde:

- M_o : moda
- L_i : límite inferior del intervalo en el cual se encuentra la moda.
- f_{i-1} : frecuencia absoluta del intervalo anterior en el que se encuentra la moda.
- f_i : frecuencia absoluta del intervalo en el que se encuentra la moda.
- f_{i+1} : frecuencia absoluta del intervalo siguiente en el que se encuentra la moda.
- A_i : amplitud del intervalo en el que se encuentra la moda.

Ejemplo del uso de la moda en datos agrupados en intervalos:

Por ejemplo, en la tabla 1.1.1 se muestra una serie de datos agrupados en intervalos y la moda es 100, que es el punto medio del rango entre 90-109 que se presenta con mayor frecuencia.

Tabla 1.1.1: Distribución de frecuencias agrupadas en intervalos de puntuaciones de cociente intelectual.

Intervalo de cociente intelectual	f
130 - 139	3
120 - 129	5
110 - 119	15
90 - 109	47
80 - 89	14
70 - 79	6
50 - 69	2
35 - 49	0
20 - 34	0

Nota: La tabla muestra el intervalo de puntuaciones de cociente intelectual en la primera columna y la frecuencia (f) en la segunda columna.

Pero ¿cómo llegamos a ese resultado? Esto se realizó de la siguiente manera:

Se tiene que encontrar el intervalo en el cual se encuentra la moda, es decir, el intervalo con mayor frecuencia absoluta. Por lo que, se puede observar que el intervalo 4, tiene la mayor frecuencia absoluta (47), por lo tanto, aquí se encontrará la moda.

$$M_o = 90 + \frac{47 - 15}{47 - 15 + 47 - 14} \cdot 20$$

Después se realizan las sumas y restas de las frecuencias como indica la fórmula.

$$M_o = 90 + \frac{32}{65} \cdot 20$$

Finalmente, dará como resultado la moda.

$$M_o = 100$$

Mediana

En cuanto a la mediana, se define como, el valor intermedio de una serie de datos. En otras palabras, dicha serie de datos es dividida en dos grupos con la misma frecuencia.

En primera instancia, para calcular la mediana, es necesario darles a nuestros datos un orden de magnitud, ya sea de mayor a menor o de menor a mayor (Triola, 2009).

En algunas ocasiones la serie de datos puede presentarse par o impar. En el caso de que la serie de datos sea par, la mediana se obtiene a través de la media de los dos números que se encuentran a la mitad de nuestra serie de datos. Mientras que, cuando la mediana es impar, la mediana es el número que se encuentra exactamente a la mitad de la lista de datos.

A continuación, para determinar la mediana a partir de una evaluación de la atención de un grupo de 15 niños de siete años, mediante la prueba de cancelación (CL), donde nuevamente sus puntuaciones escalares fueron las siguientes:

6, 5, 7, 7, 6, 6, 6, 6, 5, 4, 8, 3, 5, 5, 7

Primero es necesario ordenar los valores, en este caso están ordenados de menor a mayor.

~~3, 4, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 8~~

Ya que el número de valores es impar (15), la mediana será el valor que está exactamente a la mitad de la lista ordenada: 6.

El siguiente ejemplo considera un grupo de valores pares:

Se desea evaluar el grado de impulsividad de un grupo de 12 niños de 10 años déficit de atención, a través de la prueba de búsqueda de símbolos (BS), del WISC-IV, en la cual, los niños deben indicar en el menor tiempo posible la semejanza entre uno o varios de los símbolos que se presenten. Los tiempos empleados en segundos que presentaron los pacientes son los siguientes:

100, 99, 101, 105, 100, 98, 107, 110, 100, 110, 105, 108

Nuevamente es necesario ordenar los valores, en este caso están ordenados de mayor a menor.

110, 110, 108, 107, 105, 105, 101, 100, 100, 100, 99, 98

Debido a que el número de valores es par (12), la mediana se obtiene calculando la media de los dos valores intermedios: 105 y 101.

$$\text{Mediana} = \frac{105+101}{2} = \frac{206}{2} = 103$$

Ya que el número de valores es par (12), la mediana se determina a partir de la media de los dos valores intermedios.

Debido a que la mediana tiene una menor sensibilidad a los valores extremos en comparación con la media, que sí considera todos los valores para su cálculo; frecuentemente la mediana es utilizada para conjuntos de valores que cuentan con un número relativamente pequeño de valores extremos (Triola, 2009).

Asimismo, para calcular la mediana en datos con intervalos, hay que seguir dos pasos:

- Primero, se debe encontrar el intervalo en el que se encuentra la mediana utilizando la siguiente fórmula:

$$posición = \frac{n + 1}{2}$$

- Posteriormente, se debe utilizar la fórmula de la mediana:

$$M_e = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot A_i$$

En donde:

- M_e : la mediana del conjunto de datos
- L_i : es el límite inferior del intervalo en donde se localiza la mediana.
- n : es el número total de datos, es decir, es la sumatoria de las frecuencias absolutas.
- F_{i-1} : es la frecuencia acumulada del intervalo anterior, en la que se localiza la mediana.
- f_i : es la frecuencia absoluta del intervalo en el que se localiza la mediana.
- A_i : es la amplitud del intervalo en el que se localiza la mediana.

Ejemplo del uso de la mediana en datos agrupados en intervalos:

Se obtuvieron datos de puntaje acumulado de 50 personas que realizaron una tarea de toma de decisiones diseñada por Cortes et al. (2021). De manera breve, la tarea consta en que el participante acumule la mayor cantidad de puntos posibles eligiendo de un banco de seis estímulos neutros, con los que se hicieron 40 permutaciones de tres estímulos cada una, para formar los ensayos de la

tarea. Cada estímulo tuvo un valor (0, 2, 4, 6, 8 y 10) que se mantuvo fijo durante la tarea.

Figura 1.2.1: Distribución en intervalos de puntaje acumulado.

Puntaje acumulado	Frecuencia absoluta	Frecuencia acumulada
60 – 79	3	3
80 – 99	5	8
100 – 119	18	26
120 – 139	9	35
140 – 159	10	45
160 – 179	2	47
180 – 199	1	48
200 – 219	1	49
220 – 239	1	50

Nota: La tabla muestra los intervalos de puntaje acumulado en la primera columna, frecuencia absoluta en la segunda columna y la frecuencia acumulada en la tercera columna (que es la suma de por renglón de la frecuencia absoluta y acumulada, esta última se refiere al resultado de sumar sucesivamente las frecuencias absolutas, basado al orden en el que se encuentren en la tabla).

Se utiliza la fórmula para identificar el intervalo en el que está la mediana:

$$posición = \frac{50 + 1}{2} = 25.5$$

Este valor, se debe de buscar en la columna de frecuencias acumuladas. Si no encontramos, buscamos el valor que sigue, es decir, después del 25.5 sigue el 26. Por lo tanto, la mediana se ubica en el intervalo 3 (100 – 119).

$$M_e = 100 + \frac{\frac{50}{2} - 8}{18} \cdot 20$$

$$M_e = 100 + \frac{25 - 8}{18} \cdot 20$$

$$M_e = 100 + \frac{17}{18} \cdot 20$$

$$M_e = 118.888$$

Entonces, el valor de la mediana del puntaje acumulado de la tarea de toma de decisiones es 118.88, es decir, del tercer intervalo (100 – 119) el valor intermedio de todos los datos es el 119.

Media

Dentro de las medidas de tendencia central, la medida más utilizada e importante es la *media aritmética*, también conocida como *promedio*. Definida como la suma de una serie de datos, dividida entre el número total de los mismos datos (Triola, 2009).

La media es el valor más representativo de un conjunto de datos, y se dice que es la medida que determina su centro de gravedad (Cebrián, 2009).

Cuando se habla de la media es necesario hacer una distinción entre la media de una población o la media muestral. La media de una población se representa mediante la letra mu minúscula del alfabeto griego “ μ ”. Mientras que, la media muestral (conjunto de datos provenientes de una población), está representada por el símbolo de equis barra “ \bar{x} ” (Devore, 2019).

La fórmula para obtener la media es la siguiente:

$$\bar{x} = \sum_{i=1}^n x_i$$

En dicha fórmula, Sigma (Σ), representa la sumatoria de un conjunto de valores; x , es el símbolo comúnmente utilizado para representar los datos individuales. Si se van a considerar los datos de una muestra, n , representa a dicha muestra. Mientras que la letra i minúscula, está representando cualquier valor. De esta manera, la formula $\bar{x} = \frac{\Sigma x}{n}$, nos indica la media de un conjunto de valores muestrales (Kirk, 2008).

A continuación, se proporciona una serie de casos que ejemplifican el uso de la media:

Se evaluó la atención de un grupo de 15 niños de siete años, mediante la prueba de cancelación (CL), la cual es una prueba neuropsicológica para evaluar la capacidad atencional y que básicamente consiste en que el experimentador le presenta una serie de letras o figuras al participante, y este tiene que tachar la letra o figura que el técnico le determine. En esta prueba se registran los errores de comisión (tachar letras no indicadas) o error por omisión (no tachar letras indicadas) (de la Torre Benítez y de la Torre, 2003).

Para este ejemplo tomaremos en cuenta solo los errores de comisión y se desea obtener la media de dicho grupo. Por lo que, primero debemos de saber el puntaje de los niños. Sus puntuaciones fueron las siguientes:

6, 5, 7, 7, 6, 6, 6, 6, 5, 4, 8, 3, 5, 5, 7

Ahora bien, en primera instancia, debemos sumar cada uno de los datos que conforman nuestra muestra. Y posteriormente dividir el resultado de la suma entre el número total de datos, que en este caso sería 15.

$$\begin{aligned}\bar{x} &= \frac{\Sigma x}{n} = \frac{6 + 5 + 7 + 7 + 6 + 6 + 6 + 6 + 5 + 4 + 8 + 3 + 5 + 5 + 7}{15} = \frac{86}{15} \\ &= 5.733\end{aligned}$$

La media de las puntuaciones en la prueba CL para medir la atención en el grupo de niños es de 5.733.

A continuación, un segundo ejemplo:

Se desea evaluar el grado de impulsividad de un grupo de 12 niños de 10 años con déficit de atención, a través de la prueba de búsqueda de símbolos, del WISC-IV, en la cual, los niños deben indicar la semejanza entre uno o varios de los símbolos que se presenten. Los tiempos empleados en segundos que presentaron los pacientes son los siguientes:

100, 99, 101, 105, 100, 98, 107, 110, 100, 110, 105, 108

Recordemos que en primer lugar se deben sumar todos nuestros valores.

$$\bar{x} = \frac{\sum x}{n} = \frac{100+99+101+105+100+98+107+110+100+110+105+108}{12} = \frac{1243}{12} = 103.58$$

En este caso la media del tiempo en segundos empleado por los niños para resolver la prueba de búsqueda de símbolos fue de 103.58 segundos.

Fuera de los ejemplos anteriores, también es importante mencionar que dentro de las principales desventajas de la media es que cada valor es muy sensible. Es decir, con un solo valor que salga de la generalidad se puede ver afectado el resultado de la media de manera drástica (Triola, 2009). Sin embargo, la medida de tendencia central que resuelve dicha desventaja es la mediana (Triola, 2009).

Por otro lado, dentro de las ventajas, su cálculo es fácil de hacer, es la medida más conocida, y es la base para el cálculo de muchas otras mediciones estadísticas (Cebrián, 2009).

Un valor sólo adquiere relevancia una vez que es comparado con otros valores, por lo que conocer la media y la moda es importante pero sólo si con ellas podemos determinar si un valor es menor o mayor que éstas y, qué tan menor o mayor es, y para saber esto tenemos a las medidas de dispersión.

Medidas de dispersión

Una vez que se han observado alrededor de qué valores se distribuyen los datos, se necesita complementar con otras mediciones que ayuden a observar y analizar otras características. Es por ello, que existen mediciones que además detectan la dispersión de los valores de la población o muestra.

La importancia del uso de las medidas de dispersión está en el hecho de que el análisis estadístico no es suficiente con el cálculo e interpretación de las medidas de tendencia central, ya que, por ejemplo, cuando pretendemos representar una descripción global de los datos con los que se está trabajando, no estamos abarcando lo que podría estar pasando con los datos extremos inferiores y superiores a la media aritmética; es decir, si son cercanos o alejados a la media; esto no podría explicarse solamente con el análisis de la tendencia central (Runyon, y Haber, 1992).

De esta manera, las medidas de dispersión permiten la representación de la dispersión de puntajes alrededor de las medidas de la tendencia central. Se han descrito las siguientes medidas de dispersión: desviación media, varianza, desviación y error estándar (Cáceres, 2007; Kirk, 2008).

Desviación media

La desviación media (DM) se define como la sumatoria de los valores absolutos de todas las desviaciones, dividido entre el número de datos. En otras palabras, esta medida de dispersión calcula la distancia absoluta promedio entre cada uno de los datos y su media.

Necesariamente, la media es primordial para calcular la desviación media, creándose así su fórmula para la muestra:

$$DM = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Así como, desviación media para la población:

$$DM = \frac{\sum_{i=1}^N |x_i - \mu|}{N}$$

En donde:

- DM: es la desviación media
- x_i : cada dato u observación
- μ : es la media población
- \bar{x} : es la media muestral
- N: es el número de observaciones o datos de la población
- n: es el número de observaciones o datos de la muestra
- \sum : es el signo de sumatoria
- $| |$: es el signo del valor absoluto; es decir, no se consideran los signos (negativos o positivos), solo el valor.

Ejemplo para el cálculo de la desviación media para muestra:

En un laboratorio de Neurociencias, se obtuvo una serie de datos de pruebas de Cociente intelectual (CI) de 7 niños; los puntajes fueron de:

103, 99, 91, 98, 100, 105, 88

Tomando en cuenta que la media de este conjunto de datos es de 97.7, se realiza el cálculo de la DM.

$$DM = \frac{|(103 - 97.7) + (99 - 97.7) + (91 - 97.7) + (98 - 97.7) + (100 - 97.7) + (105 - 97.7) + (88 - 97.7)|}{7}$$

$$DM = \frac{|(5.3) + (1.3) + (6.7) + (0.3) + (2.3) + (7.3) + (9.7)|}{7}$$

$$DM = \frac{35.2}{7} = 5.02$$

La desviación media de este ejemplo fue de 5.2 mayor y menor que la media, es decir, el coeficiente intelectual de los niños se expresa respecto a la media, de la siguiente manera: $\bar{x} \pm DM$. De acuerdo con los datos del ejemplo, esto se representa así:

97.7 ± 5.2 puntaje de CI de los niños, en otras palabras, el puntaje del CI de los niños va desde 92.5 a 102.9, respecto a su media.

Varianza

Como se mostró previamente, en la DM se toman en cuenta los valores absolutos de las diferencias de los datos con respecto a la media, hecho que significa que se pierden los signos negativos. Hecho que hace que muchas

personas prefieran usar la varianza, además, de que es útil en el ajuste de modelos estadísticos que usan formas cuadráticas.

Asimismo, la varianza, es definida como una medida estadística que representa la dispersión de los miembros de un grupo y nos muestra la variabilidad de los datos. Es decir, determina el grado en que cada dato varía de la media. Se he descrito que cuando la varianza de una serie de datos es pequeña, muestra la proximidad de los puntos de datos a la media, mientras que cuando existe un valor grande de varianza, esta se interpreta como la lejanía que presentan los datos alrededor de la media (Bologna, 2013).

Dentro del cálculo de la varianza se ha mencionado que es la medida correspondiente a los datos, pero elevada al cuadrado. Es importante mencionar que en la varianza al elevarse los residuos al cuadrado no es posible que salga negativa. Por lo tanto, esta no puede ser menor que cero (Runyon y Haber, 1992).

La fórmula para calcular la varianza es:

$$\sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2}{n}$$

En donde:

- σ^2 : varianza
- x_i : cada observación o dato, en donde la i toma valores entre 1 y n .
- n : número total de observaciones de la variable de interés.
- \bar{x} : Es la media de la variable de interés.

O bien, la fórmula también puede observarse de la siguiente manera:

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Ejemplo para calcular la varianza:

Vamos a tomar de ejemplo una investigación sobre Toma de decisiones, tomando en cuenta el tiempo de reacción (TR) ante la elección de un estímulo. Tenemos cinco participantes, cada uno con un tiempo de reacción diferente: 7, 6.8, 8, 8, 5

En este ejemplo la variable de interés es el TR. Pero, antes de iniciar con el cálculo de la varianza, primero se tiene que calcular la media del TR, que fue de 6.8, obtenida por:

$$7+6+8+8+5=34/5=6.8$$

Una vez obtenida la media, y de acuerdo con la fórmula que se presentó anteriormente, se calculará la varianza siguiendo los siguientes pasos:

1. A cada valor de la variable (TR) se le restara la media, como se muestra enseguida.

$$\sigma^2 = \frac{(7 - 6.8)^2 + (6 - 6.8)^2 + (8 - 6.8)^2 + (8 - 6.8)^2 + (5 - 6.8)^2}{5}$$

2. Una vez obtenido el valor de la resta, cada valor se elevará al cuadrado.

$$\sigma^2 = \frac{(0.2)^2 + (-0.8)^2 + (1.2)^2 + (1.2)^2 + (-1.8)^2}{5}$$

3. Finalmente, se suman todos los valores obtenidos.

$$\sigma^2 = \frac{0.04 + 0.64 + 1.44 + 1.44 + 3.24}{5}$$

4. Y finalmente, la sumatoria de todos los valores de TR se divide entre n; es decir, 5.

$$\sigma^2 = 1.36$$

El resultado es de 1.36 segundos cuadrados. Sin embargo, siempre que se calcula la varianza el resultado se obtiene con unidades al cuadrado, por lo que, para pasarlo a la medida original de segundos (sin cuadrados), se debe de calcular la desviación estándar, de la cual, se hablará en el siguiente subtema del presente capítulo.

Desviación estándar

Mientras que la varianza muestra cuán lejos están los individuos en un grupo, la desviación estándar pretende mostrar cuántas observaciones de un conjunto de datos difieren de su media.

Por consiguiente, la desviación estándar muestra el grado de dispersión de los valores respecto a la media, lo que permite comparar la variabilidad de distintas distribuciones, interpretar los valores dentro una distribución y, ser usada en otros métodos estadísticos. Es una de las medidas de dispersión más utilizada. Se expresa con S para una muestra y con σ para una población (Kirk, 2008).

Las fórmulas para S y σ son;

$$S = \frac{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}{n} \qquad \sigma = \frac{\sqrt{\sum_{i=1}^n (x_i - \mu)^2}}{n}$$

En donde:

- \bar{x} : es la media de la muestra
- μ : es la media de la población
- x_i : cada observación o dato, en donde la i toma valores entre 1 y n .

- n : número total de observaciones de la variable de interés.
- \bar{x} : Es la media de la variable de interés.

Las fórmulas de la desviación estándar se pueden comprender intuitivamente, sin embargo, hay que tener en cuenta que:

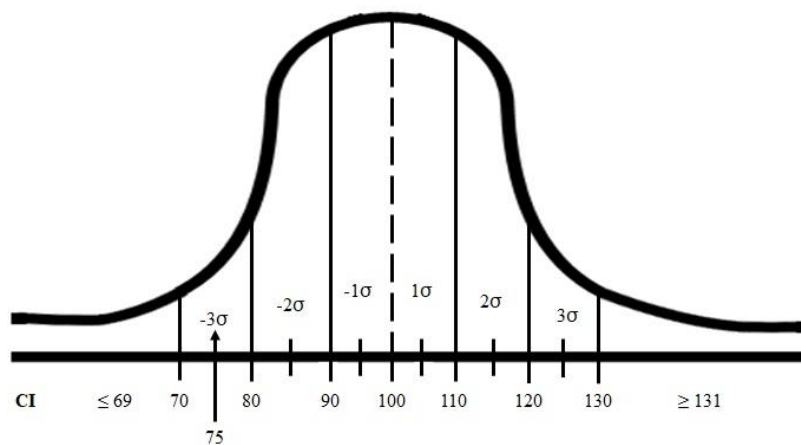
1. Se calcula a partir de cada valor en una distribución.
2. Cada valor es una desviación de la media.
3. Cada desviación se eleva al cuadrado.
4. Hay que sumar las desviaciones al cuadrado (para obtener un valor distinto a 0).
5. La suma de las desviaciones al cuadrado se divide en n , para obtener la distancia media al cuadrado por la cual los valores se desvían de la media.
6. Finalmente, se saca su raíz cuadrada para convertir el valor obtenido de las desviaciones a la unidad de medida del valor inicial.

Ejemplo para el uso de la desviación estándar:

Como se mencionó anteriormente, la desviación estándar sirve para comparar la dispersión de un valor respecto a la media. Un ejemplo del uso de la desviación estándar es la evaluación del cociente intelectual (CI), ya que las puntuaciones obtenidas se comparan con el rendimiento típico, el cual se ha establecido que la media es de 100, cuando tenemos una distribución normal. Al aplicar una escala (WISC o WAIS) para evaluar la inteligencia y aptitudes de una persona, se busca comparar su rendimiento con el rendimiento típico, buscando qué tanto se aleja la puntuación obtenida de la media y permitiendo una interpretación de ésta.

En la figura 2.3.1 se muestra la ubicación de una puntuación de CI obtenido respecto a la media, con esta figura se puede deducir que la persona que obtuvo dicha puntuación se encuentra en un rango inferior a 2 desviaciones estándar respecto a la media, e interpretar que hay una probable discapacidad, dificultad del aprendizaje o disfunción neuropsicológica.

Figura 2.3.1. Ubicación de una puntuación de CI respecto a la media y desviación típica



Nota: la figura representa la ubicación de una puntuación de CI obtenido respecto a la media y las desviaciones estándar que se aleja dicha puntuación de la media.

Error estándar

Al realizar una investigación casi nunca se cuenta con los datos de la población, por lo que se trabaja con los datos de una muestra “representativa”, que teóricamente se obtiene tomando los componentes de la muestra de forma aleatoria. De esta manera su desviación, media y distribución serán similares a las de la población y, mientras mayor sea la muestra más semejante será (Runyon y Haber, 1992).

El error estándar es la búsqueda de la discrepancia entre el resultado de una población y el de la muestra, haciendo una estimación de la desviación estándar a partir de una muestra. Es la medida del error que se pudo cometer al tomar la media calculada de una muestra como estimación de la media poblacional (Runyon y Haber, 1992).

Mientras que la variabilidad de una población se mide en desviaciones estándar, la variabilidad de una muestra se mide en errores estándar. Una muestra abarca sólo una parte de la población, si se toman varias muestras de una población, al cambiar los valores de dichas muestras cambia la media.

Si se tiene la desviación estándar de la población, la fórmula para obtener el error estándar es:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

En donde:

- σ = es la desviación estándar de la población
- n = es el tamaño de la muestra

Si se desconoce la desviación estándar de la población, la estimación del error se realiza con la fórmula:

$$\sigma_{\bar{x}} \approx \frac{s}{\sqrt{n}}$$

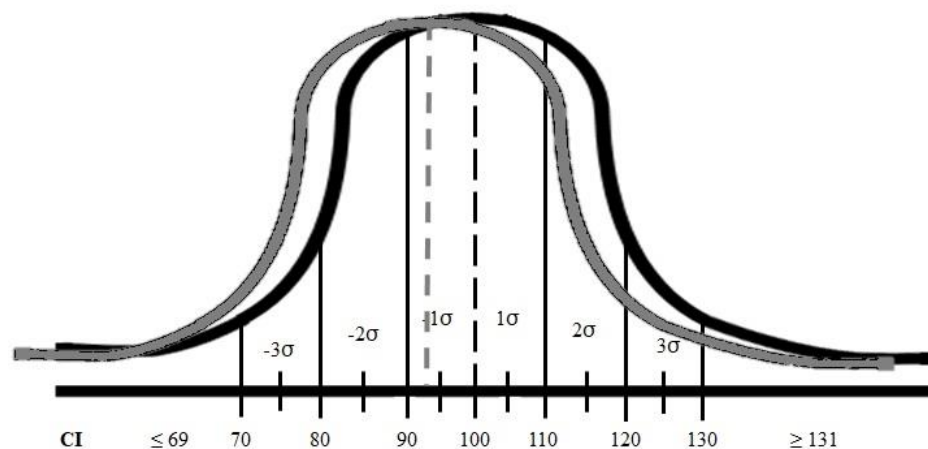
En donde:

- S = es la desviación estándar de la muestra
- n = es el tamaño de la muestra

Ejemplo para el uso del error estándar:

En la figura 2.4.1 se muestra la distribución de CI de una muestra aleatoria de 100 niños, donde la media es 94 y la desviación estándar es 15, respecto a la distribución de CI de la población, donde la media es 100 y la desviación estándar es 10.

Figura 2.4.1. Distribución de CI de una muestra respecto a la distribución de una población



Nota: la figura representa en color gris la distribución de CI de una muestra ($\bar{x}=94$, $S=15$) y en color negro la distribución de CI de una población ($\mu=100$, $\sigma=10$). Donde el error estándar de la media estimada es 1.

Conclusión

Dentro del desarrollo de proyectos de investigación científica, es muy importante el uso de métodos estadísticos para analizar resultados. Ahora bien, en el área de ciencias de la salud, al momento de investigar sobre un fenómeno, no solo basta el análisis cualitativo, el cual, se interpreta de forma subjetiva, esto se refiere a que el resultado queda a interpretación de la persona que lo observa; sino que es necesario del apoyo de la estadística para realizar de una manera más clara y objetiva un resumen de los resultados obtenidos, y así,

complementar el análisis del objeto de estudio. Es por ello que, es muy necesario el conocimiento de estas medidas estadísticas y otras que se verán en capítulos posteriores, para tomarlas como base en el análisis de datos en un reporte de investigación.

Entonces, de acuerdo con este capítulo ya sabemos un poco de la estadística descriptiva, desde la descripción de un conjunto de datos que se centran en la distribución muestral hasta las medidas de dispersión o variabilidad que se distribuyen alrededor de la media. Una vez sabiendo esto, en una investigación se pueden observar los alcances y limitaciones que tiene la información recolectada y de esta manera, realizar más análisis estadísticos o futuras intervenciones que enriquecerían los proyectos de investigación.

Referencias bibliográficas

- Ávila, C. y Polaino-Lorente, A. (2002). *Niños hiperactivos. Comportamiento, diagnóstico, tratamiento, ayuda familiar y escolar*. Alfaomega.
- Bologna, E. (2013). *Estadística para psicología y educación*. Editorial Brujas.
- Candia, J. A. C., Sepúlveda, M. L. A., y Orozco, R. A. R. (2015). *Neurociencia y educación especial. Conceptos, procesos y principios básicos*. Universidad Autónoma de Tlaxcala-Universidad de Guadalajara.
- Cáceres, R.Á. (2007). *Estadística aplicada a las ciencias de la salud*. Ediciones Díaz de Santos.
- Cebrián, L.F. (2009). *Análisis estadístico descriptivo*. Universidad Inca Garcilaso de la Vega Facultad de Ciencias de la Comunicación Turismo y Hotelería.
- Cortes, P.M., García-Hernández, J.P., Iribe-Burgos, F.A., Hernández-González, M., Sotelo-Tapia, C., y Guevara, M.A. (2021). Temporal division of the decision-making process: An EEG study. *Brain Research*, 1769, 147592.

- de la Torre Benítez, G.G., y de la Torre, A.G. (2003). Evaluación de la atención mediante el Test de Cancelación Simple y Cancelación Condicionada [Batería Neuropsicológica Sevilla (BNS)] en niños con Trastorno por Déficit de Atención e Hiperactividad (TDAH). *Revista Española de Neuropsicología*, 5(2), 177-193.
- Devore, J. (2019). *Fundamentos de probabilidad y estadística (1ª ed.)*. Cengage learning editores.
- Fernández, S.F., Sánchez, J.M. C., Córdoba, A., y Largo, A.C. (2002). *Estadística descriptiva*. Esic Editorial.
- García-Cedillo, I., y Romero-Contreras, S. (2019). Influencia de la Declaración de Salamanca sobre la Atención a la Diversidad en México y Situación Actual. *Revista Latinoamericana de Educación Inclusiva*, 13(2), 123-138.
- Kirk, R. E. (2008). *Statistics. An Introduction. Statistics. An Introduction (Fifth edit)*. Thomson Wadsworth.
- Montero-Lorenzo, J.M. (2007). *Estadística descriptiva*. Editorial Paraninfo.
- Parra, J.M. (1995). *Estadística descriptiva e inferencial I*. Colegio de Bachilleres.
- Ricardi, F.Q. (2011). Medidas de tendencia central y dispersión. *Medwave*, 1-6.
- Runyon, R., y Haber, A. (1992). *Estadística para las ciencias sociales*. Addison Wesley Iberoamericana.
- Triola, M. (2009). *Estadística (10ª ed)*. Pearson educación.

PROCEDIMIENTOS PARA EL CÁLCULO DE LAS MEDIDAS DE POSICIÓN

Marai Pérez Hernández [‡]

Edgar Oswaldo Zamora González

Luz Berenice López Hernández

Resumen

Las medidas estadísticas permiten conocer las características de una variable, contribuyendo a un primer análisis descriptivo, que da luz a información que no se puede ver a simple vista, y que posteriormente podría ser utilizado para análisis más profundos.

En este capítulo se abordará particularmente, aquellas medidas estadísticas que dan sentido a una variable en la medida que su cálculo permite referenciar a una escala, las llamadas medidas de posición, dentro de las cuales podemos encontrar los cuartiles, deciles, percentiles y el puntaje z.

Las medidas de posición son herramientas de la estadística descriptiva que nos permiten dividir la distribución en diferentes segmentos, llamados cuantiles, para conocer el lugar exacto de un elemento dentro de un grupo de datos, en este caso se hace referencia específicamente de los cuartiles, deciles, percentiles. Por su parte, el puntaje Z también indica la posición de un dato, pero en esta medida de posición, lo indicará con relación a la media, y sus valores son expresados en unidades de

[‡] Laboratorio de Neurociencias Cognitivas, Departamento de Bienestar y Desarrollo Sustentable, Centro Universitario del Norte de la Universidad de Guadalajara. Carretera Colotlán-Huejúcar, Km 191, Colotlán, Jalisco. marai.perez@academicos.udg.mx

desviación estándar, por lo que describe cuánto se aleja un valor de la media de la distribución.

A lo largo del manuscrito se presentan los procedimientos para obtener las medidas de posición y en seguida se hace uso de ellas mediante ejemplos, pretendiendo acompañar su cálculo de una manera clara y concisa.

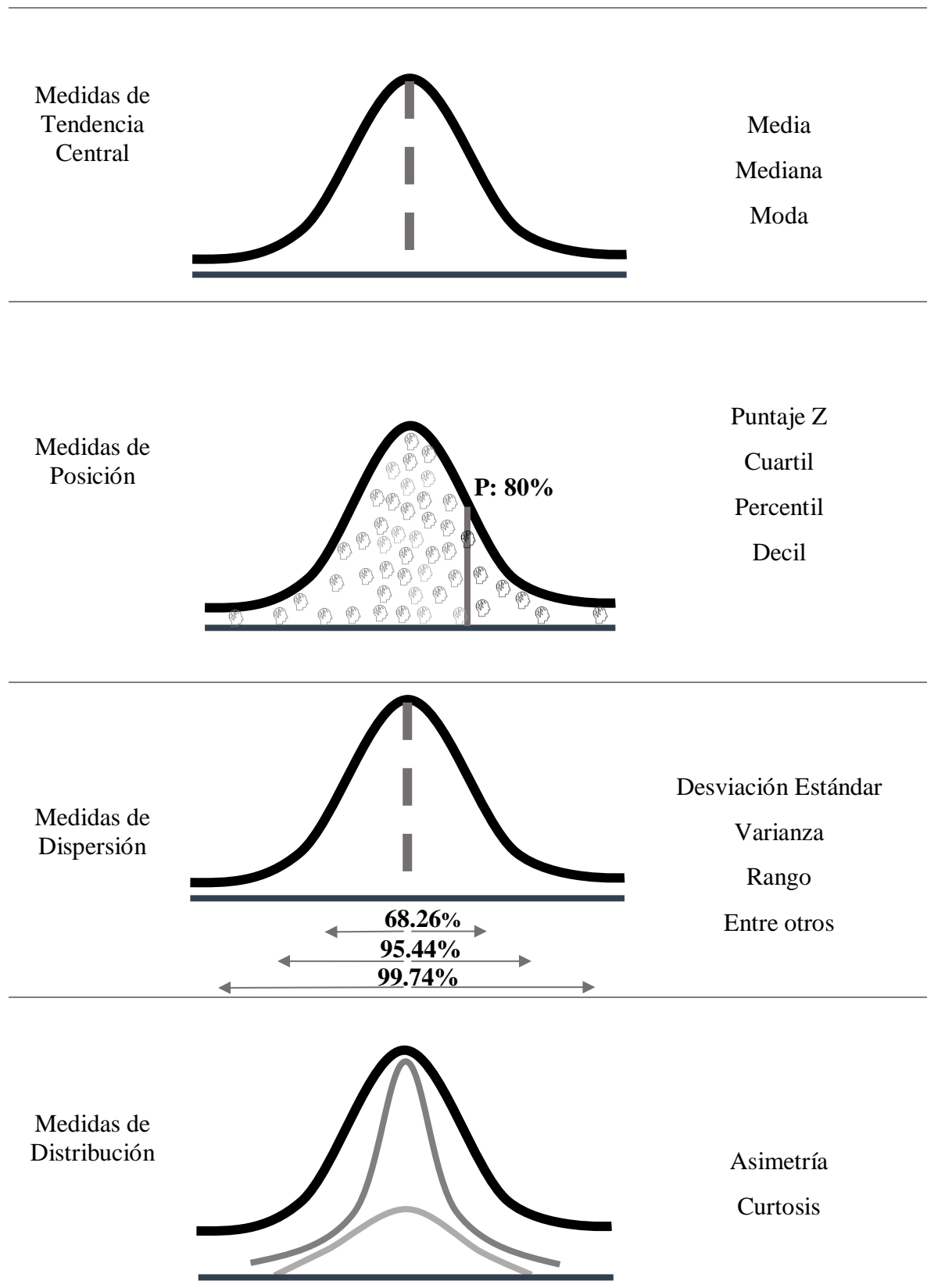
Palabras clave: Medidas estadísticas; medidas de posición, puntajes z.

Introducción

Las medidas numéricas, son usadas para resumir información de un conjunto de datos, que se obtienen a partir de observar y medir una o diversas variables de un fenómeno determinado. Pueden ser adquiridas de una población, y son denominadas parámetros; o pueden ser adquiridas a partir de una muestra (subconjunto de datos de una población) a los cuales se les denominan estadísticos o estadígrafos (William, 2014; Levin, 2010). La obtención o cálculo de los parámetros o estadísticos da un panorama bastante aproximado de la distribución y características que presenta un conjunto de datos (Torres, 2018).

Los parámetros o estadísticos son parte de la estadística descriptiva numérica y se pueden dividir de acuerdo con sus características en: Medidas de tendencia central; Medidas de dispersión; Medidas de posición; Medidas de distribución (Figura 1) (William, M., 2014; Juarez Garcia, 2014).

Figura 1. Principales medidas numéricas de la estadística descriptiva analítica.



Todos estos métodos numéricos, permiten dar un orden a un grupo de datos desorganizados, pero además ayudan a resumir información básica que a simple vista no es posible observar y que posteriormente se puede utilizar para análisis más profundos (Juares Garcia, 2014).

Las medidas de tendencia central y las medidas de posición permiten obtener la posición relativa de los datos (Cobo y cols., 2007). Las medidas de tendencia central son la media, moda y mediana y son descritas en el capítulo previo, por lo que en este apartado nos enfocaremos en las medidas de posición. Y solo se hará uso o mención de las medidas de tendencia central cuando sea necesario.

Cualquier medida de una variable carece de sentido si no está referenciada a una escala, la información de si esa medida es alta o es baja, se obtendrá a partir de ubicarla en el marco de su población de referencia (Ballesteros, 2012). Por ejemplo, en el ámbito de la educación especial, hablar del cociente intelectual, del desempeño escolar o de cualquier otro resultado de un infante no nos diría nada si no existe un marco de referencia claramente definido y uniforme (Anastasi y Urbina, 1988).

Para darle significado a una variable es necesario hacer divisiones, en sentido ascendente o descendente, que permitan caracterizar el valor de una observación facilitando la ubicación de orden dentro de un conjunto de datos (Ballesteros, 2012). Precisamente las medidas de posición son herramientas de la estadística descriptiva que nos permiten dividir la distribución en diferentes segmentos, llamados cuantiles, para conocer el lugar exacto de un elemento (medida) dentro de un grupo de datos. En otras palabras, en una distribución dada de un conjunto de datos, se puede identificar la ubicación exacta de un dato en relación con los datos que conforman el grupo (Juares Garcia, 2014;

Triola, 2018), los más comunes son los cuartiles, deciles y percentiles (centiles) (Ballesteros, 2012; Cobo y cols., 2007). Sin embargo, utilizando el mismo procedimiento podrían calcularse otros cuantiles (unidades divisionales) como quintiles o sextiles. Un importante paso en el proceso de análisis, que comparten estas herramientas es la disposición de los datos que debe ser en orden de magnitud e identificar la cantidad de mediciones obtenidas y número de datos (n) (Cobo, 2007).

La elección de entre las principales medidas de posición dependerá de la finalidad del análisis exploratorio y de las características de la variable que se está analizando, además de otros dos aspectos: el nivel de detalle con el que se quiera contextualizar la posición de una observación (variable) en el conjunto de datos referenciados y la heterogeneidad de las distribuciones, es decir, entre mayor variabilidad haya en los datos de una observación, será más útil y descriptivo un mayor número unidades divisionales (Ballesteros, 2012). Así podemos utilizar los cuartiles que dividen al conjunto de valores en 4 partes iguales, los deciles en 10 o los percentiles en 100, por lo que en este sentido los percentiles no serían una herramienta adecuada para dividir conjuntos de mediciones reducidas en comparación a los cuartiles que sí podrían ser utilizados para grupos de datos con menor número de observaciones (William, 2014; Jay, 2015; Triola, 2018).

Otra medida de posición es el puntaje Z, el cual permite determinar la posición de un dato de acuerdo con la media, es decir es utilizado para analizar la distancia de un valor obtenido respecto a la media, expresada en unidades de desviación estándar (Hernández Sampieri, Fernández Collado y Baptista Lucio, 2008). El puntaje Z, los cuartiles, percentiles y deciles son herramientas de la estadística descriptiva que nos permite realizar un análisis exploratorio de los

datos. Permitiendo obtener información relevante sin necesidad de hacer uso de la estadística inferencial (Jay, 2015).

Cálculo de los cuantiles

Existen varios métodos para calcular los cuantiles (cuartiles, deciles y percentiles) que dependen de la fuente consultada y de las características de la distribución, es decir, si los datos se encuentran desagrupados o agrupados en intervalos. En el primer caso, la determinación de la posición será por localización; en el caso de datos agrupados en intervalos se llevará a cabo un tratamiento aritmético más complejo (Tabla 1). Es importante precisar que independientemente del método que se elija, este será el mismo para todos los cuantiles únicamente cambiando el número de partes en que se quiera dividir el conjunto de datos, ya sea 4 (cuartil), 10 (decil) o 100 (percentil) partes (Ballesteros, 2012).

Tabla 1. Fórmulas para el cálculo de los cuantiles de diferentes fuentes (Johnson y Kuby, 2012; Ballesteros, 2012; Johnson, 2012; Triola, 2018) y dependiendo la distribución de los datos. Las fórmulas que se utilizan en el presente documento son las sombreadas. El significado y la descripción de las fórmulas se retomará en el apartado correspondiente.

CUANTILES	DATOS DESAGRUPADOS	DATOS AGRUPADOS EN INTERVALOS
Cuartiles	$Q_n = \frac{(Qn)n}{4}$	$Q_n = \frac{Qn}{4}N$
		$Q_m = Li + \frac{\frac{mN}{4} - Nai - 1}{nm}(a)$
Deciles	$D_n = \frac{(Dn)n}{10}$	$D_n = \frac{Dn}{10}N$
		$D_m = Li + \frac{\frac{mN}{10} - Nai - 1}{nm}(a)$
Percentiles	$P_n = \frac{(Pn)n}{100}$	$i = \left(\frac{p}{100}\right)n$
		$Q_m = Li + \frac{\frac{mN}{100} - Nai - 1}{nm}(a)$

En los apartados siguientes se describen y presentan ejemplos de cada uno de los cuantiles utilizando los diferentes métodos para calcularlos. En el caso de los cuartiles y percentiles se trabaja con datos desagrupados, mientras que en los deciles se presenta un ejemplo con datos agrupados en intervalos.

Cuartiles

Los cuartiles, permiten dividir a un conjunto de datos organizados de menor a mayor en 4 partes iguales, el resultado obtenido nos dará la posición en que se encuentra el valor del cuartil. Por lo que en los cuartiles el valor intermedio, cuartil 2 (Q2) corresponde a la mediana de los datos, el cuartil 1 (Q1) divide la mitad inferior de los datos y el cuartil 3 (Q3) divide la mitad superior de los datos, con lo que se logra dividir los datos en cuatro subgrupos de datos con el mismo número de mediciones. En otras palabras, para dividir los datos en cuartiles solo es necesario calcular los cuartiles Q1, Q2 y Q3 (William, 2014; Cobo, 2007; Johnson y Kuby, 2012).

Como se ha mencionado anteriormente, no hay un consenso en la fórmula y el procedimiento que se debe utilizar al momento de calcular los cuartiles, por lo que, dependiendo de la fuente consultada, se podrán observar pequeñas diferencias. Sin embargo, al momento de realizar el cálculo es necesario seguir una serie de pasos (Johnson y Kuby, 2012). El procedimiento consta de los siguientes pasos:

1. Ordenar los datos, de menor a mayor
2. Calcular los cuartiles Q1, Q2 y Q3 con sus fórmulas correspondientes II.2.2; II.2.2.1; II.2.2.2 (Figura 1).

3. Determinar si el resultado es un número entero (caso A) o es un número con fracción (caso B).

Caso A: Cuando el Cuartil calculado ($Q_{\text{calculado}}$) es un número entero, el valor del cuartil es el promedio de los valores de la posición del $Q_{\text{calculado}}$ y el $Q_{\text{calculado}}$ más una posición.

Caso B: Cuando el Cuartil calculado es un número con una fracción, debe ser redondeado al próximo superior. Por lo que el primer entero mayor que Q indica la posición del $Q_{\text{calculado}}$.

4. Localizar el valor que le corresponde a la posición calculada.

$$\text{Cuartil 1 (Q1): } Q1 = \frac{n}{4} \quad (\text{II.2.2})$$

$$\text{Cuartil 2 (Q2): } Q2 = \frac{2n}{4} \quad (\text{II.2.2.1})$$

$$\text{Cuartil 3 (Q3): } Q3 = \frac{3n}{4} \quad (\text{II.2.2.2})$$

Para clarificar el uso de los cuartiles se expone la siguiente investigación. Un investigador utiliza la escala para valorar la cultura, gestión y las prácticas inclusivas de 200 profesores de educación básica (García, 2015) del distrito escolar de su localidad. De los cuales 45 son profesores de educación preescolar, 75 de educación primaria y 80 de educación secundaria (Tabla 2). Dicha escala valora a los profesores con puntuaciones que van de 0 a 10.

Tabla 2. Calificación obtenida por cada profesor agrupado de acuerdo con el nivel educativo en el que se desempeñan.

Profesores Educación Preescolar					Profesores Educación Primaria					Profesores Educación secundaria				
Calificaciones					Calificaciones					Calificaciones				
7	10	8	8	8	5	4	6	4	6	4	3	6	2	5
10	8	10	10	9	6	8	4	6	7	7	6	5	5	5
8	5	10	6	7	7	4	4	8	6	1	1	1	7	3
10	6	8	9	10	4	5	5	4	4	7	1	5	3	3
8	9	7	10	5	6	4	7	6	6	7	2	7	6	4
10	10	8	8		5	4	4	4	6	5	1	3	6	2
9	6	8	10		8	1	4	4	5	2	7	7	2	6
10	8	5	7		7	5	5	4	5	7	1	7	6	5
10	9	8	5		7	6	6	8	1	4	7	4	1	4
10	9	8	6		6	6	4	8	6	5	4	2	4	3
					8	7	5	4	5	6	4	7	1	2
					8	6	4	7	8	3	4	1	2	5
					8	8	5	8	6	5	1	6	2	3
					7	8	4	5	8	1	3	3	1	4
					8	6	6	8	4	1	6	1	6	1
										5	3	4	5	1

Se desea analizar la distribución de los resultados obtenidos, de forma independiente para cada grupo de profesores de acuerdo con el nivel de educación en el que se desempeñan. Se utilizan los cuartiles con la finalidad de realizar la evaluación de cada grupo.

Para analizar los datos de forma independiente se debe seguir el procedimiento previamente descrito para la obtención de los cuartiles Q1, Q2 y Q3 de cada grupo.

Paso 1. Ordenar los datos de menor a mayor en cada grupo de profesores, como se muestra en la Tabla 3.

Tabla 3. Grupos de datos ordenados de menor a mayor de las Calificación obtenida por cada profesor agrupado de acuerdo con el nivel educativo en el que se desempeñan en

Profesores educación Preescolar					Profesores Educación Primaria					Profesores Educación Secundaria				
Calificaciones					Calificaciones					Calificaciones				
5	7	8	9	10	1	4	5	6	8	1	2	3	5	6
5	7	8	10	10	1	4	5	6	8	1	2	3	5	6
5	8	8	10	10	4	4	5	6	8	1	2	3	5	6
5	8	8	10	10	4	4	5	6	8	1	2	3	5	6
6	8	8	10	10	4	4	6	6	8	1	2	4	5	6
6	8	9	10		4	4	6	6	8	1	2	4	5	7
6	8	9	10		4	4	6	6	8	1	2	4	5	7
6	8	9	10		4	5	6	7	8	1	2	4	5	7
7	8	9	10		4	5	6	7	8	1	2	4	5	7
7	8	9	10		4	5	6	7	8	1	3	4	5	7
					4	5	6	7	8	1	3	4	5	7
					4	5	6	7	8	1	3	4	6	7
					4	5	6	7	8	1	3	4	6	7
					4	5	6	7	8	1	3	4	6	7
										1	3	5	6	7

Paso 2. Cálculo de Cuartiles para cada grupo de profesores de educación básica.

Calcular los cuartiles para el grupo de profesores de educación preescolar:

$$Q1: Q1 = \frac{n}{4} = \frac{45}{4} = 11.25 = 12$$

$$Q2: Q2 = \frac{2n}{4} = \frac{2(45)}{4} = 22.5 = 23$$

$$Q3: Q3 = \frac{3n}{4} = \frac{3(45)}{4} = 33.75 = 34$$

Paso 3. Debido a que los cuartiles calculados son números con una fracción, debe ser redondeado al próximo superior, como lo sugiere el paso 3, caso B. Por lo que el primer entero mayor que Q indica la posición del $Q_{\text{calculado}}$.

Paso 4. Localizar el valor que le corresponde a la posición calculada. Entonces:

Q1: 11.25 = 12 posición, que corresponde a la calificación de 7

Q2: 22.50 = 23 posición, que corresponde a la calificación de 8

Q3: 33.75 = 34 posición, que corresponde a la calificación de 10

Se determinan los cuartiles para el grupo de profesores de primaria:

$$Q1: Q1 = \frac{n}{4} = \frac{75}{4} = 18.75 = 19$$

$$Q2: Q2 = \frac{2n}{4} = \frac{2(75)}{4} = 37.50 = 38$$

$$Q3: Q3 = \frac{3n}{4} = \frac{3(75)}{4} = 56.25 = 57$$

Paso 3: Debido a que los cuartiles calculados son números con una fracción, debe ser redondeado al próximo superior, como lo sugiere el caso B

del paso 3. Por lo que el primer entero mayor que Q indica la posición del $Q_{\text{calculado}}$.

Paso 4. Localizar el valor que le corresponde a la posición calculada.

Entonces:

Q1: 18.75 = 19 posición, que corresponde a la calificación de 4

Q2: 37.50 = 38 posición, que corresponde a la calificación de 6

Q3: 56.25 = 57 posición, que corresponde a la calificación de 7

Se determinan los cuartiles para el grupo de profesores de secundaria:

$$Q1 = \frac{n}{4} = \frac{80}{4} = 20$$

$$Q2 = \frac{2n}{4} = \frac{2(80)}{4} = 40$$

$$Q3 = \frac{3n}{4} = \frac{3(75)}{4} = 60$$

Paso 3: Debido a que los cuartiles calculados son números enteros, el valor del cuartil es el promedio de los valores de las posiciones del $Q_{\text{calculado}}$ y el $Q_{\text{calculado}}$ más una posición. Por lo que se identifican los $Q_{\text{calculado}}$ y el $Q_{\text{calculado}}$ más una posición de cada cuartil.

Paso 4. Localizar el valor que le corresponde a la posición calculada.

Entonces:

Q1: 20 posición, corresponde a la calificación de 2 y $Q1 + 1$ posición corresponde a 2

Q2: 40 posición, corresponde a la calificación de 4 y $Q2 + 1$ posición corresponde a 4

Q3: 60 posición, corresponde a la calificación de 6 y Q3 + 1 posición corresponde a 6

Se determina los cuartiles Q1, Q2 y Q3:

$$Q2 = \frac{2 + 2}{2} = 2$$

$$Q2 = \frac{4 + 4}{2} = 4$$

$$Q3 = \frac{6 + 6}{2} = 6$$

Después de calcular los cuartiles para cada grupo de profesores podemos obtener una serie de observaciones:

Observación 1: En el grupo de educación preescolar, el 75% de los profesores obtuvieron una calificación igual a mayor 7 y solo el 25% obtuvieron una calificación menor o igual a 7. Por lo que la mayoría de los profesores obtuvieron una calificación aprobatoria.

Observación 2: En el grupo de educación primaria, el 50% de los profesores obtuvieron una calificación igual o menor a 6 y el otro 50% obtuvo una calificación igual o mayor a 6. Por lo que la cantidad de profesores que obtuvieron una calificación aprobatoria al parecer es igual a la cantidad de profesores que obtuvieron una calificación no aprobatoria.

Observación 3: En el grupo de educación secundaria, el 75% de los profesores obtuvieron una calificación menor o igual a 6 y solo el 25% obtuvieron calificaciones iguales o superiores a 6. Por lo que el 75% de los profesores obtuvieron una calificación no aprobatoria.

Deciles

Los deciles (D) son los nueve valores que fraccionan la distribución en diez partes iguales, conteniendo cada una de ellas la décima parte de las observaciones y se nombran como D1, D2, D3, D4, D5, D8 y D9 (Ballesteros, 2012). El D1 corresponde al décimo percentil, dejando por debajo el 10% de los datos, el D2 es igual al veinteavo percentil, dejando por debajo el 20% de los datos y así sucesivamente con los nueve deciles (Suárez y Tapia, 2012; Mesa Guerrero y Caicedo Zambrano, 2020).

Para el obtener la posición de los deciles en una distribución de datos desagrupados se seguirá los mismos pasos que en los cuartiles y percentiles, ordenar los datos en orden ascendente, en seguida hacer el cálculo usando la siguiente fórmula (Ballesteros, 2012):

$$Dn = \frac{Dn}{10} N$$

Donde:

- Dn: Es el índice Dn calculado (elemento en el puesto Dn)
- n: Decil deseado
- N: Número de datos

Para calcular los deciles cuando se tiene datos agrupados en intervalos se deberá seguir el siguiente procedimiento (Ballesteros, 2012):

1. A partir de los datos agrupados en intervalos y su frecuencia absoluta, completar la tabla de distribución de frecuencias.
2. Localizar los intervalos donde se encuentran los valores deciles estableciendo los nueve puntos que dividen la distribución en diez

segmentos iguales. Para obtenerlo, se divide el total de los datos entre diez (número de segmentos iguales) y se multiplica el resultado por el decil que se desea.

3. Calcular el valor exacto aplicando la correspondiente fórmula:

$$D_m = L_i + \frac{\frac{mN}{10} - N_{ai-1}}{nm} (a)$$

Donde:

- D_m : Es el índice D_m calculado (elemento en el puesto D_m)
- m : Decil deseado
- L_i : Límite inferior del intervalo donde se encuentra el decil.
- N : Número total de observaciones (frecuencia absoluta total)
- N_{ai-1} : Frecuencia absoluta acumulada del intervalo anterior a aquel que se encuentra el decil.
- nm : Frecuencia absoluta total del intervalo decil.
- a : amplitud del intervalo

Para ejemplificar el procedimiento de los deciles cuando se tienen datos agrupados en intervalos se retomará los datos descritos en Ballesteros (2012). En la Tabla 4 se presenta de manera sintetizada el valor de la tasa de fecundidad temprana en 112 países de los cinco continentes, para el año 2006. A partir de estos datos se determinan los deciles 1, 5 y 9.

Tabla 4. Valor de la tasa de fecundidad temprana en 112 países de los cinco continentes, para el año 2006. Tomado de Ballesteros (2012).

Valor de la fecundidad temprana %	Frecuencia absoluta
3-13	28
14-24	17
25-50	21
51-75	17
76-100	12
101-219	17
	N=112

Paso 1. Completar la tabla de distribución de frecuencias (Tabla 5).

Tabla 5. Tabla de distribución de frecuencias de la tasa de fecundidad temprana. Modificado de Ballesteros (2012).

Valor de la fecundidad temprana %	Frecuencia absoluta	Porcentaje	Frecuencia absoluta acumulada	Porcentaje acumulado	Intervalos deciles
3-13	28	25.00	28	25.00	D1
14-24	17	15.18	45	40.18	
25-50	21	18.75	66	58.93	D2
51-75	17	15.18	83	74.11	
76-100	12	10.71	95	84.82	
101-219	17	15.18	112	100	D9
	N=112				

Paso 2. Localizar los intervalos donde se encuentran los valores deciles. Se busca en qué frecuencia acumulada (observaciones) estaría contenido el resultado (ver intervalos deciles en Tabla 5). En este caso D1=11.2 observaciones en el intervalo 3-13%, D5=56 observaciones correspondiente al

intervalo 25-50% y D9= 100.8 observaciones que corresponde al intervalo 101-219%.

Paso 3. Calcular el valor exacto aplicando la fórmula:

$$Dm = Li + \frac{\frac{mN}{10} - Nai - 1}{nm} (a)$$

$$D1 = 3 + \frac{\frac{(1)(112)}{10} - 0}{28} (10) = 3 + \frac{11.2}{28} (10) = 3 + 4 = 7 \%$$

$$D5 = 25 + \frac{\frac{(5)(112)}{10} - 45}{21} (25) = 25 + \frac{11}{21} (25) = 25 + 13.09 = 38.09 \%$$

$$D9 = 101 + \frac{\frac{(9)(112)}{10} - 95}{17} (118) = 101 + \frac{5.8}{17} (118) = 101 + 40.25 = 141.26 \%$$

A partir de estos resultados se puede observar que los países que se encuentran en el D1 son aquellos que presentan una tasa de fecundidad temprana igual o menor al 7%; los que están dentro del D5 su tasa de fecundidad temprana es igual o menor al 38% y por último los que se localizan en el D9 tendrían una tasa de fecundidad temprana igual o menor al 141%.

Percentiles

Los percentiles son otra medida de posición que permite dividir a un conjunto de datos en 100 partes iguales (subconjuntos) teniendo en cuenta que el grupo de datos no debe tener muchos valores repetidos (Johnson y Kuby, 2012; Triola, 2013; Anderson, y cols., 2015). Es de observarse que se obtienen subgrupos con el 1% de los valores del conjunto de datos (Johnson y Kuby, 2012; Triola, 2013). De esta manera, el percentil p-ésimo (entendiendo por p-ésimo como el elemento en el puesto p, en otras palabras, se refiere al percentil que se desea

calcular) es un valor tal que por lo menos p por ciento de las observaciones son menores o iguales que este valor, y por lo tanto al menos $(100-p)$ por ciento de las observaciones son mayores o igual que este valor. De ahí que para calcular el p -ésimo de un percentil es necesario utilizar la siguiente fórmula:

$$i = \left(\frac{p}{100} \right) n \quad (\text{II.2.3})$$

Donde:

- i : Es el índice i calculado (elemento en el puesto i)
- P : Percentil deseado
- n : Número de datos

El procedimiento para calcular el Percentil es muy similar al del cuartil, la diferencia radica en la fórmula y el percentil que se desea calcular por lo que aplicamos el siguiente procedimiento:

1. Ordenar los datos, de menor a mayor
2. Calcular el percentil p -ésimo deseado utilizando la fórmula.
3. Determinar si el resultado es un número entero (caso A) o es un número con fracción (caso B)

Caso A: Cuando el i calculado es un número entero, el valor del percentil es el promedio de los valores de la posición del índice $i_{\text{calculado}}$ e i más una posición.

Caso B: Cuando el i calculado es un número con una fracción, debe ser redondeado al próximo superior. Por lo que el primer entero mayor que i denota la posición del percentil (Johnson y Kuby, 2012; Anderson, y cols., 2015).

Es importante mencionar, que es posible también calcular el percentil de un valor dado (x) en un grupo de datos y para ello es necesario utilizar la siguiente fórmula (Triola, 2018):

$$\text{Percentil de } x = \frac{\text{número de valores menores que } x}{\text{número total de valores}} * 100 \quad (\text{II.2.3.1})$$

Utilizando la información contenida en la tabla 2, se desea realizar un análisis de toda la información en su conjunto sin importar el grado en el que imparten clases los 200 profesores de educación básica, con el fin de evaluar el porcentaje total de profesores que lograron obtener una valoración aprobatoria igual o mayor a (mínima) 6 en el sector escolar, así como la cantidad de profesores que obtuvieron valoraciones iguales o menores de 3 y valoraciones iguales o mayores a 9 que se consideran bajos y sobresalientes, respectivamente. Así mismo el investigador tiene la intención de observar la distribución de la totalidad de los datos al obtener el percentil 25, 50 y 75.

Encontrando el percentil de un valor x

Con el fin de obtener los percentiles de 6, 3 y 9 se utiliza el procedimiento previamente descrito.

Se ordenan los datos de menor a mayor (Tabla 6).

Tabla 6. Total de datos agrupados de las calificaciones obtenidas por los profesores de educación básica en el distrito, ordenados de menor a mayor.

Profesores de educación Básica							
Puntaje de valoración							
1	2	4	5	6	6	7	8
1	2	4	5	6	6	7	8
1	3	4	5	6	6	8	8
1	3	4	5	6	6	8	8
1	3	4	5	6	7	8	8
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	9
1	3	4	5	6	7	8	10
1	3	4	5	6	7	8	10
1	4	4	5	6	7	8	10
1	4	4	5	6	7	8	10
1	4	4	5	6	7	8	10
1	4	4	5	6	7	8	10
1	4	4	5	6	7	8	10
2	4	4	5	6	7	8	10
2	4	5	5	6	7	8	10
2	4	5	5	6	7	8	10
2	4	5	5	6	7	8	10
2	4	5	6	6	7	8	10
2	4	5	6	6	7	8	10
2	4	5	6	6	7	8	10

En la lista ordenada de menor a mayor, se observa que la primera valoración de 6 se encuentra arriba de la posición 97, por lo que se procede a calcular el percentil utilizando la fórmula II.2.3.1.

$$\text{Percentil de } x = \frac{\text{número de valores menores que } x}{\text{número total de valores}} * 100$$

Despeje:

$$\text{Percentil de } 6 = \frac{97}{200} * 100 = 0.485 * 100 = 48.5 \approx 49$$

Resultado:

$$\text{Percentil de } 6 = 49$$

$$\% \text{ de profesores aprobados} = (100 - P) = 100 - 49 = 51$$

Por lo que podemos concluir que el 51% de los profesores de educación básica del distrito escolar, pasaron la valoración.

Se calculan los percentiles de 3 y 9:

Para encontrar el percentil de 3, se observa que la última valoración de 3 se encuentra en la posición 38, por lo que se procede a calcular el percentil de 3:

Despeje:

$$\text{Percentil de } 3 = \frac{38}{200} * 100 = 0.19 * 100 = 19$$

Resultado:

$$\text{Percentil de } 3 = 19$$

Por lo que 19% de los profesores presentan una baja valoración, de acuerdo con la prueba aplicada.

Para encontrar el percentil de 9, se observa que la primera valoración de 9 se encuentra arriba de la posición 180, por lo que se procede a calcular el percentil de 9:

Despeje:

$$\text{Percentil de 9} = \frac{180}{200} * 100 = 0.9 * 100 = 90$$

Resultado:

$$\text{Percentil de 9} = 90$$

$$\% \text{ de profesores sobresaliente} = (100 - P) = 100 - 90 = 10$$

Por lo que podemos concluir que el 10% de los profesores de educación básica del distrito escolar, obtuvieron una valoración sobresaliente.

Determinando el índice i del percentil deseado

Paso 1. Ordenar los datos de mayor a menor como se presentan en la tabla 4

Paso 2. Se calcula el percentil p -ésimo para cada percentil 25, 50 y 75, de acuerdo con el procedimiento establecido y utilizando la fórmula:

$$i = \left(\frac{p}{100} \right) n$$

Percentil 25:

$$i = \left(\frac{p}{100} \right) n = \left(\frac{25}{100} \right) * 200$$

$$i = 0.25 * 200 = 50$$

Paso 3. Determinar si el resultado es un entero o es un número con fracción, en este caso es un entero y se procede a calcular el índice i de acuerdo con el procedimiento planteado en el paso 4, caso A:

De acuerdo con la Tabla 6, el índice $i_{\text{calculado}}$: 50, corresponde a la calificación de 4 y $i_{\text{calculado}} + 1$ posición corresponde a 4.

Por lo que se procede a calcular el promedio:

$$i = \frac{4 + 4}{2} = 4$$

$$i = 4$$

Por lo que se concluye que el 25% inferior de los datos es igual o menor a una valoración de 4 y son profesores que no tienen los conocimientos elementales que evalúa la prueba.

Percentil 50:

$$i = \left(\frac{p}{100}\right) n = \left(\frac{50}{100}\right) * 200$$

$$i = 0.5 * 200 = 100$$

Paso 3. Determinar si el resultado es un entero o es un número con fracción, en este caso es un entero y se procede a calcular el índice i de acuerdo con el procedimiento planteado en el caso A:

De acuerdo con la Tabla 6, el índice $i_{\text{calculado}}$: 100, corresponde a la calificación de 6 y $i_{\text{calculado}} + 1$ posición corresponde a 6.

Por lo que se procede a calcular el promedio:

$$i = \frac{6 + 6}{2} = 6$$

$$i = 6$$

Por lo que se concluye que el 50% de las valoraciones son igual o mayores a una valoración de 6 y son profesores que sí tienen los conocimientos elementales que evalúa la prueba.

Percentil 75:

$$i = \left(\frac{p}{100}\right)n = \left(\frac{75}{100}\right) * 200$$

$$i = 0.75 * 200 = 150$$

Paso 3. Determinar si el resultado es un entero o es un número con fracción, en este caso es un entero y se procede a calcular el índice i de acuerdo con el procedimiento planteado en el caso A:

De acuerdo con la Tabla 6, el índice $i_{\text{calculado}}$: 150, corresponde a la calificación de 7 y $i_{\text{calculado}} + 1$ posición corresponde a 7.

Por lo que se procede a calcular el promedio:

$$i = \frac{7 + 7}{2} = 7$$

$$i = 7$$

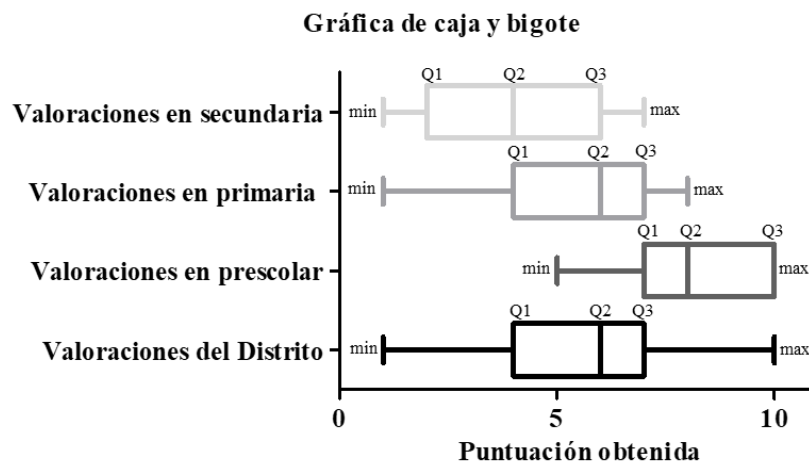
Por lo que se concluye que el 25% de las valoraciones más altas obtenidas son igual o mayores a 7 y son profesores que sí tienen los conocimientos que evalúa la prueba.

Finalmente, si es de interés resumir, describir y explorar un conjunto de datos con mayor claridad y poder realizar comparaciones más exactas de dos grupos de datos o más, es necesario utilizar el resumen de los 5 números y la gráfica de caja y bigotes, todo esto sin hacer uso de la estadística inferencial. De hecho, podríamos darnos una idea de la asimetría que tienen los datos (Johnson, 2012; Triola, 2018).

Consiste en calcular los cuartiles Q1, Q2 y Q3 o lo que es lo mismo, los percentiles 25, 50 y 75, así como de, hacer uso del valor mínimo y máximo (resumen de los 5 números). Con esta información se construye una gráfica que contiene desde el valor mínimo de los datos hasta el valor máximo, los cuales se grafican de forma lineal. Así como, los datos intermedios que corresponden a las posiciones del Q1, Q2 y Q3, con los que se construye la caja, de tal manera que se crea una gráfica mejor conocida como caja y bigotes (Triola, 2013).

Como ejemplo práctico, se utilizan los datos de las Tablas 3 y 6, de los cuales se obtiene el resumen de los 5 números para cada grupo de datos y se se construyen las gráficas de caja y bigotes. Esto permitirá observar el sesgo o dispersión de los datos en su conjunto y su distribución entre los cuartiles (Johnson, 2012).

Figura 3. Gráfica de caja de los diferentes resultados obtenidos en la prueba para valorar la cultura, gestión y las prácticas inclusivas, en el distrito y en los diferentes niveles de educación básica.



Como se observa en la Figura 3, cada nivel de educación presenta una distribución diferente en sus evaluaciones, en la que se puede apreciar que las puntuaciones más altas las encontramos en profesores de nivel preescolar, así mismo, se observa que los profesores de nivel secundaria presentan una distribución menos asimétrica en comparación a los otros grupos.

Puntaje Z

El puntaje Z es una medida de posición relativa a la media y se expresa en unidades de desviación estándar. El resultado obtenido de su cálculo señala la dirección (izquierda o derecha) y el grado en que un valor individual se aleja de la media en una escala de unidades de desviación estándar. De esta manera, se estandariza la escala de una variable, permitiendo comparar mediciones de distribuciones diferentes (Hernández Sampieri, Fernández Collado y Baptista Lucio, 2008).

Para entender de forma práctica el puntaje Z, se presenta el siguiente ejemplo: Un grupo de investigación está desarrollando un estudio piloto con diez padres de familia (PF), con el fin de evaluar la comunicación sobre sexualidad entre padres e hijos con discapacidad, utilizando la herramienta metodológica reportada por Barcena Gaona y cols., (2017). Los resultados obtenidos en el cuestionario se describen en la Tabla 7.

Tabla 7. Datos obtenidos en el estudio piloto. Los datos se obtienen a partir de un cuestionario de 16 preguntas. Los valores se obtienen a partir de la transformación de los datos.

PF	Valor	PF	Valor	PF	Valor	PF	Valor	PF	Valor
I	13	III	3	V	7	VII	4	IX	6
II	2	IV	8	VI	16	VIII	15	X	3

Sin embargo, por falta de experiencia en la aplicación de la metodología tienen dudas de continuar con el protocolo a una escala mayor. Una de las principales cuestiones, es si los datos obtenidos en el estudio piloto son comparables en relación con los datos obtenidos por Barcenas Gaona y cols., (2017) y la única información que se tiene es que, en niños de 12 a 18 años, la media de comunicación sobre sexualidad es de 9.02 y una desviación estándar de 4.87 (Barcenas Gaona y cols., 2017).

Por lo que para evaluar su información es conveniente estandarizar los datos a una función normal estándar típica, como lo es el puntaje Z.

El puntaje Z, permite identificar la posición relativa de los datos con respecto de su media, por lo que es una unidad o fracción de desviaciones estándar de un

valor dado x , con respecto a la media de los datos (William, 2014; Triola, 2018) y está determinada por las siguientes fórmulas:

Muestra:

$$z = \frac{x - \bar{x}}{s} \quad (\text{II.2.1})$$

O

Población:

$$z = \frac{x - \bar{x}}{\sigma} \quad (\text{II.2.1.1})$$

Donde:

- Z: Es la puntuación estandarizada
- X: Es la variable de interés
- \bar{X} : Es la media aritmética de la muestra
- S: Es la desviación estándar muestral
- σ : Es la desviación estándar poblacional

El puntaje Z presenta las siguientes características (William, 2014; Triola, 2018):

1. Es una variable adimensional, por lo que un valor Z es el número de desviaciones estándar, que un valor presenta con respecto a la media.
2. Es una medida puntual, puede estar indicada con signo de más (+) o menos (-) indicando la posición relativa a la izquierda de la media cuando es negativa o a la derecha de la media cuando es positiva.
3. Los valores con más de dos desviaciones estándar positivos o negativos se consideran datos atípicos o eventos raros.

Debido a las características de la puntuación Z, puede ser utilizada para comparar valores que proceden de diferentes poblaciones como el problema planteado en un inicio (Triola, 2018) por lo que se procede a desarrollar el ejemplo.

Se utilizan los datos obtenidos en el estudio piloto con padres de niños de 12 a 18 años de la Tabla 7. Se toma en cuenta los datos obtenidos en el trabajo de Barcenás Gaona y cols., (2017): media 9.02 y Desviación estándar: 4.87

Se utiliza la ecuación:

$$z = \frac{x - \bar{x}}{s}$$

Se despeja con el Dato I= 13:

$$z = \frac{13 - 9.02}{4.87}$$

Resultado:

$$Z= 0.82$$

Es de notar que el valor se redondea a dos dígitos después del punto decimal, esto debido a las características y el formato que presentan las tablas del puntaje Z, donde los valores se expresan con dos decimales.

Esto se realizaría para cada uno de los datos y podremos obtener los resultados presentados en la Tabla 8.

Tabla 8. Resultados obtenidos después de aplicar la fórmula de puntaje Z.

Valor = 13 Z = 0.82	Valor = 3 Z = -1.24	Valor = 7 Z = -0.41	Valor = 4 Z = -1.03	Valor = 6 Z = -0.62
Valor = 2 Z = -1.44	Valor = 8 Z = -0.21	Valor = 16 Z = 1.43	Valor = 15 Z = 1.23	Valor = 3 Z = -1.24

A partir de los resultados obtenidos podemos inferir que ninguno de los datos en el estudio piloto se debe considerar un dato atípico y esto permite deducir que los resultados del estudio piloto pueden ser comparados con los datos publicados por Barcenás Gaona y cols. (2017).

A manera de conclusión: Aplicación de las medidas de posición en la educación especial

A lo largo del presente documento se han descrito los procedimientos para el cálculo de las medidas de posición y las observaciones que se pueden hacer a partir de los resultados obtenidos, sin embargo, queda pendiente concretar cómo estas herramientas de la estadística descriptiva son aplicables en el ámbito de la educación especial. En otras palabras, qué información podría dar la obtención de resultados de cuantiles y puntaje Z en esta especialidad.

Con la Declaración de Salamanca por la UNESCO en 1994 se reconoce el derecho de todas las personas con necesidades educativas especiales a ser integradas en escuelas regulares, con la visión de desaparecer el estigma existente por los rasgos distintos, proponiendo un modelo de educación integradora (Cedillo y Contreras, 2019). Hasta antes de Salamanca, a todos los estudiantes diagnosticados con alguna deficiencia o discapacidad se les enviaba

a escuelas de educación especial, esto sin valorar las capacidades con las que sí contaban (Mateos, 2008). Este modelo educativo inclusivo pretende integrar a todas las personas a la educación, independientemente de sus características particulares, buscando intervenir en el proceso enseñanza-aprendizaje de tal manera que este favorezca a todos por igual.

Desde esta perspectiva, contar con herramientas estadísticas que describan las características de la población con la que se trabaja en el aula y situar las dificultades de aprendizaje, así como las habilidades de cada estudiante respecto al resto de sus pares y a su vez, respecto al norma poblacional; podría ser de utilidad al plantear programas educativos que incorporen recursos suficientes para que todos los alumnos adquieran los conocimientos escolares para su grado académico.

Retomando la interrogante de qué información podrían darnos las medidas de posición, un ejemplo se puede observar al aplicar un instrumento que evalúe ya sea el cociente intelectual, lenguaje, razonamiento matemático o cualquier otro dominio de interés. Mediante estas herramientas estadísticas se indicará la posición del individuo en relación con la muestra normativa, permitiendo evaluar el desempeño en comparación con otras personas; además, en el caso específico de la puntuación Z, proporciona medidas equivalentes que permiten la comparación del desempeño del individuo en diferentes pruebas (Anastasi y Urbina, 1988; Hernández Sampieri, Fernández Collado y Baptista Lucio, 2008). En términos generales, los resultados obtenidos mediante estas pruebas darían información sobre el grado de severidad de una alteración o el grado de desarrollo de una habilidad en un individuo.

Referencias bibliográficas

- Anastasi, A., y Urbina, S. (1998). *Tests psicológicos*. Pearson Educación.
- Anderson, D., Sweeney, D., y Williams, T. (2015). *Estadística para negocios y economía*. Cengage Learning Editores S.A. de C.V.
- Ballesteros Doncel, E. (2012). *Cuartiles, Deciles y Percentiles: Cálculo, aplicaciones y prácticas resueltas para enseñar y aprender*. Universidad Complutense de Madrid. <https://eprints.ucm.es/id/eprint/15707/1/eprint.pdf>
- Bárcena, S.X., Guevara, Y., y Álvarez, E. (2007). Características de padres e hijos con discapacidad que intervienen en la comunicación sobre sexualidad. *Journal of Behavior, Health y Social Issues*, 9 (2), 36-44. <https://doi.org/10.1016/j.jbhsi.2018.01.002>
- Cedillo, I. G., y Contreras, S. R. (2019). Influencia de la Declaración de Salamanca sobre la Atención a la Diversidad en México y Situación Actual. *Revista latinoamericana de educación inclusiva*, 13(2), 123-138.
- Cobo E., Muñoz P., González J. A., Bigorra J., Corchero C., Miras F., Selva A., y Videla S. (2007). *Estadística descriptiva*, Cobo E., Muñoz P., González J. A., Bigorra J., Corchero C., Miras F., Selva A., Videla S., *Bioestadística Para no Estadísticos*. Elsevier España.
- Henry, B. C., y Pellillo, B. C. (2019). *Estadística básica*. Cengage Learning Editores SA de CV.
- Hernández Sampieri, R., Fernández Collado, C., y Baptista Lucio, P. (2008). *Metodología de la investigación* (4ª edición). McGrawHill. <http://dx.doi.org/10.4067/S0718-73782019000200123>
- Jay, D. (2015). *Probabilidad y Estadística para Ingeniería y Ciencias* (9th ed.). Cengage Learning Editores SA de CV.
- Johnson, R., y Kuby, P. (2012). *Estadística elemental*. Cengage Learning Editores S.A. de C.V.

- Juares Garcia, F. (2014). *Apuntes para la Investigación en Salud*. (1er ed). Universidad Nacional Autónoma de México.
- Levin, R. I. (2010). *Estadística para Administración y economía* (7th ed.). Pearson HispanoAmerica.
- Levine, D. M., Krehbiel T. C. y Berenson, M. L. (2014). *Estadística: Para administración*. (6th ed). Pearson Educación de México, SA de CV.
- Mesa Guerrero, J. A., y Caicedo Zambrano, S. J. (2020). *Introducción a la estadística descriptiva*. Universidad de Nariño.
- Papis, G. M. (2008). Educación especial. *Revista Intercontinental de Psicología y Educación*. 10(1), 5-12. <http://www.redalyc.org/articulo.oa?id=80210101>
- Suárez, M. O. I., y Tapia, F. A. Z. (2012). *Interaprendizaje de estadística básica*. (2ª edición). Ibarra-Ecuador.
- Torres, R.H.S.Y.C.P. M. (2018). *Metodología de la investigación*. Las rutas de la investigación cuantitativa, cualitativa y mixta. McGraw-Hill Interamericana.
- Triola, M. F. (2013). *Estadística (11th ed.)*. Pearson Educación, México.
- Triola, M. F. (2018). *Estadística (12th ed.)*. Pearson HispanoAmerica.
- UNESCO. (1994). *Declaración de Salamanca y marco de acción*. http://www.unesco.org/education/pdf/SALAMA_S.PDF
- William, M. (2014). *Introducción a la Probabilidad y Estadística (14th ed.)*. Cengage Learning Editores SA de CV.

LA DISTRIBUCIÓN DE FRECUENCIAS Y SU REPRESENTACIÓN GRÁFICA EN LA EDUCACIÓN ESPECIAL

Claudia Amezcua Gutiérrez

Rosa María Hidalgo Aguirre [‡]

Jahaziel Molina del Río

Resumen

En el trabajo relacionado con la educación especial es importante realizar proyectos de investigación, a fin de responder preguntas cuyas respuestas no son inmediatamente obvias o cuantificables. Para lo cual, es indispensable contar con herramientas estadísticas que faciliten la organización, el análisis y la interpretación de datos que permitan evaluar críticamente, y visualizar de forma más clara la información procedente de poblaciones o muestras específicas. Esto es clave para la comparación y representación de datos en crudo, estandarizados o en percentiles de acuerdo con datos de la población en general. Dichos datos pueden ser presentados ya sea en tablas, histogramas, gráficas de barras o de pastel, polígonos de frecuencia, o en cajas de Tukey, dependiendo de la información que se desee visualizar. Finalmente, podemos mencionar que la representación gráfica nos puede llegar a facilitar la interpretación de datos acerca de algún padecimiento en particular, para mejorar la aplicación de tratamientos, entrenamientos cognitivos o terapias en la práctica profesional.

Palabras clave: Representación gráfica, distribución de frecuencias, JASP.

[‡]Laboratorio de Neuropsicología, Centro Universitario de los Valles, Universidad de Guadalajara. Carretera Ameca-Guadalajara, Km. 45.5, Ameca, Jalisco. rosa.hidalgo@academicos.udg.mx

Introducción

La realización de investigaciones en torno a la educación especial requiere presentar resultados cuantificables. La claridad de dicha presentación es de vital importancia para la comprensión de los resultados y la interpretación de éstos, además de realizar un análisis estadístico adecuado. Comúnmente, la presentación de datos numéricos se hace por medio de tablas, en ocasiones un diagrama o un gráfico pueden ayudar a representar de un modo más eficiente nuestros datos.

En este capítulo se describirán algunas técnicas de representación de las frecuencias de datos, así como los distintos tipos de gráficos que se pueden utilizar para presentar diversos datos en torno a la educación especial.

Normalmente, al llevar a cabo un estudio científico independientemente de si es cuantitativo o cualitativo, los datos que se suelen obtener son numerosos y requieren ser ordenados y organizados. Así pues, el primer paso que debemos efectuar ante cualquier estudio es sintetizar el conjunto de datos mediante tablas o gráficos, con el fin de poder identificar el comportamiento característico de un fenómeno y facilitar su análisis exhaustivo.

A continuación, se explicará a detalle que son las tablas de frecuencia, como se realizan y algunos ejemplos.

Distribución de frecuencias

Normalmente, al obtener datos de alguna investigación, obtenemos muchos datos que quizá en un inicio, no sepamos qué hacer con ellos o cómo analizarlos para poder interpretar lo que obtuvimos; la estadística descriptiva nos permitirá

describir la relación existente entre dos o más variables, mientras que la estadística inferencial, nos permitirá llegar a conclusiones acerca de una población, a partir del análisis de datos de una muestra.

En análisis descriptivos, se estudian los datos arrojados por un grupo determinado, sin referirnos a una muestra o población, por lo que se presentan las conclusiones referentes al grupo con el que se está trabajando. Una de las formas de presentar la distribución de datos, puede hacerse mediante la construcción de tablas de frecuencias.

Por ejemplo, al obtener grandes cantidades de datos, puede ser difícil el manejo de los mismos, principalmente la interpretación de la información, por lo que ordenar los datos en distribuciones de frecuencias nos permite observar los datos de manera ordenada y clara, podremos analizar cuántos datos se repiten, cuál es el dato medio, cuántos hay por encima o por debajo, también se podrá obtener el rango de los puntajes cuando se muestran como distribución de frecuencia.

Específicamente, cuando existen muchos datos y la amplitud de distribución de los mismos es muy amplia, enlistar los datos de manera individual suele arrojar diversos valores con una frecuencia de cero y suele ser difícil de visualizar la forma de distribución, así como su tendencia central (punto medio de distribución de los datos). En este tipo de situaciones, se recomienda construir tablas de frecuencia, para limitar el análisis de la información; al construir una tabla de frecuencia se deben considerar como elementos de la misma: el título, el cual debe contener la variable que se está analizando y el grupo de estudio evaluado (Mejía Vélez, 2016).

Además, se suelen agrupar los datos individuales en intervalos de clase y se presentan en una distribución de frecuencia de puntajes agrupados, para limitar o resumir los datos presentados, denominados intervalos. Al agruparlos, se debe considerar qué tan amplio necesitamos el límite para los grupos de datos, ya que cada vez que se agrupan los datos, parte de la información puede perderse, por ejemplo, si se acomodan en un intervalo muy grande, no sabemos hacia dónde se encuentran la mayor cantidad de datos, o si están distribuidos uniformemente, entre más grande sea el intervalo, mayor será la ambigüedad que se tenga de los datos y por obvias razones, entre más estrecho sea el intervalo, mayor será la fidelidad de los datos originales.

Cuando el intervalo es demasiado estrecho, se agrupan datos casi de manera individual, valores con frecuencia cero, por lo que se busca elegir un intervalo que no sea ni muy amplio ni demasiado estrecho. En la práctica, por lo regular, se determina la amplitud del intervalo mediante la división de la distribución en un rango de entre 10 y 20 intervalos, y aún dentro de ese rango, la cantidad específica de intervalos utilizados depende de la cantidad y el rango de los puntajes en bruto (Pagano, 2015), con las consideraciones antes mencionadas.

Tabla 1. Puntajes de cociente intelectual en niños con autismo

67	57	81	76	86	80	67
93	76	65	74	94	96	77
70	79	60	56	73	82	63
89	79	71	77	52	90	68
72	88	84	70	83	93	76
82	96	99	58	89	77	81
87	63	77	72	56	78	78
69	54	82	82	66	72	79
86	76	65	46	62	87	93
82	76	75	76	92	74	95

Para realizar un ejemplo de tabla de distribución de frecuencia de puntajes agrupados, se sugieren algunos pasos, usaremos los datos de la tabla 1 para agruparlos en intervalos:

Primero, se debe seleccionar el rango de los puntajes, este se define como la diferencia entre el dato mayor y el dato menor del conjunto de información, en este caso:

$$\text{Rango} = \text{el dato mayor} - \text{el dato menor} = 99 - 46 = 53$$

Después, se tiene que determinar la amplitud de cada intervalo de clase (i), para lo que se tiene que considerar la cantidad de intervalos de clase que se desean agrupar los datos, en este ejemplo quedarán en aproximadamente 10, por lo que:

$$i = \frac{\text{Rango}}{\# \text{ intervalos de clase}} = \frac{53}{10} = 5.3$$

Cuando el valor de i tiene decimales, se redondean los datos, el intervalo i a la misma cantidad de decimales que tienen los puntajes en bruto, para este ejemplo será 5 (sin decimales).

Posteriormente, se enlistan los límites de cada intervalo de clase, los mayores y menores, se colocan en el intervalo que contenga al valor dado. Al enlistar los intervalos, se comienza con el intervalo inferior, primero determinando el nivel inferior este intervalo. Se necesitan considerar dos cosas en este punto: a) el límite inferior de este intervalo debe ser tal que el intervalo contenga al menor puntaje y, b) es costumbre hacer que el límite inferior de este intervalo sea divisible entre el rango (i). Por lo que, al límite inferior, se le asigna el valor del menor puntaje en la distribución si es divisible entre i ; o en caso de no ser así, al límite inferior se le asigna el siguiente valor más bajo que sea divisible entre i .

Al enlistarlos, es necesario que se asegure de que sean continuos y mutuamente excluyentes, es decir, que no tengan puntajes que se contengan en otro intervalo. Algo que se debe tomar en cuenta, es que muchas veces los puntajes no necesariamente serán números enteros, pueden tener decimales, y en esos casos, se debe considerar que los límites reales para cada intervalo serán de .5 antes y .5 después del puntaje, por ejemplo, si nuestro intervalo va de 95-99, realmente abarcan de 94.5 a 99.5, y el resto se incluirán en los intervalos aledaños (Tabla 4).

Una vez realizada la tabla, se recomienda colocar una marca de conteo a los puntajes en bruto en los intervalos de clase apropiados, para asegurarse que no se están considerando o sumando en otro intervalo (esto no es obligatorio, solo es una opción para asegurarnos de una correcta distribución de datos y no repetir valores al asignarlos en los intervalos). Los puntajes deben ser registrados en los intervalos de clase apropiados a su valor, una vez que se realiza el acomodo, se hace el marcaje de conteo. Este es un procedimiento en el que se debe revisar la distribución de manera sistemática y para cada puntaje

en bruto, se anota una marca junto al intervalo que lo contiene, para saber cuántos puntajes hay en cada intervalo (Pagano, 2015).

1. En algunas ocasiones es necesario hacer la sumatoria de las marcas de conteo para cada intervalo, con el fin de obtener la frecuencia del intervalo, este dato es con el que comúnmente se trabaja para presentar datos agrupados, por ejemplo, Tabla 2 y 3.
2. Por último, se deben sumar las marcas de conteo, convirtiendo las frecuencias de cada intervalo.

Tabla 2. Datos de frecuencia de sexo, prevalencia y puntajes de inteligencia en niños con autismo (Duncan y Bishop, 2013).

	Frequency	Percent
Male	377	90.4
Female	40	9.6
Autism	363	87.1
ASD	30	7.2
Asperger's disorder	24	5.7
WASI	47	11.3
WISC-IV	20	4.8

Tabla 3. Datos de características demográficas de pacientes con autismo (Hong, Singh, Kalb, Ashkar y Landa, 2021).

	[ALL]	No ASD	ASD	N
	N = 412	N = 107	N = 305	
Age	25.1 (3.85)	24.6 (4.17)	25.3 (3.72)	412
12–20 months (Algorithm 1)	67 (16.3%)	23 (21.5%)	44 (14.4%)	
21–30 months with <5 words (Algorithm 1)	230 (55.8%)	38 (35.5%)	192 (63.0%)	
21–30 months with 5 words or more (Algorithm 2)	115 (27.9%)	46 (43.0%)	69 (22.6%)	
Sex:				412
Female	92 (22.3%)	27 (25.2%)	65 (21.3%)	
Male	320 (77.7%)	80 (74.8%)	240 (78.7%)	
Race:				412
White	204 (49.5%)	56 (52.3%)	148 (48.5%)	
Black	90 (21.8%)	23 (21.5%)	67 (22.0%)	
Asian	44 (10.7%)	7 (6.54%)	37 (12.1%)	
Hispanic/Spanish	12 (2.91%)	3 (2.80%)	9 (2.95%)	
Multi-Racial	29 (7.04%)	12 (11.2%)	17 (5.57%)	
Other	33 (8.01%)	6 (5.61%)	27 (8.85%)	

Tabla 4. Elaboración de los intervalos de clase y reales, así como la distribución de frecuencia para puntajes agrupados.

Intervalo de clase	Intervalos reales	Marca de conteo	Frecuencia
95-99	94.5-99.5	IIII	4
90-94	89.5-94.5	IIII I	6
85-89	84.5-89.5	IIII II	7
80-84	79.5-84.5	IIII IIII	10
75-79	74.5-79.5	IIII IIII IIII I	16
70-74	69.5-74.5	IIII IIII	9
65-69	64.5-69.5	IIII II	7
60-64	59.5-64.5	IIII	4
55-59	54.5-59.4	IIII	4
50-54	50.5-54.5	II	2
45-49	44.5-49.5	I	1
			N=70

Distribuciones de frecuencia relativa, de frecuencia acumulativa y de porcentaje acumulativo. Además de las tablas de distribución, se pueden agrupar los datos en frecuencias absolutas, relativas, y porcentajes acumulativos. Primero, se debe tener claro a qué nos referiremos con cada frecuencia; como distribución de la frecuencia absoluta denominamos al número de observaciones que se encuentran dentro de cada intervalo, por lo que la suma de todas las frecuencias absolutas, debe ser igual al número total de datos analizados.

Como frecuencia absoluta acumulada, entendemos el número de observaciones que está por debajo del límite superior de cada intervalo, con sumar los valores de las frecuencias absolutas de los intervalos menores. Mientras que la frecuencia relativa, se hace referencia a la proporción del número total de puntajes que se incluyen en cada intervalo, se obtiene dividiendo la frecuencia absoluta de cada intervalo entre el número de datos analizados, también suele expresarse como porcentaje, multiplicando cada proporción obtenida por 100.

La frecuencia relativa acumulada hace referencia a la proporción o porcentaje de datos que están incluidos por debajo del límite superior de cada intervalo, el cual se obtiene dividiendo la frecuencia absoluta acumulada entre el número total de datos, o acumulando la frecuencia relativa.

Para convertir una distribución de frecuencia en una de frecuencia relativa, la frecuencia de cada intervalo se divide entre la cantidad total de puntajes:

$$f \text{ relativa} = \frac{f}{N}$$

Si observamos la tabla 4, la frecuencia relativa para el intervalo de 45 a 49 puntos se obtiene al dividir su frecuencia entre la cantidad total de puntajes, para la cual es 70, entonces la frecuencia relativa para ese intervalo es= $1/70=0.01$. La frecuencia relativa es útil, ya que nos ayuda a identificar la proporción de datos contenidos en ese intervalo.

Mientras que la frecuencia acumulativa para cada intervalo se calcula sumando la frecuencia de ese intervalo de interés, con las frecuencias de todos los intervalos de clase inferiores a ese. Por lo que la frecuencia acumulativa para el intervalo de 70 a 74, sería de $9+7+4+4+2+1=27$.

El porcentaje acumulativo para cada intervalo se encuentra al convertir las frecuencias acumulativas en porcentajes acumulativos con esta fórmula:

$$\% \text{ acumulativo} = f \frac{\text{acumulativa}}{N} \times 100$$

Tabla 5. Distribuciones de frecuencia relativa, acumulativa y de porcentaje acumulativo para los puntajes de cociente intelectual agrupados de la tabla 1.

Intervalo de clase	<i>f</i>	<i>f</i> relativa	<i>f</i> acumulativa	% acumulativo
95-99	4	0.06	70	100
90-94	6	0.09	66	94.29
85-89	7	0.10	60	85.71
80-84	10	0.14	53	75.71
75-79	16	0.23	43	61.43
70-74	9	0.13	27	38.57
65-69	7	0.10	18	25.71
60-64	4	0.06	11	15.71
55-59	4	0.06	7	10.00
50-54	2	0.03	3	4.29
45-49	1	0.01	1	1.43
	N=70	1.00		

Usando el dato para el mismo intervalo del ejemplo anterior, de 70 a 74 sería,

$$\% \text{ acumulativo} = f \frac{\text{acumulativa}}{N} \times 100 = \frac{27}{70} \times 100 = 38.57 \%$$

Las distribuciones de frecuencia acumulativa y de porcentaje acumulativo son particularmente útiles para ubicar percentiles y rangos percentiles.

Percentiles. Particularmente, los percentiles son medidas de posición relativa, que suelen usarse en educación para comparar el desempeño de un individuo con el de un grupo de referencia, quizá en su rango de edad o con las características que tiene ese sujeto, que ya han sido evaluadas previamente en grandes cantidades de participantes. Se definen como el valor en la escala de medición por debajo del cual queda un porcentaje específico de los datos en la distribución. Por lo que, el 70-ésimo punto percentil es el valor en la escala de medición debajo del cual queda el 70 % de los datos en la distribución (Pagano, 2015).

Con las tablas previamente mencionadas, haremos un ejemplo, calcularemos el 50-ésimo punto percentil para los puntajes de cociente intelectual (Tabla 1). En la tabla 6, se presentan los datos como distribuciones de frecuencia acumulativa y de porcentaje acumulativo, para lo cual se utilizará el símbolo P_{50} para que se represente el 50-ésimo punto percentil, lo que significa que nos interesa saber el valor de la escala debajo del cual queda el 50 % de los datos, suponiendo que la distribución de éstos están distribuidos de manera uniforme en cada intervalo.

Si observamos la tabla 5, la frecuencia acumulada del intervalo 70 a 74, 27 de los puntajes quedan por debajo del valor 74.5, por lo que necesitamos llegar a 35, es decir requerimos 8 puntajes más para tener la mitad de los datos o el punto medio, que es el dato que estamos buscando (P_{50}). Realmente, existen 16 puntajes en el intervalo y cada uno, tiene una amplitud de cinco unidades de la escala, por lo que cada dato en el intervalo representa $5/16$ de una unidad de la escala. Por lo que, para llegar al percentil 50, necesitamos obtener ocho

puntajes más, necesitamos movernos dentro del intervalo $5/16 \times 8 = 2.5$ unidades, y al sumar 2.5 al límite inferior de 74.5, llegamos a P_{50} .

$$P_{50} = 74.5 + 2.5 = 77.0$$

Para encontrar un punto percentil, necesitamos:

1. Determinar la frecuencia de los puntajes por debajo del punto percentil dado.

$$f_P \text{ acum} = (\% \text{ puntajes}) \times N$$

Por lo que, para P_{50} :

$$f_P P_{50} \text{ acum} = 50 \% \times N = (0.50) \times (70) = 35$$

2. Determinar el límite real inferior del intervalo que contiene al punto percentil. Al conocer la cantidad de puntajes por debajo del punto percentil, podemos localizar el intervalo que lo contiene al comparar la $f_P \text{ acum}$ con la frecuencia acumulativa para cada intervalo. Una vez localizado el intervalo, podemos comprobar su límite real inferior: X_L . En este caso, el intervalo que lo contiene es el de 75 a 79 y su límite real inferior X_L es 74.5.
3. Determinar el número de puntajes adicionales que necesitamos obtener en el intervalo para alcanzar el punto percentil, es decir el 50 % de los datos.

Cantidad de puntajes adicionales $= f_P \text{ acum} - f_L \text{ acum}$, en donde $f_L \text{ acum} =$ la frecuencia de puntajes por debajo del límite real inferior del intervalo que contiene al punto percentil.

$$\text{Cantidad de puntajes adicionales} = f_P \text{ acum} - f_L \text{ acum} (35 - 27 = 8).$$

- Determinar la cantidad de unidades adicionales de la escala dentro del intervalo que debemos recorrer para obtener la cantidad adicional de puntajes:

Unidades adicionales

= (cantidad de unidades por puntaje) x cantidad de puntajes adicionales

$$= (i/f_i) \times \text{cantidad de puntajes adicionales}$$

$$= (5/16) \times 8 = 2.5$$

Considerando que, f_i es la cantidad de puntajes en el intervalo, e i/f_i nos da la cantidad de unidades por puntaje para el intervalo.

- Determinar el punto percentil, esto se consigue al sumar las unidades adicionales al límite real inferior del intervalo que contiene al punto percentil.

Punto percentil = X_L + Unidades adicionales

$$\text{Punto percentil} = X_L + (i/f_i) (f_P \text{ acum} - f_L \text{ acum})$$

$$P_{50} = 74.5 + \left(\frac{5}{16}\right) (35 - 27) = 74.5 + 2.5 = 77.00$$

Tabla 6. Cálculo de los puntos percentiles para los puntajes de cociente intelectual agrupados de la tabla 1.

Intervalo de clase	<i>f</i>	<i>f</i> acumulativa	% acumulativo	Cálculo del percentil
95-99	4	70	100	
90-94	6	66	94.29	
85-89	7	60	85.71	Punto percentil = $X_L + (i/f_i)(f_{pacum} - f_{Lacum})$
80-84	10	53	75.71	
75-79	16	43	61.43	
70-74	9	27	38.57	
65-69	7	18	25.71	$P_{50} = 74.5 + (5/16)(35-27) = 77.00$
60-64	4	11	15.71	
55-59	4	7	10.00	
50-54	2	3	4.29	
45-49	1	1	1.43	$P_{20} = 64.5 + (5/7)(14-11) = 66.64$
	N=70			

Rango percentil. Por último, el rango percentil es un contraste entre dos percentiles determinados, hace referencia al porcentaje de datos con valores inferiores al punto en cuestión, lo que nos permite evaluar la posición relativa del dato de interés en un conjunto de datos. Esto nos sirve para conocer el rango percentil de un dato en bruto, por ejemplo, dado que un dato en la evaluación del cociente intelectual fue 86, le resultará útil saber el rango percentil de 86.

Para calcular el rango percentil, tenemos el dato del puntaje, y ahora debemos calcular el porcentaje de datos inferiores a ese punto, por lo que asumimos que los puntajes dentro de cualquier intervalo están distribuidos uniformemente.

$$Rango\ percentil = f_L\ acum + \frac{\left(\frac{f_i}{i}\right)(X - X_L)}{N} \times 100$$

En donde:

- $f_L\ acum$ = frecuencia de puntajes por debajo del límite real inferior del intervalo que contiene al puntaje X.

- X = puntaje cuyo rango percentil será determinado
- XL = valor de la escala del límite real inferior del intervalo que contiene al puntaje X
- i = amplitud del intervalo
- f_i = frecuencia del intervalo que contiene al puntaje X
- N = número total de puntajes en bruto

Por lo que, para encontrar el rango percentil de 86, obtenemos:

$$\text{Rango percentil} = 53 + \frac{(77.5)(86 - 84.5)}{70} \times 100 = 78.81$$

Representación gráfica

La representación gráfica de los datos se utiliza para demostrar, con mayor claridad, las características de la distribución de los datos o de otras estadísticas; permitiendo así que el lector tenga un resumen visual de los resultados obtenidos (Coolican, 2005). Para elegir el tipo de gráfico adecuado es necesario saber el tipo de variable que vamos a analizar y su respectiva escala de medición.

Para las variables cualitativas se utilizan los gráficos de barras y de sectores; en el caso de las variables cuantitativas los gráficos más adecuados son gráfico de barras, el histograma y el polígono de frecuencias (Mejía Vélez, 2016).

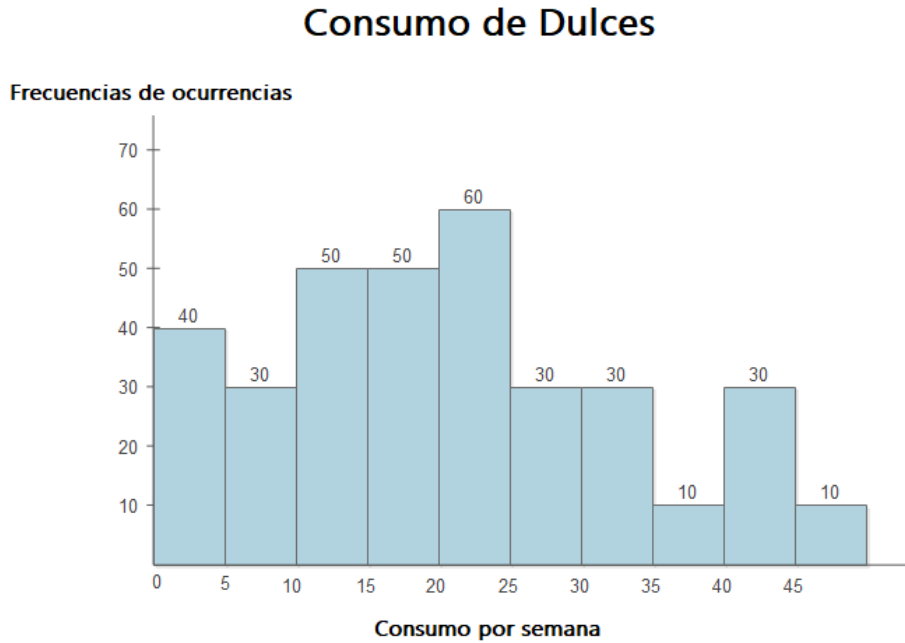
Existen muy diversos tipos de gráficas, generalmente aplicándose unas u otras en función de lo que se pretenda representar o simplemente de las preferencias del autor.

A continuación, indicamos algunas de las más conocidas y comunes.

Histograma de barra. Es una figura desarrollada en el plano cartesiano, en el que en el eje de las abscisas (X) se representan los intervalos de clase y en el eje de las ordenadas (Y) puede representarse la frecuencia absoluta (f_i) o la frecuencia relativa (h_i), de acuerdo con lo que quiera presentar el investigador (número de sujetos o porcentaje). Se forman barras que en el ancho representan la amplitud del intervalo y en la altura el número o porcentaje de observaciones dentro de cada intervalo (Mejía Vélez, 2016).

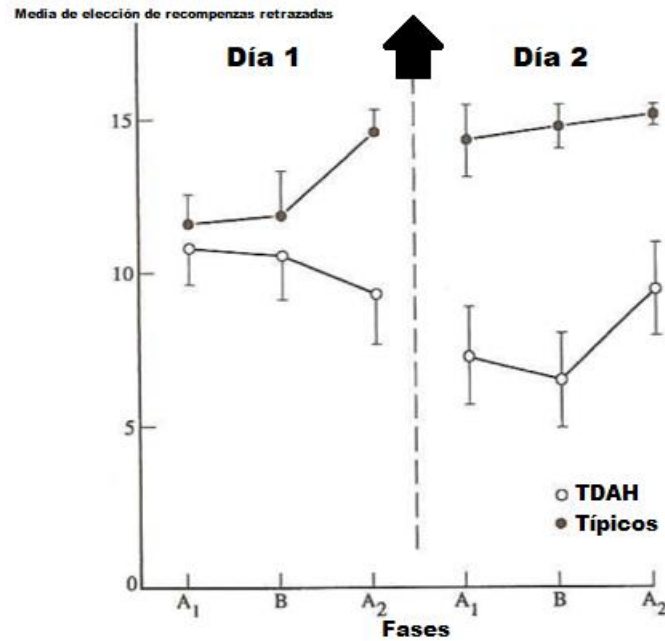
Ejemplo. - En la figura 1 se presenta el consumo de dulces por semana de un niño de 7 años con Déficit de Atención con Hiperactividad (TDAH). En el eje de las Y se representa la ocurrencia es decir cuantas veces comió dulces y en el eje de las X el consumo por semana. Como puede observarse en el histograma de la semana 20 a 25 el niño comió 60 dulces lo que probablemente tuvo un efecto en su hiperactividad y de la semana 35 a 40 solo 10. Al observar este tipo de datos el investigador puede explorar que pasó esas semanas que se incrementó o decrementó el consumo de dulces. Este hecho le permite tal vez crear estrategias de intervención.

Figura 1. Ejemplo de un histograma de barras



Polígono de frecuencias. Este gráfico es similar al histograma, pero se diferencian en que el polígono de frecuencias es un gráfico de líneas y no de barras; para graficarlo, se ubica en el eje “x” los intervalos de la variable que queremos graficar y en el eje “y” se ubica la frecuencia (Espinoza-Garzón, 2017).

Un ejemplo de cómo graficar un polígono de frecuencias se ilustra en la figura 2, presentada en un trabajo realizado por Schweitzer y Sulzer-Azarofft (1995). Ellas estudiaron niños entre 5 y 6 años típicos y niños con Trastorno por Déficit de Atención con Hiperactividad (TDAH) y evaluaron la preferencia por recompensas más grandes-retrasadas versus más inmediatas-pequeñas durante diferentes fases: A1, elección sin música o juguetes disponibles, B, elección con música y juguetes disponibles, y A2, elección sin música ni juguetes disponibles.

Figura 2. Media de la frecuencia de elecciones durante cada fase

Nota. La recompensa más grande retrasada se eligió de 16 ensayos para cada grupo. La fase B fue con música y juguete mientras que las otras no. Traducida de Schweitzer y Sulzer-Azaroff (1995)

De igual manera, en la a figura 2 se puede observar que los niños típicos o control, aumentaron su elección para las recompensas retrasadas el día 2 con respecto al día 1, mientras que los niños con TDAH disminuyeron su elección de las recompensas retrasadas del día1 al día 2. Los niños típicos eligieron las recompensas retrasadas más de la mitad del tiempo. Los niños con TDAH respondieron de forma más variable. Del total de 60 fases, eligieron las recompensas retrasadas alrededor de la mitad del tiempo (56,7%). En la gráfica también se puede observar el efecto de agregar estimulación durante la elección. Los niños con TDAH eligieron recompensas retrasadas más grandes significativamente menos que los chicos típicos. Agregar estimulación redujo el comportamiento motor, pero no afectó la elección.

A medida que aumentaron las tasas de actividad motora en los niños con TDAH, su preferencia por las recompensas más grandes disminuyó con el tiempo.

Cabe destacar que la gráfica está representando la media de la frecuencia y el error estándar. Estos conceptos fueron explicados en este volumen en la Parte II. Estadística descriptiva, capítulo II.1 Implicación de las medidas de tendencia central y de dispersión en la educación especial.

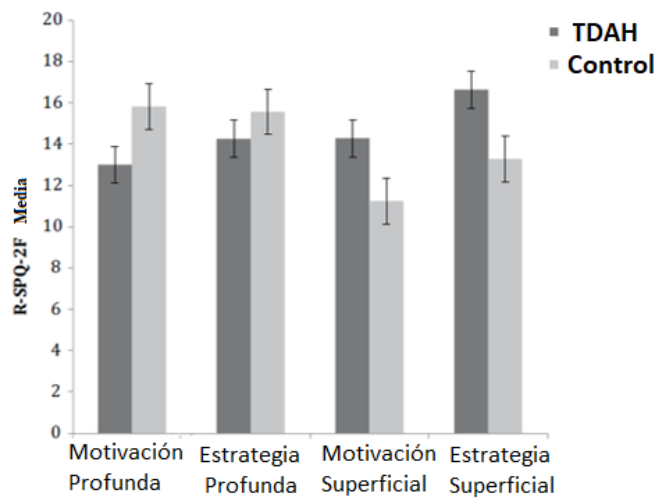
Gráfico de barras. El más conocido y utilizado de todos los tipos de gráficos es el gráfico o diagrama de barras. En éste, se presentan los datos en forma de barras contenidas en dos ejes cartesianos (coordenada y abscisa) que indican los diferentes valores de la variable. Las categorías de la variable se ubican en el eje X o abscisas y los valores de la variable presentados como frecuencias, porcentajes, rangos, entre otros, se ubican en el eje Y u ordenada (Coolican, 2005; Mejía Vélez, 2016).

Para interpretar este gráfico Ritchey (2006) propone algunas recomendaciones. La primera tiene que ver con la altura de las barras, la barra más alta indica que es la categoría que tiene la frecuencia más alta; en segundo lugar, es recomendable hacer comentarios sobre el orden de las categorías, bien sea comparándolas entre sí o mencionando las categorías espacialmente altas o bajas.

Para ejemplificar este tipo de gráficas describiremos una investigación realizada por Dack y colaboradores (2014) titulada *Study Habits, Motives, and Strategies of College Students with Symptoms of ADHD*. El objetivo del estudio fue examinar la relación entre los síntomas del TDAH y el rendimiento académico del estudiante al considerar el enfoque y la motivación del estudiante

para aprender. Participaron 361 estudiantes universitarios quienes completaron un cuestionario que desglosa las estrategias de aprendizaje en un enfoque superficial o profundo. Cada enfoque se divide en una motivación y una estrategia profunda o superficial.

Figura 3. Medias y errores estándar para las cuatro subcategorías del R-SPQ-2F. R-SPQ-2F = Cuestionario revisado del proceso de estudio de dos factores. (Traducido de Dack y cols., 2014.)



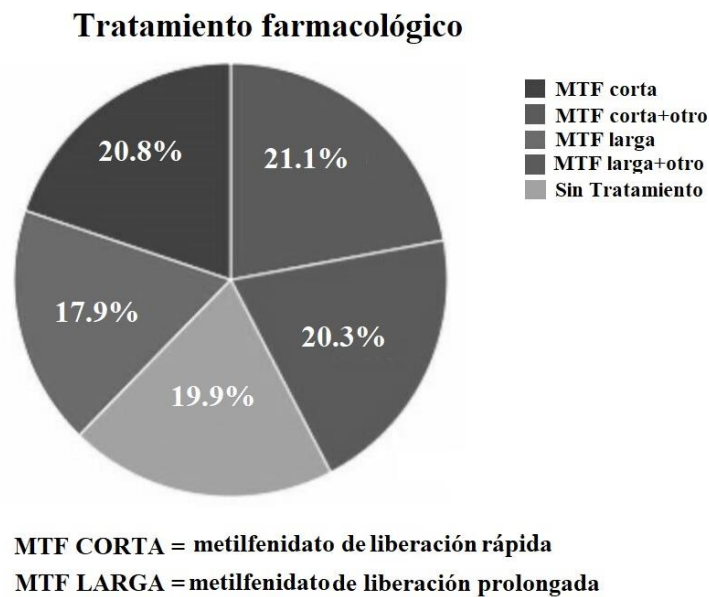
De manera general la gráfica (figura 3) revela que el grupo control utiliza más motivación profunda que el grupo TDAH, y que este grupo utiliza más motivación y estrategias superficiales en comparación con el control.

Gráfico de sectores o de pastel. Esta figura consiste en dividir una circunferencia en porciones, cada porción debe coincidir con el porcentaje de observaciones en cada una de las categorías (Mejía Vélez, 2016); por lo que la variable es cualitativa.

Para realizar la interpretación de este gráfico debemos concentrarnos en la porción más grande o con mayor porcentaje, también podemos comparar entre si los porcentajes de las categorías (Ritchey, 2006).

Siguiendo con la evaluación del TDAH, en el siguiente ejemplo ficticio se les preguntó a padres de niños con TDAH el tipo de tratamiento farmacológico que se les da a sus hijos. Los niños con TDAH y el porcentaje de niños que recibe cada uno de estos tratamientos.

Figura 4. Gráfica de pastel.



Como puede observarse en la figura 4 el mayor porcentaje de niños recibe metilfenidato de liberación rápida en combinación con otro medicamento. Casi el mismo porcentaje de niños recibe metilfenidato de liberación prolongada en combinación con otro y el menor porcentaje de niños toma metilfenidato de liberación prolongada (17.9%). Sólo el 19.9% de los niños con TDAH no recibe tratamiento.

Diagrama de caja y bigotes de Tukey. El diagrama de caja y bigotes, también conocido como gráfica de caja trama, o simplemente por el nombre de “boxplot”, el cual proviene de su denominación en inglés, fue desarrollado por John Tukey en los 70s. Es una herramienta gráfica conveniente en el análisis descriptivo para mostrar un grupo o grupos de datos numéricos a través de sus medianas, medias, cuartiles y mínimos y máximas observaciones. Un diagrama de caja es útil para mostrar la distribución de datos continuos, examinar la simetría e indicar valores atípicos potenciales y también se puede utilizar para comparar grupos de datos (Benjamini, 1988).

En su forma más simple, el gráfico de caja representa cinco datos estadísticos en una sola dimensión: el valor mínimo, el cuartil inferior (percentil 25), la mediana, la media, el cuartil superior (percentil 75) y el valor máximo.

Los cuartiles son los valores que dividen los datos en cuatro partes iguales o cuartos. Los percentiles, en cambio, dividen los datos en cien partes iguales. De esta forma, el cuartil inferior (Q1) se corresponde con el percentil 25 y el cuartil superior (Q3) con el percentil 75 (Liu, 2008).

Caja: la longitud de la caja representa la distancia entre los percentiles 25 y 75: el tramo de la escala que va desde el primer cuartil al tercer cuartil de la distribución (rango intercuartílico o rango intercuartil; RIC). Esto incluye el 50 % de las observaciones centrales.

- Punto interior de la caja: representa la media.
- Línea interior de la caja: representa la mediana.

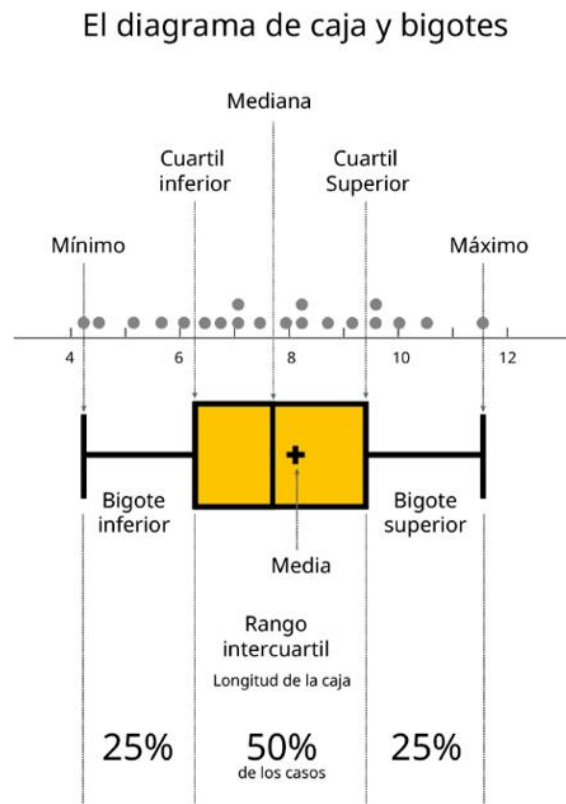
Líneas verticales (bigotes): representan los valores máximos (corte superior) y mínimos (corte inferior) del conjunto de datos. Sirven como referencia para ubicar las observaciones que están por fuera del 50% central de

la distribución. Por tanto, estos diagramas descriptivos son muy valiosos para observar los valores atípicos ('outliers') y las dispersiones de los puntos con respecto a los percentiles.

Todos estos datos pueden observarse gráficamente en la figura 5.



Figura 5. Diagrama de caja y bigotes y sus componentes.



Fuente: (2017, juantoxocruz.com) <http://juantoxocruz.com/charts-diagrama-caja-bigotes-john-w-tukey/>

Básicamente el diagrama de caja tiene grandes ventajas:

- 1.- Muestra información detallada sobre las observaciones en las colas. Si hay un interés en el valor de una observación, por lo general es uno en la cola.

- 2.- Las distribuciones de muchos lotes o grupos de datos pueden ser fácilmente comparadas mostrando sus diagramas de caja uno al lado del otro. Además, el diagrama de caja es fácil de calcular.
- 3.- Su significado o interpretación pueden ser aplicados por usuarios de estadísticas que, por lo general, los dibujan e incorporan en sus publicaciones (Benjamini, 1988).

¿Cómo hacer un diagrama de caja? Para realizar un diagrama de caja y bigotes, debemos seguir una serie de pasos al pie de la letra, tal como indicamos a continuación:

- 1.- Ordenar los valores: cada uno de los datos tiene su propio valor numérico, los cuales deben organizarse desde el mínimo hasta el máximo.
- 2.- Calcular los cuartiles: cuenta cada uno de los valores por unidad y divide el total entre 4, para obtener una media aritmética.
- 3.- Diseñar la caja: trázala entre el Q1 y el Q3.
- 4.- Ubicar la media y el RIC: recuerda que la primera está situada en el Q2, mientras el rango intercuartilíco va desde el Q1 hasta el Q3.
- 5.- Colocar los límites: el límite inferior (Li), se calcula restándole a Q1 la cantidad equivalente a 1,5 veces el valor del rango intercuartilíco, mientras el límite superior se obtiene sumándole esa misma cifra al Q3.
- 6.- Dibujar los brazos o bigotes: haz una línea desde el Q1 hasta el valor mínimo, y otra del Q3 al máximo.
- 7.- Escribir los valores atípicos: también se les conoce como “extremos” y van por fuera de cualquiera de los límites. Se distinguen con un punto, un “0” o un “*”.

Ejemplo de uso del Diagrama de Caja

A continuación, observemos un ejemplo donde se siguen los pasos necesarios para intentar calcular el puntaje de 20 niños en una prueba de atención, considerando las medidas actuales de cada uno, las cuales se muestran en la tabla 6:

Tabla 6. Puntaje total de los niños

Puntaje total									
7,13	5,35	10,14	4,15	5,85	4,69	6,84	0,94	3,45	5,79
5,64	4,92	5,32	6,56	6,17	4,33	6,18	6,50	3,74	2,98

Si ordenamos las medidas de menor a mayor, quedarían dispuestos de la siguiente manera (ver tabla 7):

Tabla 7. Puntaje total de los niños ordenado

Puntaje total - ordenado									
0,94	2,98	3,45	3,74	<u>4,15</u>	<u>4,33</u>	4,69	4,92	5,32	<u>5,35</u>
<u>5,64</u>	5,79	5,85	6,17	<u>6,18</u>	<u>6,5</u>	6,56	6,84	7,13	10,14

Q_2 (under 5,64) Q_1 (under 6,18) Q_3 (under 6,5) Q_2 (under 10,14)

Después de organizar los datos, se deben calcular los tres cuartiles y la mediana, así que las fórmulas serían las siguientes:

$$Q_2 = \text{Mediana}(X) = \frac{X_{(N+1)}}{2} = X_{10,5}$$

$$Q_2 = \frac{X_{10} + X_{11}}{2} = \frac{5,35 + 5,64}{2} = 5,50$$

$$Q_1 = X_{(N+1)/4} = X_{5,25} = X_5 + 0,25 \times (X_6 - X_5)$$

$$Q_1 = 4,15 + 0,25 \times (4,33 - 4,15) = 4,20$$

$$Q_3 = X_{\frac{3(N+1)}{4}} = X_{15,75} = X_{15} + 0,75 \times (X_{16} - X_{15})$$

$$Q_3 = 6,18 + 0,75 \times (6,50 - 6,18) = 6,42$$

Se calcula el rango intercuartílico así:

$$IQR = Q_3 - Q_1 = 6,42 - 4,20 = 2,22$$

Posteriormente con las siguientes formulas se calculan los 2 límites

$$LI = Q_1 - 1,5 \times IQR = 4,20 - 1,5 \times 2,22 = 0,96$$

$$LS = Q_3 + 1,5 \times IQR = 6,42 + 1,5 \times 2,22 = 9,59$$

Se dibujan los dos brazos o bigotes: desde el Q1 hasta el LI y del Q3 al LS.

De tal manera, la información del diagrama debería ser esta:

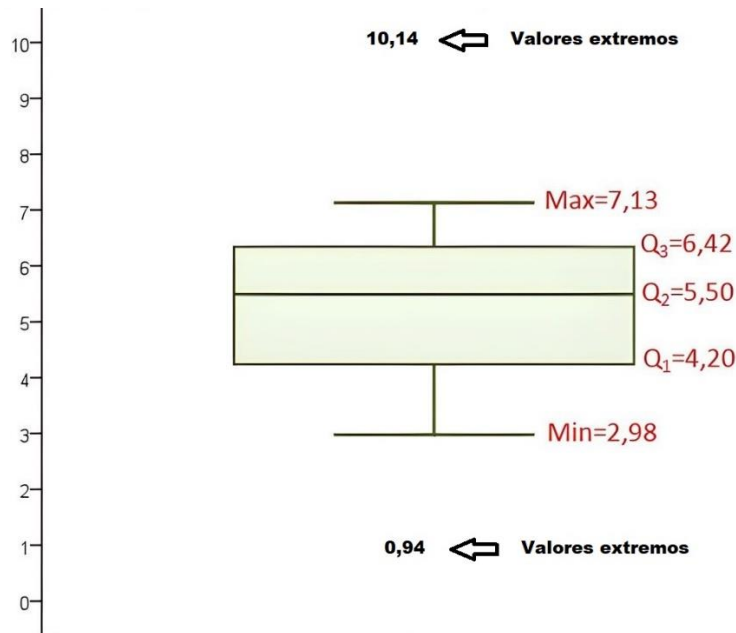
Niños con puntajes extremos: uno en el LI con un puntaje de 0,94 y otro en el LS que puntuó 10,14 metros. Ambos identificados con puntos, 0, o * en el diagrama.

Puntaje mínimo: 2,98

Puntaje máximo: 7,13

En conclusión, el diagrama de cajas como debería quedar se presenta en la figura 8.

Figura 8. Diagrama de caja de Tukey resultante.



Fuente: Modificado de <https://plantillaarbolgenealogico.net/diagramas/caja/>

Elaboración de tablas y gráficas en software libre

En la actualidad las tecnologías de la información y la comunicación han dotado a los investigadores de diversas herramientas que facilitan la visualización y la divulgación de los datos, permitiendo que sus resultados tengan un formato más accesible inclusive para personas no familiarizadas con la temática abordada.

Diversos son los softwares que podemos utilizar para la organización, análisis y graficación de datos, como pueden ser Excel, SigmaPlot y SPSS, sin embargo, no son opciones de uso libre, por otra parte encontramos software como R Studio, interfaz propia del lenguaje de programación R, diseñado

principalmente para el tratamiento estadístico de bases de datos, pero su uso requiere de conocimiento de programación, finalmente derivado del lenguaje R se han elaborado diferentes recursos de uso libre como pueden ser JAMOVI y JASP, entre otros.

Para el adecuado trabajo en cualquiera de los softwares de análisis estadístico se debe partir de una base de datos organizada propiamente, algunos de los programas permiten la elaboración de la base mediante el vaciado directo, otros como el que ejemplificaremos, requieren contar una base de datos previa, elaborada a manera de tabla en cualquier programa de hojas de cálculo o editor de texto. Lo anterior ofrece la libertad de recopilar los datos de la forma que mejor favorezca de acuerdo con su método, los cuales posteriormente se recopilaran y organizaran en un nuevo documento, mismo que al estar en un formato general podría emplearse en diferentes programas.

La organización adecuada de una base de datos de acuerdo con los programas descritos requiere que en las filas se coloquen los elementos de los que se obtienen los datos, comúnmente los participantes, en caso de que se realiza un análisis de prevalencia de alguna variable, las filas estarían ocupadas por las instituciones o inclusive las aulas, es decir, el elemento sobre el que se analiza la prevalencia; por su parte en las columnas de deben colar los datos de cada una de las mediciones que se realizaron sobre los elementos colocados en las filas.

En un ejemplo respecto al primer tipo de base de datos en los que las filas se ocupan por los participantes y las columnas por sus datos podríamos encontrar el siguiente caso (Tabla 7).

Tabla 7. Ejemplo de base de datos de muestras por sujetos o casos.

ID	Sexo	Edad	Escolaridad	Diagnóstico	CI
C01	M	8	3	TDAH	76
C02	F	6	1	TEA	75
C03	M	7	2	TDAH	76
C04	M	6	1	TEA	74
C05	F	8	3	TDAH	73

Como se puede apreciar la primera fila de la tabla anterior corresponde a los títulos de las variables que se colocan, seguido de los datos de cada uno de los sujetos registrados, las columnas incorporan los datos de cada una de las mediciones, los cuales como se ha mencionado pueden ser nominales, ordinales o escalares. La primera columna representa la identificación de cada participante, para estos casos se recomienda trabajar con códigos que nos permitan identificar las características de la muestra, cuando se realiza análisis de casos la clasificación sirve para llevar un control del número de casos, en este ejemplo se coloca la inicial “C” para referirse a “Caso” y el numeral correspondiente, cuando se registran diferentes condiciones o grupos las iniciales deberían incluir estas características, por ejemplo si fuera un grupo Rural y otro Urbano se podrían colocar los códigos “CU01” para caso urbano número uno y “CR01” para caso rural número uno. Aunque la selección de las etiquetas es arbitraria se recomienda mantener un código para su fácil identificación y manejo. La segunda columna de la tabla contiene la información del sexo de los participantes, siendo este un dato nominal la etiqueta podría incluir el término en extenso (i.e. Masculino, Femenino, Otro, etc.), sin embargo, para el vaciado y manejo de los datos se sugiere utilizar códigos cortos y concisos, en este particular caso las iniciales serían suficiente. La quinta columna del ejemplo contiene información del diagnóstico que pudieran tener los casos, al igual que la condición anterior se tratan de datos

nominales a los que corresponden las sugerencias previas de códigos cortos y concisos. La tercera, cuarta y sexta columna presentan datos escalares, los cuales implican una medición numérica de alguna de las variables de estudio. Para el caso de la edad los valores podrían ser de intervalo, sin embargo los decimales no son equivalentes a la edad en meses, por lo que es recomendable colocar solo años cumplidos, de lo contrario, o para estudios con población infantil, se puede colocar la edad en meses directamente; esta situación aplica también para el caso de la escolaridad, en este caso se podría especificar en la etiqueta de la columna si se habla de años de escolaridad o grado escolar, evitando así colocar la etiqueta en las celdas de los datos (i.e. “1 de primaria”), puesto que al incluir caracteres no numéricos el programa de análisis no reconocerá la celda como un dato numérico y podría eliminarlo. Finalmente, la última columna incorpora como ejemplo una medición del C.I. de los participantes, con las características de una medición numérica de razón como les comentadas previamente.

Como se comentó anteriormente las bases de datos pueden corresponder a diferentes elementos de los que se recaban datos, el anterior ejemplo correspondería a una situación de registro de participantes o casos, pero esos elementos de los que tomamos una muestra podrían ser instituciones o grupos que se desean describir, para estos casos la organización de los datos, aunque sigue la misma lógica podría verse con un orden distinto, como se ejemplifica a continuación (Tabla 8).

Tabla 8. Ejemplo de base de datos de muestras por grupos.

ID	Masculino	Femenino	TEL	TDAH	TEA
G01	16	14	4	7	1
G02	13	14	2	4	0
G03	17	15	3	6	1
G03	15	15	4	6	1
G05	14	16	3	5	0

En este último ejemplo la primera columna la ocupa el identificador del elemento de quien recopilamos los datos, a diferencia del ejemplo anterior en el que correspondía cada fila a un caso o participante, aquí se ejemplifican un conjunto de valores de diferentes grupos de la muestra, como podría ser el número de condiciones que se presenten en cada uno de los grupos de una escuela, en este caso los grupos se representan con el código con la inicial G de grupo y el numeral correspondiente. En esta tabla las columnas corresponden a cada una de las etiquetas de las variables que se están cuantificando, colocando en la celda correspondiente los valores de la frecuencia en que se presentaban dichas variables dentro del grupo.

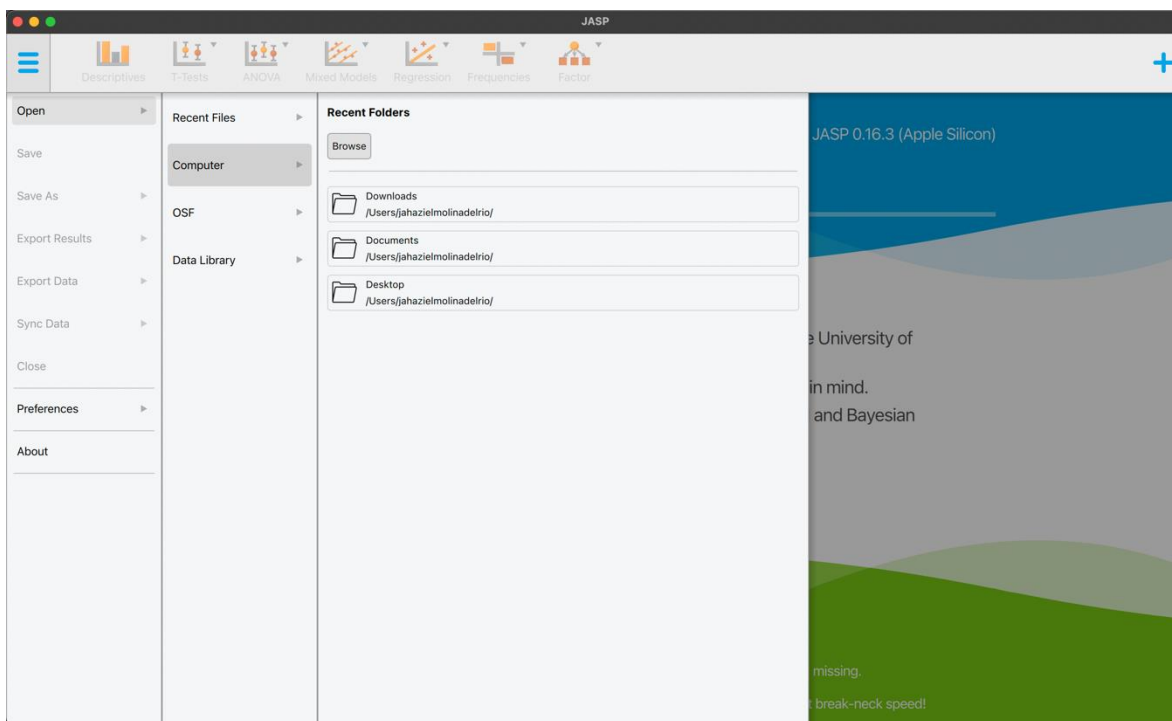
Una vez realizada nuestra base de datos de acuerdo con la característica de los datos que recolectamos, tendremos que exportarla en un archivo que pueda ser reconocido por los diferentes programas que podemos emplear para el análisis, en estos casos el formato más apropiado para las tablas de datos es el formato de valores separados por comas, al cual corresponde la extensión “.CSV”. Para esto en la mayoría de los programas de tablas de datos se puede elegir en la opción “guardar como” para posteriormente seleccionar el formato sugerido.

De entre los diversos softwares mencionados para el análisis de datos, presentaremos aquí el ejemplo utilizando el software JASP (2022), el cual es un

software de uso libre basado en el lenguaje de R para el análisis estadístico y graficación de datos, el cual puede descargarse de la siguiente dirección web <https://jasp-stats.org>.

La pantalla inicial del programa solicitará la apertura de nuestro archivo con la base de datos para poder trabajar con ella, en la esquina superior izquierda aparecerá un icono con tres rayas horizontales que al presionarlo desplegará diversas opciones, de las cuales debemos seleccionar la opción “Abrir” (Open), posteriormente “Computadora” (Computer) y después la opción “Navegar” (Browse), la cual nos desplegará una ventana de directorios en el cual debemos buscar la carpeta que contiene nuestro archivo .CSV y seleccionarlo para abrirlo dando clic en “Abrir” (Open), lo anterior se puede visualizar en la siguiente imagen (Figura 9).

Figura 9. Impresión de pantalla del menú para elegir el archivo con la base de datos.



Una vez importada la base de datos, esta se desplegará en la pantalla del programa, considerando la primera fila como las etiquetas y el resto como los valores en el mismo formato que los vaciamos, el tipo de datos se indica en la parte superior de cada columna, los cuales, aunque el programa identifica su origen, podemos modificar en caso de ser necesario, dando clic en el icono que se encuentra a la izquierda de la etiqueta, como se muestra en la siguiente imagen (Figura 10).

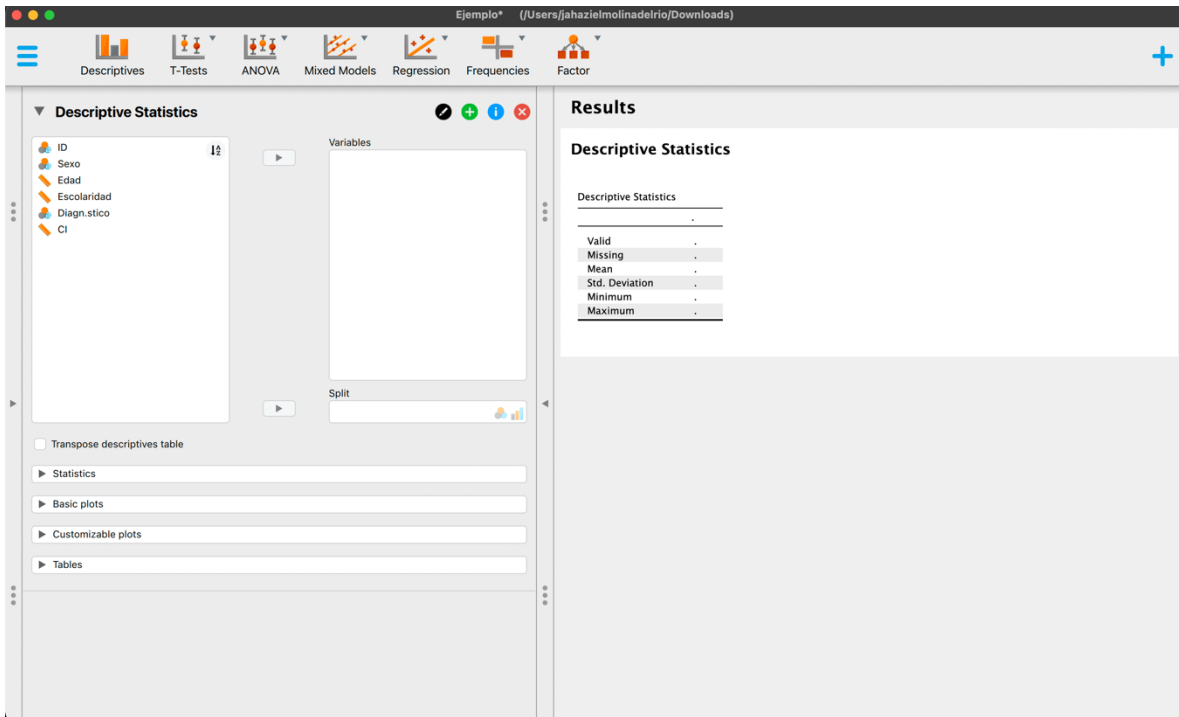
Figura 10. Impresión de pantalla de la visualización de la base de datos en el programa.

ID	Sexo	Edad	Escolaridad	Diagn.stico	CI	
1	C01	M	3	TDAH	76	
2	C02	F	1	TEA	75	
3	C03	M	2	TDAH	76	
4	C04	M	6	1	TEA	74
5	C05	F	8	3	TDAH	73
6	C06	M	8	3	TDAH	78
7	C07	F	6	1	TEA	78
8	C08	M	7	2	TDAH	79
9	C09	M	6	1	TEA	79
10	C10	F	8	3	TDAH	80
11	C11	M	8	2	TDAH	71
12	C12	F	9	1	TEA	65
13	C13	M	9	4	TDAH	68
14	C14	M	9	4	TEA	74
15	C15	F	7	4	TDAH	71
16	C16	M	7	2	TDAH	65
17	C17	F	7	1	TEA	77
18	C18	M	9	1	TDAH	74
19	C19	M	9	2	TEA	65
20	C20	F	6	3	TDAH	78

Como se aprecia en la imagen previa, en la parte superior de la ventana aparecen las diferentes opciones de análisis que podemos realizar (Descriptivos, Test-t, ANOVA, etc.), para el análisis de frecuencias elegiremos los análisis descriptivos, presionando en el ícono correspondiente, lo cual nos dirigirá a la ventana de análisis en la cual debemos elegir las variables que queremos análisis

y los análisis que deseamos se realicen de cada uno o varios de los datos que seleccionemos.

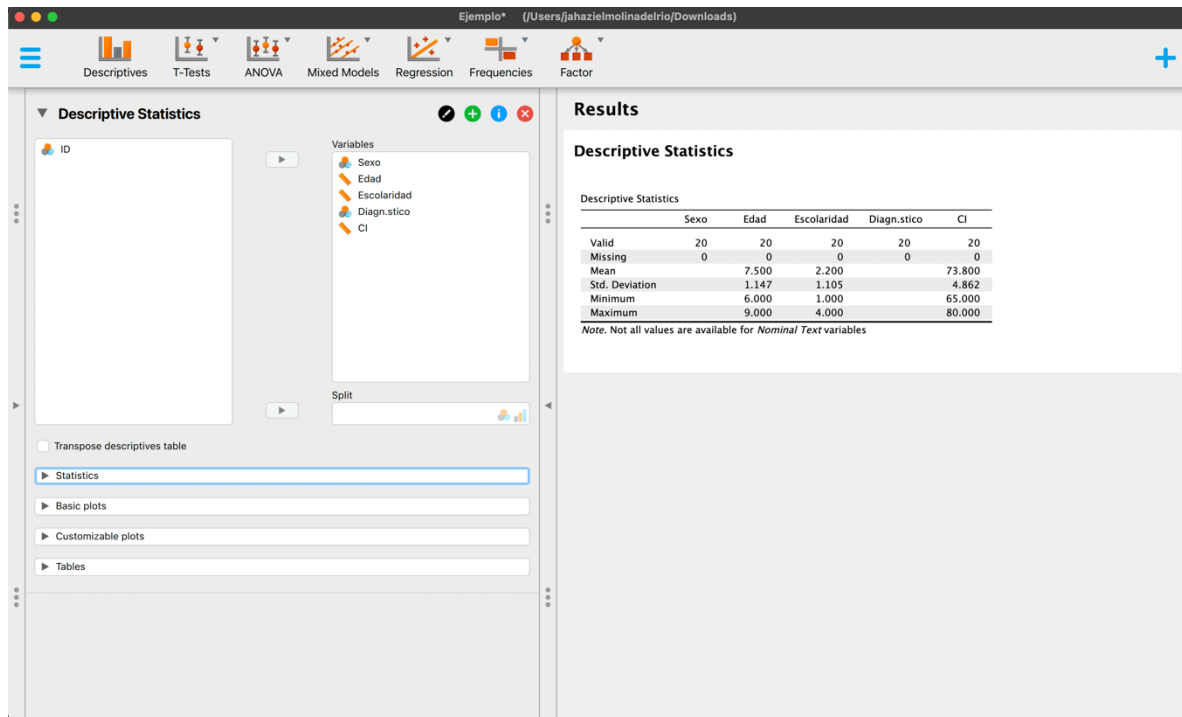
Figura 11. Impresión de pantalla al elegir la opción de estadísticas descriptivas.



Una vez abierta la ventana correspondiente al análisis descriptivo (Figura 11) podrán verse en la parte izquierda dos ventanas, en la primera se muestra las distintas variables que se registraron en la base de datos, de las cuales podemos ir eligiendo una o varias para mediante el botón del triángulo situado entre los dos recuadros se transfieran al área correspondiente a las variables sobre las que se realizarán los análisis, las cuales pueden agruparse de acuerdo con alguna otra variable (por ejemplo sexo o diagnóstico) que se colocaría en el recuadro de “Split”, nótese que los iconos que se encuentran en el recuadro de Split refieren que los datos para la división deben ser ordinales o nominales,

mientras que los que se analizarán en las variables deberán ser solo ordinales o escalares. En el área de resultados, localizada en la porción derecha de la ventana, se irán visualizando, de manera casi inmediata, los análisis correspondientes conforme se vayan seleccionando los datos (Figura 12).

Figura 12. Impresión de pantalla de los resultados a los que se le aplicaron estadística descriptiva.

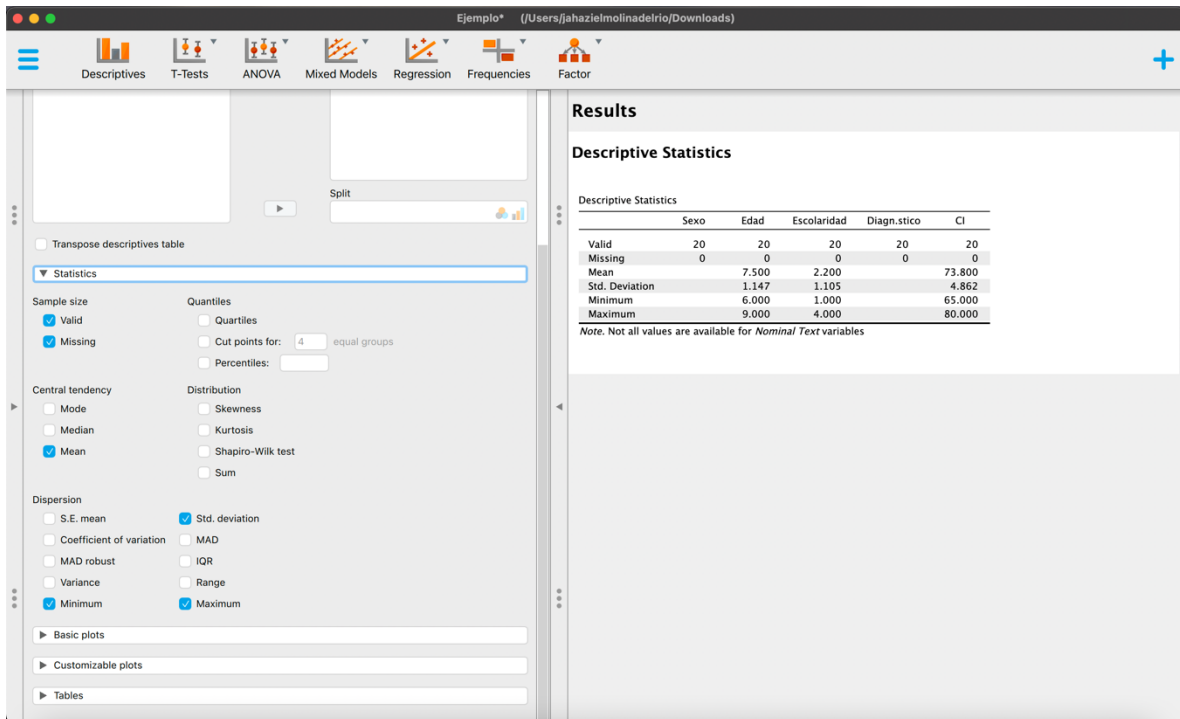


La imagen anterior muestra el resultado de la estadística descriptiva de las variables seleccionadas, como puede observarse los valores dependerán del tipo de variable, por ejemplo, en el caso de las nominales solo se presentan el conteo de casos válidos y perdidos, por otra parte, para las variables escalares se presentan además de los anteriores la media, desviación estándar, mínimo y máximo.

Debajo de las ventanas para la selección de variables, existen diferentes pestañas que se pueden desplegar para acceder a más opciones de análisis y visualización de los datos, mediante las cuales podemos obtener las diferentes gráficas indicadas anteriormente.

La primera pestaña despliega una serie de análisis descriptivos que podemos ir incorporando o modificando de la tabla previa, tales como cuartiles, medidas de tendencia central, distribución y dispersión (Figura 13).

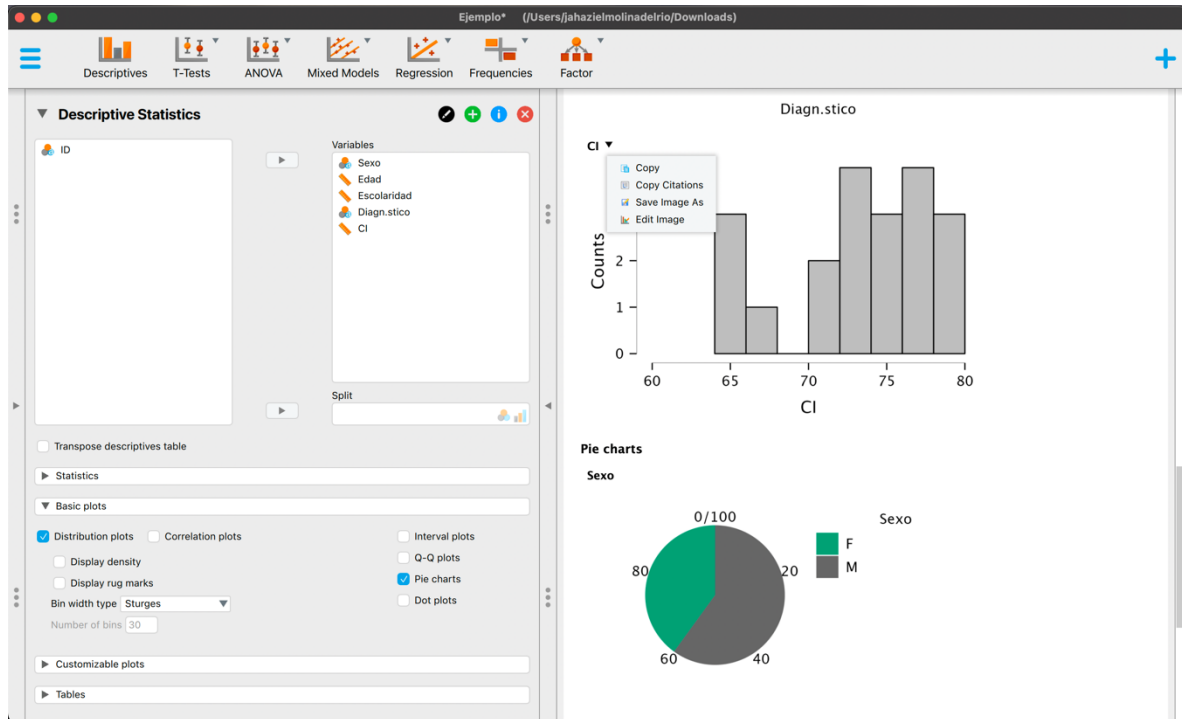
Figura 13. Impresión de pantalla de las opciones de estadísticas desplegada.



La siguiente pestaña presenta las opciones de gráficas básicas, entre las cuales podemos elegir graficas de distribución representadas con gráficas de barras, gráficas de correlación, gráficas de intervalos, Graficas Q-Q, gráficas de pastel y gráficas de puntos. Al seleccionar una o varias opciones éstas se irán

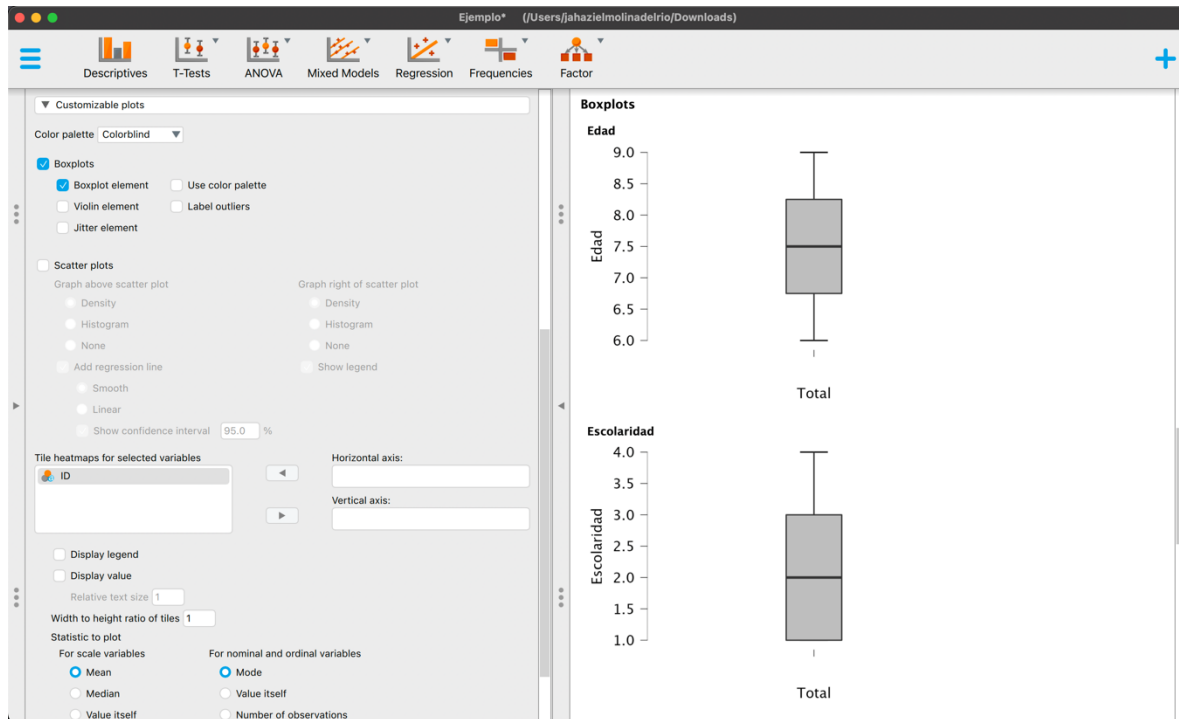
desplegando en la ventana de resultados, en la cual podemos seleccionar cualquiera de las imágenes generadas para copiarla, guardarla o editar algunos de sus elementos, como los rangos, nombres de los ejes y títulos (Figura 14).

Figura 14. Impresión de pantalla de las selecciones de la opción de gráficas básicas.



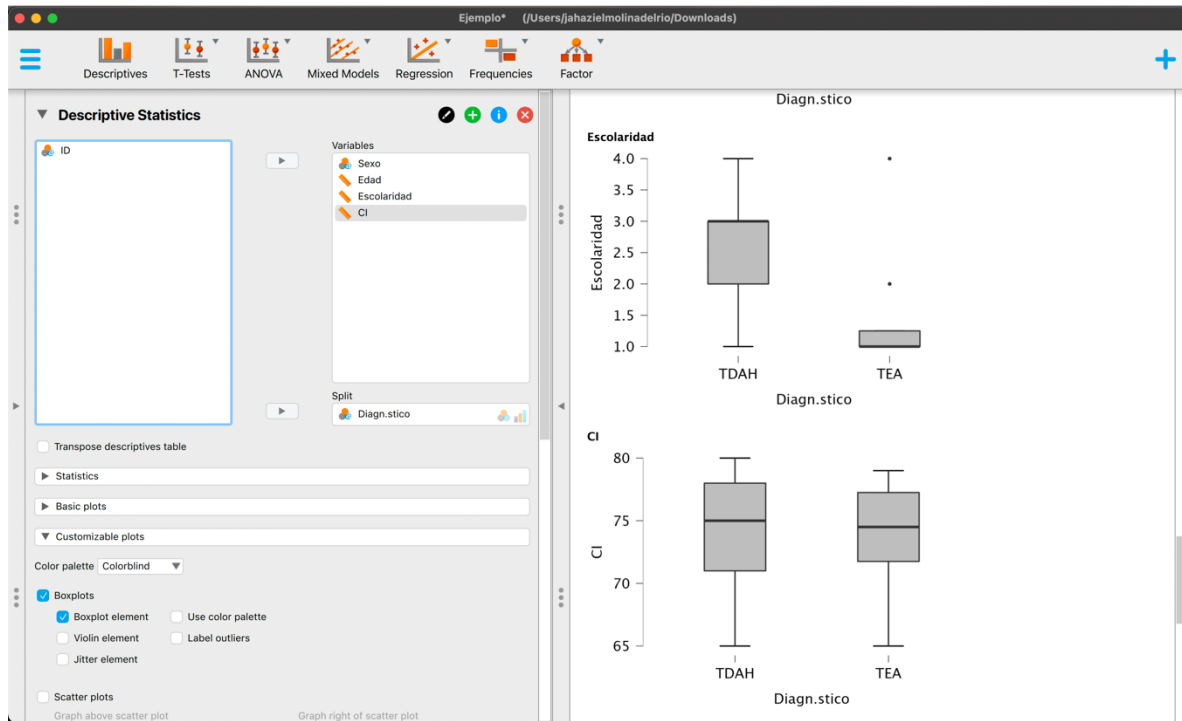
La pestaña siguiente contiene las gráficas personalizables, entre las que se encuentran los boxplots, gráficos de dispersión, así como gráficas de mosaicos. Cabe señalar que algunas de las funciones que cada tipo de gráfica presentan estarán disponibles dependiendo de las características de los datos que hayamos ingresado (Figura 15).

Figura 15. Impresión de pantalla de las selecciones de la opción graficas personalizables.

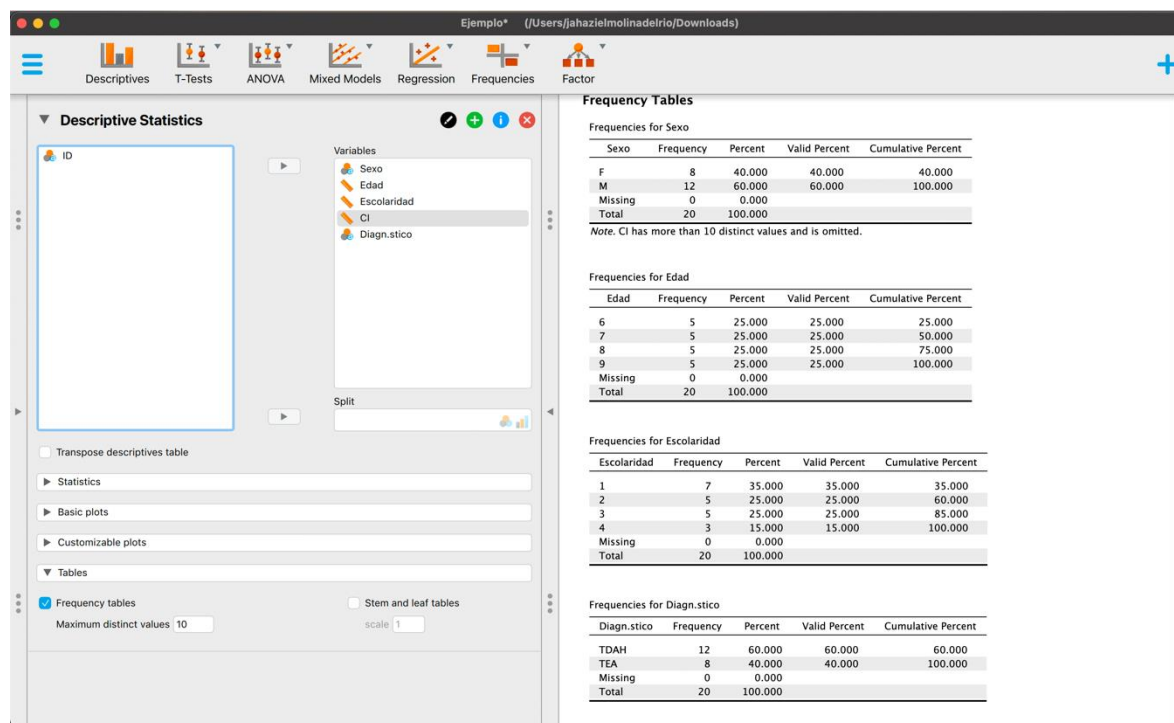


Al tratarse de datos descriptivos y haber ingresado todas las variables a la sección correspondiente para su análisis, los datos y **graficas** se realizan por cada una de manera independiente, como se ha visto en las imágenes anteriores, en caso de querer observar alguna distribución comparada por alguna otra variable de tipo nominal y ordinal, esta debe incluirse en el recuadro de “Split” como se señaló anteriormente, de esta forma el programa colocará en una misma gráfica los elementos correspondientes a las etiquetas que tengamos de dicha variable, como puede apreciarse a continuación en los boxplots (Figura 16).

Figura 16. Impresión de pantalla de los resultados al incluir una variable de separación.



La última pestaña denominada tablas, presenta la opción de las tablas de frecuencia, al igual que los análisis anteriores, estas se presentarán por cada una de las variables colocadas en el recuadro de análisis, incluyendo la frecuencia absoluta, el porcentaje de frecuencia, el porcentaje valido y el porcentaje acumulado. Si quisiéramos comparar la distribución dividida por alguna otra variable ordinal o nominal, al incluirla en el recuadro de “Split” las tablas realizaran la distribución considerando los diferentes niveles de la variable incluida (Figura 17).

Figura 17. Impresión de pantalla de la selección de la opción de tablas de frecuencia.

De esta forma podemos obtener diferentes análisis y gráficas de los datos que complementen la manera en cómo presentamos y divulgamos nuestra información. Cabe mencionar algunos aspectos importantes, en primera instancia este tipo de software es una herramienta que facilita el análisis y elaboración de gráficas, pero que no sustituye el conocimiento que se debe tener respecto a las nociones estadísticas que le permiten elegir las pruebas necesarias para su análisis, así como el sustento teórico para la interpretación de los resultados.

Finalmente, como se mencionó anteriormente el software presentado ofrece la posibilidad de diversos análisis no solo de estadística descriptiva, sino inferencial, con opciones de análisis de estadística tanto clásica como

Bayesiana, por lo que invitamos a explorar esta herramienta poniendo en práctica los diferentes análisis que se detallan en el presente libro.

Conclusiones

El presente capítulo pretende ser una guía tanto de estudiantes como de profesionales de la educación especial con el propósito de orientar su trabajo, en el abordaje y presentación de resultados de problemas de investigación. Primeramente, se explicó de manera detallada como realizar una tabla distribución de frecuencias y específicamente, se ejemplificó como aplicarla en la realización de experimentos en torno a la educación especial. Este tipo de presentación de datos en la educación especial, puede ser muy útil ya que podemos comparar los puntajes de diferentes pruebas en niños con características específicas, conocer su frecuencia, y poder comprarlos con datos de una media poblacional, así mismo, la presentación de los datos en histogramas, polígonos de frecuencia, gráficas de barras, gráficas de pastel o los diagramas de caja y bigotes de Tukey, tienen un amplio uso para observar mejor la distribución de frecuencias o medias de los datos analizados, al hacer esto, es mucho más amigable y fácil tener una noción de interpretación de los datos.

La finalidad principal del texto fue aportar estrategias y o una metodología para presentar de manera clara resultados cuantificables de diversos experimentos o investigaciones en el ámbito de la educación especial. Es importante destacar que la claridad de dicha presentación es de vital importancia para la comprensión de los resultados y la interpretación de estos, además de realizar un análisis estadístico adecuado.

Finalmente, el acceso actual a las tecnologías de la información y a la comunicación digital, han dotado a los investigadores de diversas herramientas de programas para graficar, además de los softwares tradicionales como Excel, o los de paga como SigmaPlot, SPSS, también se tiene la opción de trabajar con programas de acceso libre como R Studio, JAMOVI y JASP, los cuales nos permiten realizar análisis estadísticos no sólo de estadística descriptiva, sino inferencial, y poder graficar los datos de una manera rápida y eficiente, aunque como todo es necesario perder el miedo de usar nuevos programas y usarlos constantemente para que estos sean usados adecuadamente o conocer los atajos de los mismos. Aquí presentamos una pequeña guía del uso de JASP para estas gráficas presentadas aquí.

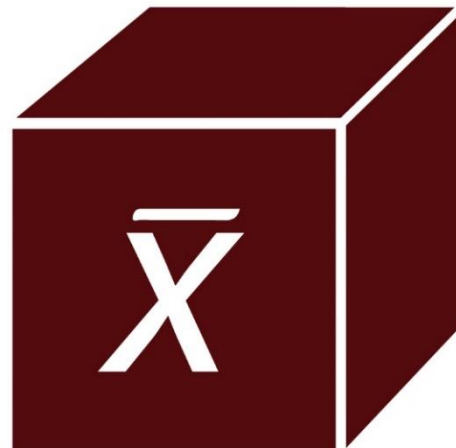
Referencias bibliográficas

- Benjamini, Y. (1988). Opening the box of a boxplot. *The American Statistician*, 42(4), 257-262.
- Carrascosa Duro, M. M. (2016). Datos en tablas y gráficas estadísticas. Tendenciosidad y errores más comunes. *Publicaciones Didácticas*, 72(1), 71-78.
- Duncan, A. W., y Bishop, S. L. (2013). Understanding the gap between cognitive abilities and daily living skills in adolescents with autism spectrum disorders with average intelligence. *Autism*, 19(1), 64–72. <https://doi:10.1177/1362361313510068>
- Espinosa Garzón, A. M. (2017). *Estadística para las ciencias sociales y humanas I*. Bogotá D.C., Fundación Universitaria del Área Andina.
- Hong, J.S., Singh, V., Kalb, L., Ashkar, A. y Landa, R. (2021). Replication study of ADOS-2 toddler module cut-off scores for autism spectrum disorder classification. *Autism Research*, 2021, e02496. <https://doi.org/10.1002/aur.2496>
- JASP Team (2022). JASP (Version 0.16.3)[Computer software].

- Liu, Y. (2008). Box plots: use and interpretation. *Transfusion* (Philadelphia, PA), 48(11), 2279-2280.
- Mejía Vélez, S. (2016). *Apuntes de estadística con aplicaciones a la psicología*. Colección CB Ciencias Básicas. Universidad Católica de Colombia: Bogotá.
- Pagano, R.R. (2011). *Estadística para las ciencias del comportamiento*. Novena edición, Thomson.
- Pértega Díaz, S., y Pita Fernández, S. (2001). Representación gráfica en el análisis de datos. *Cad Aten Primaria*, 8, 112-117.
- Ritchev, F. (2006). *Estadística para las ciencias sociales*. México: McGraw Hill.
- Schweitzer, J. B., y Sulzer-Azaroff, B. (1995). Self-control in boys with attention deficit hyperactivity disorder: Effects of added stimulation and time. *Journal of Child Psychology and Psychiatry*, 36(4), 671-686.
- Young, P. V., y Schmid, C. F. (1951). Las Técnicas de la Investigación Social. Capítulo XII: Representación Gráfica. *Revista Mexicana de Sociología*, 127-165.

Parte III:

ESTADÍSTICA INFERENCIAL



EL USO DE LA CORRELACIÓN

Jorge Carlos Hevia Orozco [‡]

Cid Ramón González González

Resumen:

La correlación es una herramienta que se ha usado en la investigación científica para establecer el grado de relación que tienen dos o más variables. Sin bien, a través de la correlación no se establece una relación causal, si nos ayuda a establecer que por distintas razones, dos variables dependientes o independientes, guardan una relación que pudiera ayudarnos a explicar el comportamiento de tales variables y aproximarnos a entender un fenómeno de cualquier índole, ya sea social, fisiológico, físico, biológico, etc. Además, con la correlación estadística se puede establecer el grado de relación que existen entre las variables, puesto que este índice va desde el +1 (correlación perfecta) hasta el -1 (anticorrelación perfecta), pasando por el 0 (ninguna correlación). Como se podrá ver en el capítulo, la correlación estadística puede ser el objetivo principal de una investigación o puede ser parte de una serie de preguntas de investigación de un proyecto científico. Este capítulo ayudará al educador especial a comprender cómo se puede usar la correlación en el campo educativo que practica y beneficiarse de esta herramienta tan útil.

Palabras clave: correlación estadística, correlación Pearson, correlación Spearman, regresión lineal

[‡] Universidad Anáhuac Mayab, Escuela de Psicología, Mérida, Yucatán, México. Centro Educativo New Hope, Mérida, Yucatán, México. jorge.hevia@anahuac.mx

Introducción

La correlación es una herramienta muy reveladora para explicar fenómenos de la vida diaria de dos variables que se relacionan entre sí. El primero en interesarse fue Adolphe-Jacques Quetelet (1796-1874), un matemático belga que dedicó gran parte de su vida a estudiar medidas antropométricas del hombre medio y su dependencia con distintas variables independientes tales como el sexo, edad, profesión o nivel de educación. Esto lo hacía estableciendo las medias y desviaciones típicas de medidas antropométricas y relacionando variables como la edad y la altura en personas de 0 a 30 años. Sin embargo, la correlación tomó un gran impulso al ser retomada por Francis Galton (1822-1911), quien estudió la variación de dos medias o promedios. Dado su parentesco con Charles Darwin, Galton se interesó por sus estudios sobre la herencia y estableció la variabilidad o la relación de los patrones físicos entre padres e hijos. No fue sino hasta 1895, cuando Karl Pearson presentó la fórmula para evaluar el coeficiente de correlación, medida mediante la cual se establece la asociación entre dos variables, en otras palabras, para evaluar si los valores de una variable cambian consistentemente conforme a los de la otra, o si, por el contrario, no hay ningún patrón que las asocie entre ellas (Mondragón, 2014).

Es importante mencionar que la correlación ha sido utilizada por distintos campos de la ciencia como la economía, la psicología y diversas ciencias naturales, entre otras. La utilidad de la correlación radica en su capacidad para conocer el comportamiento de una variable, al conocer el comportamiento de otra variable vinculada. Es decir, cuando los puntajes de un grupo de estudio pueden predecirse a partir de su asociación con los puntajes de otro grupo de individuos. Un ejemplo de esto puede ser que “a mayor tiempo de ejercicio, mayor eficacia al ejecutar una tarea” o “a menor tiempo de ejercicio, menor

eficacia al ejecutar una tarea”. Esta correlación, que está claramente expresada en un sentido positivo, también puede ser expresada en una correlación negativa como por ejemplo “a mayor tiempo requerido para terminar un ejercicio, menor eficacia al ejecutar una tarea” o “a menor tiempo requerido para la resolución de un ejercicio, mayor eficacia al ejecutar una tarea”.

Dado que es común que los estudiantes se confundan en la diferencia entre correlación y causalidad, los autores creen importante hacer una breve aclaración de este punto. Una correlación significativa entre dos variables no implica que haya una causalidad entre ellas. Tomemos de ejemplo dos trabajos de investigación.

Por ejemplo, Al-Yagón (2012), aplicó pruebas para evaluar el ajuste socioemocional y actitudes de apego en pacientes con problemas de aprendizaje (PA). Al-Yagón aplicó pruebas que medían apego seguro con la figura paterna y la figura materna y con los maestros, los sentimientos de soledad, afectos positivos y negativos personales y conductas de externalización y de internalización. En este caso, la correlación de Pearson se aplicó dentro de un modelo llamado modelo de ecuación estructural (structural equation model o SEM, por sus siglas en inglés). Este modelo tiene el objetivo de usar la correlación de Pearson de manera simultánea para los dos grupos. Entre los resultados más importantes encontraron que los adolescentes con PA que se sentían más apegados a la madre, presentaron una correlación negativa con sentimientos negativos, positiva con sentimientos positivos, negativa con sentimientos de soledad con respecto a los amigos, negativo con problemas de internalización y negativo con problemas de externalización de la conducta. Con respecto al apego con el padre, en este grupo se encontró sólo una correlación positiva con los sentimientos positivos. Por otro lado, en el grupo

de adolescentes sin PA, se encontró a través de la correlación, que el apego con el padre estaba relacionado de manera negativa con los sentimientos negativos, de manera positiva con los sentimientos positivos, de manera negativa con los sentimientos de soledad con respecto a un compañero y con respecto al grupo. A través de la correlación, estos resultados nos permiten entender el factor relacionado (apego con las madres y apego con los padres), con algunas conductas de los adolescentes con desarrollo normal, sin embargo, vemos como no hay el mismo patrón de correlaciones en el grupo con PA, con respecto al apego con el padre. Esto es, en el grupo con PA, el apego con la figura paterna no guardó relación con otras variables como con los sentimientos de soledad o los estados afectivos, como sí sucedió en el grupo con desarrollo típico, en donde el apego con la figura del padre, se relacionó de manera negativa con situaciones negativas y de manera positiva con situaciones positivas para la vida del adolescente.

Considerando lo anterior, es lógico pensar que el apego con la figura paterna o con la figura materna es la causa de los sentimientos negativos o positivos de los pacientes o con los sentimientos de soledad de estos. Sin embargo, hay muchos factores que pudieran estar involucrados de manera simultánea como otras comorbilidades, la alimentación, el sueño, otros hábitos diarios, problemas familiares, etc. De tal forma que no podremos hablar de una causalidad directa entre variables y mucho menos entre conductas.

Otro ejemplo es el estudio llevado a cabo por García-Villamizar (2017). El autor buscaba encontrar la relación entre la percepción del dolor en pacientes con autismo y algunas variables tales como la psicopatología comórbida (ansiedad o depresión), las disfunciones ejecutivas y el nivel de estrés. En los resultados se encontró que la percepción del dolor se encontraba relacionada de

manera positiva con la sintomatología del autismo, las comorbilidades asociadas al trastorno autista, y con la disfunción ejecutiva autista en la vida diaria. La única variable que no demostró ninguna relación con la percepción del dolor, fue el nivel de estrés. De esta forma se puede entender que, en el grupo de pacientes con autismo, el estar estresado, no es un factor que favorezca la percepción del dolor, pero sí, el hecho de que el paciente presente alguna comorbilidad y también alguna disfunción ejecutiva.

De la misma forma, no es lógico pensar que la percepción del dolor es causante de la sintomatología del autismo, puesto que no tiene coherencia lógica esta asociación. Por otro lado, se puede pensar que el sentido inverso de la aseveración puede tener cierto tinte de relación: la sintomatología del autismo se relaciona con la percepción del dolor. Pero también en este caso, la causalidad no sería lógica, puesto que la manifestación clínica de un trastorno, no puede ser causante de un fenómeno tan complejo como la percepción del dolor.

Como pudimos comprender, un error común para todo investigador en sus primeros pasos es confundir la correlación con la causalidad. El hecho de que dos variables dependientes se encuentren asociadas no significa que una cause la aparición de la otra. Muchas veces esta equivocación se da al intentar explicar un fenómeno que es de gran interés, sin embargo, el alcance de la correlación no es tal que permita obtener esta conclusión. Ciertamente, la correlación permite ver si dos variables están relacionadas y también el grado de asociación, pero esto no representa una causalidad entre dos o más variables. Inclusive puede darse el caso de encontrar relaciones entre puntajes y que esto no tenga un significado lógico, por ejemplo, encontrar una correlación entre la estatura de un grupo de adolescentes y su coeficiente intelectual. Muchas veces,

esto se da por la gran cantidad de datos analizados y el subsecuente aumento de la probabilidad automática. Este es un detalle que debe de tomarse en cuenta al llevar a cabo un análisis guiado por los datos, puesto que en estos análisis la gran cantidad de datos es la base para su realización.

Es importante mencionar que, en el desarrollo de un proyecto, el análisis correlacional puede hacerse en conjunto con otro tipo de análisis como por ejemplo, los explicativos (experimentales y no experimentales). Existen cuatro tipos de alcances de una investigación: el exploratorio, el descriptivo, el correlacional y el explicativo. Cada uno de ellos tiene su propio propósito, su propio valor y también sus propias limitaciones (Hernández - Sampieri, 2014). En la práctica, cada una de ellas tiene sus propias metodologías, pero también en un mismo trabajo, pueden coexistir cualquiera de las cuatro, ya sea en diferentes fases del proyecto, o también, dando diferentes tratamientos a los datos. Es decir, un proyecto para determinar las causas y efectos a través de la manipulación de las variables independientes, puede complementarse con un análisis correlacional de tal forma que las conclusiones del proyecto sean cada vez más profundas y la explicación del fenómeno estudiado sea más completa. De la misma forma, un estudio puede iniciar como descriptivo, recopilando distintas características de un grupo en términos cuantificables numéricos y posteriormente, pueden establecerse las correlaciones estadísticas entre distintas de las variables recogidas (siempre y cuando tengan un valor y una explicación lógica). Por último, un estudio que inicia con hipótesis correlacionales puede ser complementado con la aplicación de una metodología explicativa (experimental o no experimental), al obtener resultados interesantes que requieren una comparación de variables independientes para un mayor

entendimiento. Todo dependerá del interés del investigador, de la capacidad logística, de los materiales y de los recursos económicos para hacerlo.

La herramienta de la correlación sirve para comprobar hipótesis correlacionales. Es decir, aquellas hipótesis que especulan acerca de la asociación entre dos variables dependientes. Por ejemplo, “el alcoholismo está asociado con la presencia de padecimientos hepáticos”, o “el clima laboral está asociado con la satisfacción laboral”, o “la exposición al asbesto está asociada con la incidencia de cáncer pulmonar”. Establecen no sólo que hay una relación, sino también cómo están asociadas. En el contexto de la educación especial, las hipótesis correlacionales responderían suposiciones tales como “mientras más temprano se dé un diagnóstico de TDAH, menos repercusiones académicas se presentan”, “a mejor dinámica familiar, mejores puntajes de desarrollo psicomotor en pacientes con parálisis cerebral”. En el caso de los estudios correlacionales se busca establecer la asociación entre variables mediante un patrón predecible para un grupo o población (Hernández - Sampieri, 2014).

Coefficiente de correlación de Pearson

El coeficiente de correlación de Pearson es una prueba estadística que es utilizada para analizar la relación entre dos variables las cuales son de intervalo o de razón. Se simboliza con la letra r . Como ya se ha mencionado anteriormente, esta prueba no establece causalidad entre variables, sino que expresa el grado de asociación que existen entre las variables dependientes. La hipótesis que busca establecer es de tipo correlacional en donde los altos puntajes de una variable “x” están asociados con altos puntajes de una prueba “y” o los altos puntajes de una variable “x” están asociados con bajos puntajes

de una prueba y”. Así el coeficiente de correlación puede variar de +1 a -1. La interpretación de los valores en este rango aparece en la tabla 1.

Tabla 1. Valor del coeficiente de correlación y su interpretación empírica.

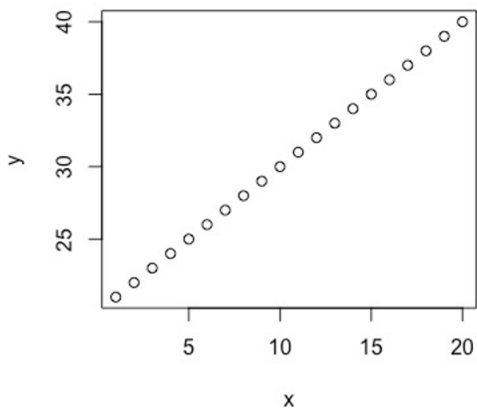
Valor	Interpretación
-1.00	<i>Correlación negativa perfecta.</i> ¹
-0.90	Muy fuerte correlación negativa.
-0.75	Considerable correlación negativa.
-0.50	Correlación negativa media.
-0.25	Débil correlación negativa.
-0.10	Muy débil correlación negativa.
0.00	No existe ninguna correlación entre las variables dependientes.
+0.10	Muy débil correlación positiva.
+0.25	Débil correlación positiva.
+0.50	Correlación positiva media
+0.75	Correlación positiva considerable.
+0.90	Muy fuerte correlación positiva.
+1.00	<i>Correlación positiva perfecta.</i> ²

Si consideramos a dos variables como variable “x” y variable “y”, podemos afirmar que:

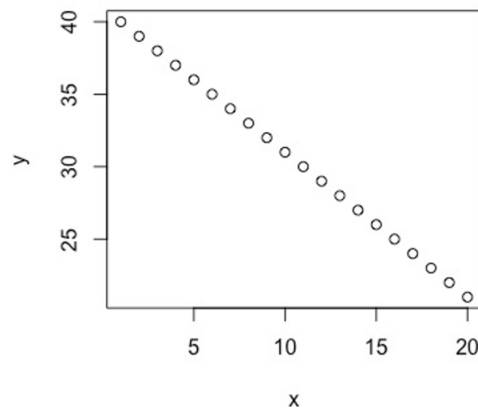
- 1) “A mayores puntajes obtenidos de una prueba X , estos pudieran asociarse a menor puntajes obtenidos de una prueba Y ”, de manera proporcional. Es decir, cada vez que los puntajes de una prueba X aumentan en una unidad, los puntajes obtenidos de una prueba Y , disminuyen siempre una cantidad constante (correlación positiva). Esto también se aplica “a menor X , mayor Y ” (correlación negativa).
- 2) “A mayores puntajes de X , mayores puntajes de Y ” o “a menores puntajes de X , menores puntajes de Y ”, de manera proporcional. Cada vez que X aumenta, Y aumenta siempre una cantidad constante.

Las relaciones podrían expresarse gráficamente como se muestra en la figura 1.

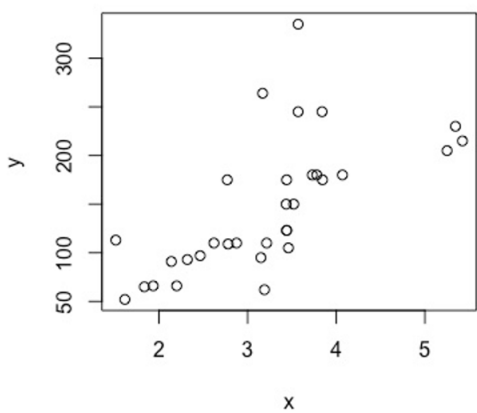
Figura 1. Diagramas de dispersión indicando distintas correlaciones.



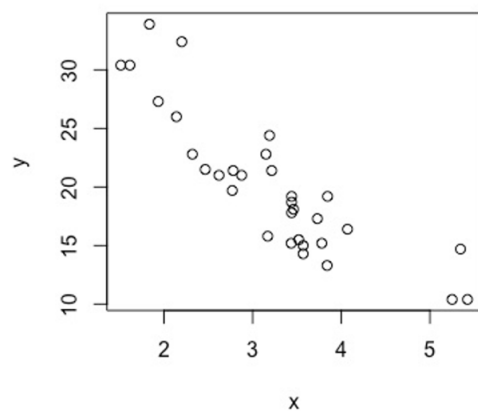
a) Correlación positiva perfecta entre x y y



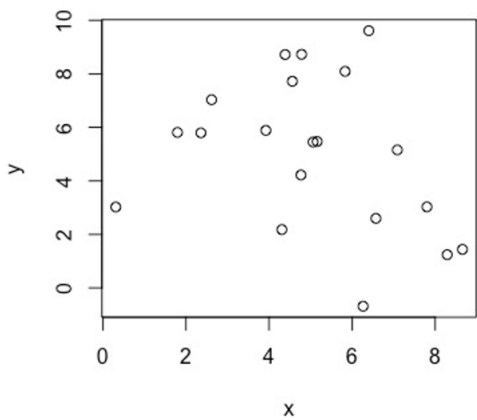
b) Correlación negativa perfecta entre x y y



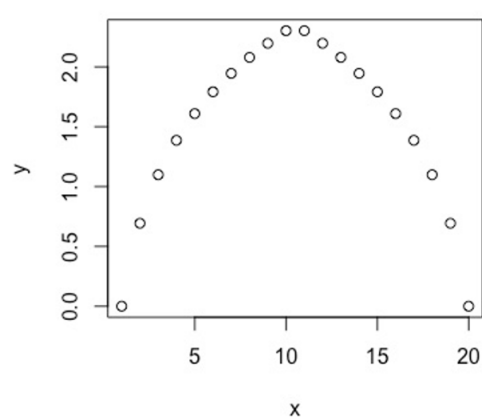
c) Correlación positiva media entre x y y



d) Correlación negativa media entre x y y



e) Sin correlación entre x y y



f) Correlación no lineal entre x y y

Ecuación para calcular el coeficiente de correlación

La ecuación para calcular el coeficiente de correlación r esta expresado en la ecuación 1.

$$r = \frac{\Sigma(x - \underline{x})(y - \underline{y})}{\sqrt{\Sigma(x - \bar{x})^2 \cdot \Sigma(y - \bar{y})^2}}$$

Donde:

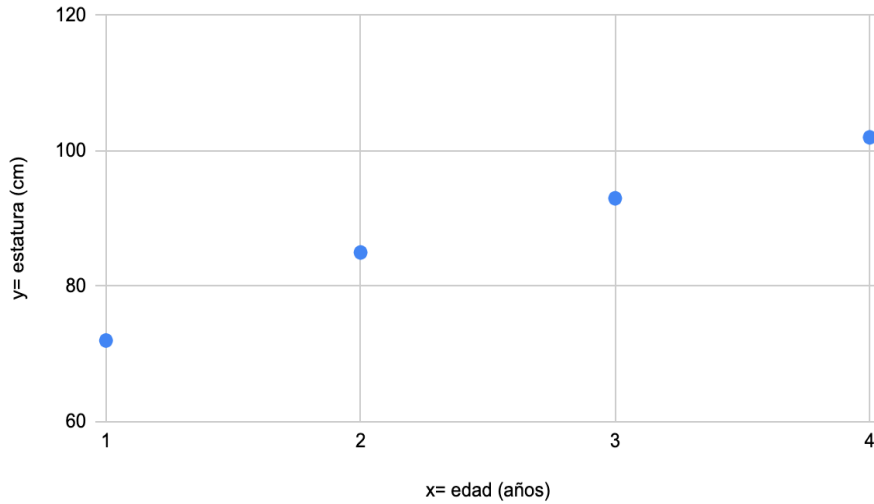
- \bar{x} es la media de los valores de x
- \bar{y} es la media de los valores de y.

Como ejemplo vamos a hallar la correlación que existe entre la edad de un niño varón y su estatura durante los primeros 4 años en un crecimiento promedio normal (Tabla 2). Aquí tendremos dos variables, donde x es la edad y y la estatura en cm.

Tabla 2. Valores de la edad en años y estatura correspondiente de un niño varón con crecimiento normotípico (Centers for Disease Control and Prevention, National Center for Health Statistics, 2017).

x= edad (años)	1	2	3	4
y= estatura (cm)	72	85	93	102

Haciendo una vista previa de la correlación de los datos en una gráfica de x,y tenemos:

Figura 2. Gráfica x,y de la correlación entre estatura y edad de niños (CDC, 2017).

Aplicando la ecuación, lo primero que tenemos que hallar son los promedios para la edad (x) y la estatura (y).

$$\bar{x} = \frac{1 + 2 + 3 + 4}{4} = 2.5$$

$$\bar{y} = \frac{72 + 85 + 93 + 102}{4} = 88$$

Segundo, se calcula la distancia de valor de x y y con respecto de la media (Tabla 3):

Tabla 3. Valores de la diferencia entre las variables x (edad) y y (estatura) respecto a su media correspondiente.

Edad(x)	Estatura(y)	$x_i - \bar{x}$	$y_i - \bar{y}$
1	72	$1 - 2.5 = -1.5$	$72 - 88 = -16$
2	85	$2 - 2.5 = -0.5$	$85 - 88 = -3$
3	93	$3 - 2.5 = 0.5$	$93 - 88 = 5$
4	102	$4 - 2.5 = 1.5$	$102 - 88 = 14$

Tercero, resolvemos el numerador de la ecuación 1 multiplicando los valores correspondientes desde el primero hasta el último valor de las diferencias y se suman los valores de acuerdo con la ecuación:

$$\Sigma[(x_i - \bar{x})(y_i - \bar{y})]$$

$$[(-1.5)(-16)] + [(-0.5)(-3)] + [(0.5)(5)] + [(1.5)(14)] = 24 + 1.5 + 2.5 + 21 = 49$$

Cuarto, resolvemos el denominador de la ecuación:

$$\sqrt{\Sigma(x - \bar{x})^2 \cdot \Sigma(y - \bar{y})^2}$$

Resolviendo los cuadrados de las distancias por separado:

$$\Sigma(x - \bar{x})^2 = -1.5^2 + (-.5)^2 + 0.5^2 + 1.5^2 = 2.25 + 0.25 + 0.25 + 2.25 = 5$$

$$\Sigma(y - \bar{y})^2 = -16^2 + (-3)^2 + 5^2 + 14^2 = 256 + 9 + 25 + 196 = 486$$

Multiplicamos el resultado de las dos expresiones:

$$5 \times 486 = 2430$$

Al aplicar la raíz cuadrada al resultado:

$$\sqrt{2430} = 49.29$$

Finalmente resolvemos la ecuación al dividir el número sobre el denominador:

$$r = \frac{49}{49.29} = 0.994$$

Esto significa una correlación casi perfecta, muy cercana a 1, donde por cada año aumenta una parte proporcional de estatura de crecimiento en

centímetros. Esto es de esperar en los primeros años de crecimiento. Sin embargo, llegará una edad en la que existirá una desaceleración en el crecimiento y esta correlación irá disminuyendo hasta que deje de haber crecimiento y en ese momento esa correlación se convertirá en casi cero.

Cabe mencionar que la correlación estadística se puede establecer entre distintos tipos de datos. Un ejemplo de esto es el realizado por Hevia-Orozco y Sanz-Martín (2018), en donde se establecieron las correlaciones entre variables electroencefalográficas y variables conductuales provenientes de la ejecución de un test. Ellos establecieron la correlación entre la actividad eléctrica cerebral de la corteza prefrontal durante un estado de reposo y los puntajes derivados de pruebas que evaluaban niveles de ansiedad, depresión y trastorno de estrés postraumático (TEPT). Esto lo llevaron a cabo en adolescentes que habían vivido durante más del 80% de su vida en orfanatorios. Los investigadores habían usado la correlación electroencefalográfica para medir la “comunicación” entre áreas cerebrales y esos mismos índices fueron usados para asociarlos con puntajes conductuales. De esta forma se podría ver si los puntajes de ansiedad, depresión o TEPT se asociaban con la comunicación anormal entre dos áreas cerebrales. Entre varios resultados, encontraron que distintos síntomas de estrés postraumático como la reexperimentación ($r=0.625$), el alertamiento ($r=0.540$) y la evitación ($r=0.599$) están correlacionadas con la comunicación entre regiones frontopolares (las regiones más anteriores de la corteza frontal) derechas e izquierdas. También, que distintos síntomas de ansiedad tales como pánico agorafóbico ($r=0.482$) y ansiedad por separación ($r=0.452$), se correlacionaron con la comunicación prefronto - parietal izquierda. Por último, encontraron que el síntoma depresivo de disforia se correlacionó con la comunicación intra - prefrontal derecha e

izquierda en bandas electroencefalográficas desde delta ($r=0.520$) hasta alfa 1 ($r=0.540$). Como podemos observar, las correlaciones encontradas son interpretadas como una correlación positiva media y una correlación positiva débil, acorde a la tabla mencionada más arriba. Estos resultados permitieron a los autores, obtener un entendimiento más claro acerca del potencial desarrollo anormal de distintas regiones prefrontales y su relación con algunas conductas patológicas.

El valor P y la comprobación de hipótesis

Otro valor importante que tenemos que considerar al interpretar una correlación es el valor p que aporta evidencia estadística de que el coeficiente de correlación es diferente de cero de acuerdo con las observaciones realizadas. El valor p es una medida que nos indica la probabilidad de que la hipótesis nula sea verdadera. Esto es, cuando no existe ningún efecto. Una correlación puede resultar como significativa (p.ej., $p=0.005$), pero con una correlación débil (p.e., $r= -0.16$). O bien, con una correlación fuerte (p.e = 0.89), pero no significativa (p.e., $p = 0.34$).

Este valor p se basa en un sistema de teorías de probabilidades de distribución normal y pueden consultarse en una tabla con base en el número de grados de libertad disponibles del grupo de datos que se está analizando y el nivel de confianza con la que se pretende trabajar. Este nivel de confianza es arbitrario y regularmente se utiliza el 95%. Esto es equivalente a un valor alfa (α) = 0.05 en las tablas de probabilidad. Sin embargo, pueden utilizarse otros valores tales como 0.1 , o aún menores como 0.01 . Esto dependerá del tipo de evaluaciones llevadas a cabo. Esto quiere decir que si utilizamos un valor de

referencia $p = 0.05$ como significancia estadística, entonces un valor p menor de 0.05 se consideró estadísticamente significativo. Por el contrario, un valor mayor a 0.05, entonces no existirá suficiente evidencia para rechazar la hipótesis nula y considerará como diferencia no significativa.

Correlaciones no paramétricas

Charles Edward Spearman (Londres, 1863-1945) desarrolló un coeficiente de correlación que permite correlacionar variables dependientes, considerando un orden jerárquico que requiere el ordenamiento de los datos, conforme a su rango, en lugar de medir cada uno de los datos sin considerar la interrelación entre cada uno de ellos.

Primero, hay que recordar que los ítems de una prueba pueden tener distintos niveles de medición. Es decir, por un lado, hay encuestas en las que una pregunta exija una respuesta que puede ser contestada con un “Sí”, un “No” o con un nombre en particular tal como “Católico”, “Judío”, “Protestante”, o “Local” o “Foráneo” o, por otro lado, hay pruebas que exigen una respuesta como el tiempo de reacción en milisegundos. Evidentemente, la forma de procesar para cuantificar y medir los resultados de esas respuestas será diferente para cada uno. Es en ese momento cuando tenemos que entender que la correlación puede aplicarse a distintos tipos de datos.

Retomaremos brevemente algunos conceptos de los niveles de medición de las variables. Por ejemplo, hay variables que son nominales, es decir, que son solamente categorías independientes, que, por las características de las respuestas, estas no tienen ningún orden jerárquico de mayor o menor. No se pueden procesar de manera aritmética. Este nivel de medición es el nominal

(porque sólo procesa nombres particulares). Hay variables que pertenecen a la escala ordinal, en los que cada categoría tiene asignado un número, el cual forma parte de una escala jerárquica que permite organizar de menor a mayor tales categorías. Por ejemplo, el orden del final de una carrera de Fórmula 1:

1- Max Verstappen

2- Charles Leclerc

3- Sergio Pérez

4- Lewis Hamilton

En el ejemplo, los nombres son organizados conforme el número que fueron cruzando la meta en el último lapso, el cual les confiere un orden jerárquico.

La siguiente escala de medición es la de intervalos, en los cuales, las distancias entre categorías son igual a lo largo de una escala, además de que los datos tienen orden y jerarquía. Por ejemplo, si en una prueba de baile Alma obtuvo un puntaje de 35, Gabriela un puntaje de 25 y Alejandra un puntaje de 15, estos datos tienen un orden jerárquico pero la distancia se mantiene constante. Es decir, la distancia entre Alma y Gabriela es la misma que la de Gabriela y Alejandra. Además, como podemos recordar, el 0 es arbitrario, es decir la escala se construye a partir de la asignación arbitraria de 0 a una categoría. Es importante mencionar esto puesto que se puede establecer la correlación entre variables de tipo nominal y ordinal a través de una correlación llamada correlación de Spearman.

Cabe aclarar que algunas mediciones de la conducta humana se acercan en su definición a la escala de medición de intervalo, aunque no lo son verdaderamente (como por ejemplo una escala de inteligencia o una escala de

satisfacción laboral), pero se pueden tratar como si fuera de intervalo. Este nivel de medición permite utilizar mediciones aritméticas básicas y avanzadas que de otra manera no se podrían utilizar. Aunque hay muchos investigadores, principalmente del área de las ciencias exactas, que no estarían de acuerdo en procesar, por ejemplo, una escala de Likert como variable de intervalo.

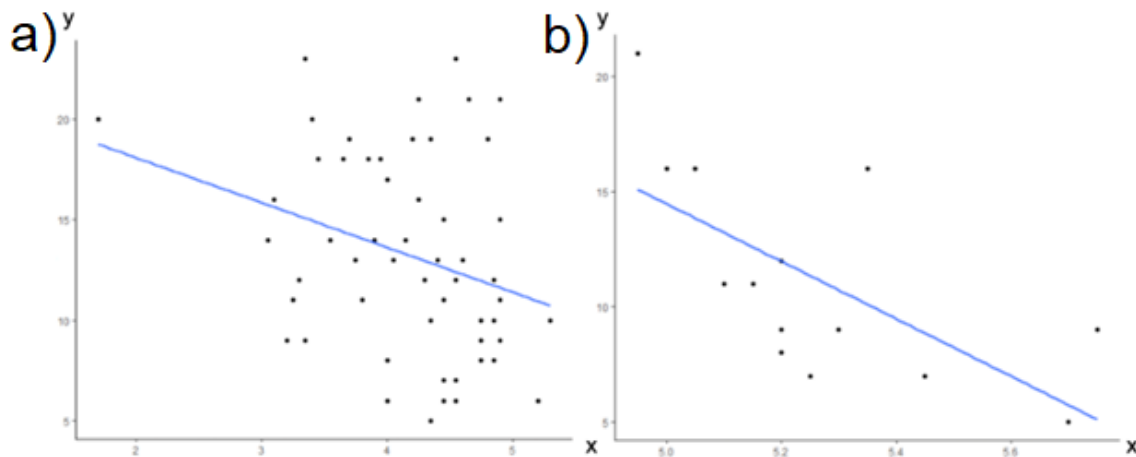
La correlación de Spearman es un coeficiente de correlación para variables en un nivel de medición ordinal, es decir, cuando para esto, es necesario primero la ordenación de manera jerárquica de los valores de calificación. Por ejemplo, supongamos que se le pide a un grupo de personas que califique del 1 al 10, la “preferencia en el olor” y la “atractivo de la botella” de 10 perfumes, puesto que queremos entender de manera detallada las preferencias de los compradores. Para establecer la relación entre estas dos variables resultará necesario hacer primero la organización jerárquica de los puntajes de la preferencia puesto que es un dato imprescindible para entender el concepto que estamos tramando investigar. Al establecer una correlación de Spearman, obtendremos una relación que va de +1 (correlación positiva perfecta, a mayor preferencia, mejor percepción de atractivo de la botella), pasando por el 0 (sin relación entre la preferencia y la presentación), hasta llegar al -1 (correlación negativa perfecta, a menor preferencia, mejor percepción de atractivo de la botella).

Otra característica importante de la correlación de Spearman es que es un estadístico no paramétrico de variables. Es decir, se usa cuando los datos de los resultados presentan una distribución no normal. En otras palabras, no exige la suposición de la normalidad de la muestra. Podemos comprenderlo más amigablemente cuando observamos que en la distribución de los datos en una

gráfica, éstos no se agrupan tendenciosamente hacia el centro de la campana de Gauss y pueden estar sesgados hacia la derecha o hacia la izquierda.

El uso de la correlación de Spearman en la psicología experimental es altamente útil para profundizar en el análisis de la conducta humana. En un estudio reciente, Pless-O'Connor y Hevia-Orozco (en preparación), usaron esta herramienta para ver la relación que hay entre los distintos roles de género y la flexibilidad cognitiva. Esta relación se ha postulado bajo la premisa de que las personas que se pueden identificar con roles de género opuestos, es decir hombres con roles femeninos (hombres femeninos) y mujeres con roles masculinos (mujeres masculinas), tienen mayor flexibilidad cognitiva en comparación a los hombres con roles masculinos (hombres masculinos) y mujeres con roles femeninos (mujeres femeninas). Otra de las opciones de roles conductuales es el indiferenciado (hombres o mujeres indiferenciadas). Con respecto en los resultados vinculados a la flexibilidad cognitiva, se encontró una correlación significativa negativa débil entre puntajes de masculinidad y perseveraciones en la prueba de Wisconsin de clasificación de cartas en mujeres indiferenciadas ($p = 0.03311$, $\rho = -0.2878$) y negativa moderada entre las mismas variables en mujeres masculinas ($p = 0.005$, $\rho = -0.6824$) (Figura 3).

Figura 3. Correlación lineal entre x = masculinidad y y =flexibilidad cognitiva en: a) mujeres indiferenciadas, b) mujeres masculinas.



Otra medida de correlación no paramétrica fue propuesta por Maurice George Kendall (1938). Esta correlación ha sido probada más robusta y eficiente que Spearman cuando existen pocos datos extremos o atípicos que puedan afectar la correlación (Croux y Dehon, 2010). La correlación de Kendall utiliza un algoritmo diferente a la correlación de Spearman y puede resolverse mediante programas de estadística. Sin embargo, la interpretación obedece a los mismos parámetros del coeficiente de correlación y el valor p para la comprobación de hipótesis ya explicadas arriba.

Regresión lineal

Pero ¿será posible correlacionar variables dependientes e independientes? Es posible gracias a un método conocido como regresión. Para hacerlo, debe tenerse un sólido sustento teórico. Este método utiliza la correlación de Pearson y sirve para explicar de forma lineal la relación entre dos variables, una independiente y otra dependiente. Las dependientes son aquellas variables que

van a cambiar en función de las variables independientes. Por ejemplo, se realiza una pregunta de investigación sobre el número de repeticiones de un ejercicio matemático y el nivel de retención del estudiante. Se establece la hipótesis de que entre mayor número de repeticiones mayor será el nivel de retención del estudiante y que este podrá resolver un problema similar con la misma estrategia, pero con distintos valores durante un examen. Aquí el número de repeticiones es la variable independiente, y la puntuación del examen es la variable dependiente que está en función del número de repeticiones.

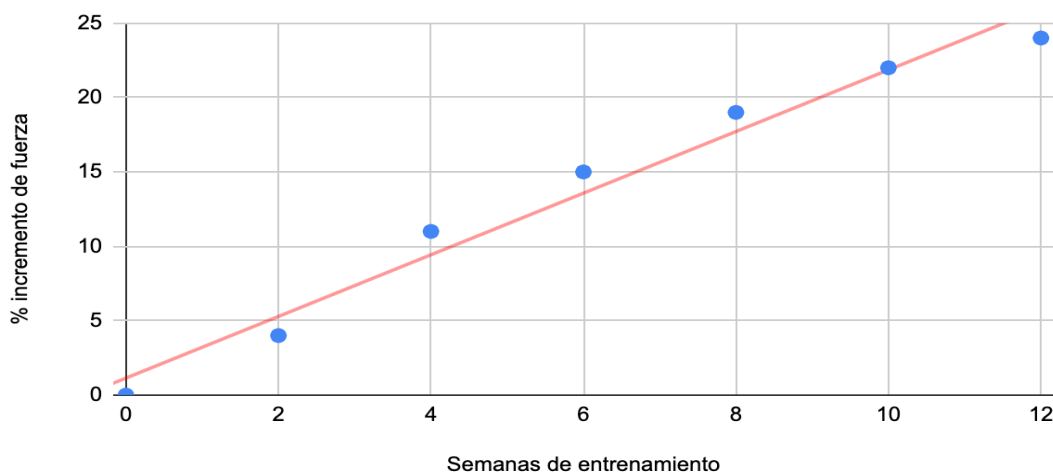
Hay que recordar que, al hacer una regresión, buscamos explicar un fenómeno al ver qué tanto nuestros datos se asemejan o se ajustan a un modelo lineal. Es decir, qué tanto mi nube de puntos en un diagrama de dispersión se ajusta a un modelo o forma ideal de explicar un fenómeno. Por ejemplo, supongamos que somos entrenadores de levantamiento de pesas y queremos explicar la relación que existe entre una variable dependiente Y (% de incremento de fuerza en levantamiento de barra) y un conjunto de variables independientes X_1, \dots, X_n (semanas de entrenamiento). Los valores se describen en la tabla 4:

Tabla 4. Valores de porcentaje de aumento de fuerza en levantamiento de barra en función de las semanas de entrenamiento.

x= Semanas de entrenamiento	0	2	4	6	8	10	12
y= % de incremento de fuerza	0	4	11	15	19	22	24

Esta relación se podría visualizar en la Figura 4.

Figura 4. Regresión lineal entre el porcentaje de incremento de la fuerza en levantamiento de barra y las semanas de entrenamiento.



La variable independiente se le considera como variable predictora. Es decir, trata de predecir el comportamiento de la variable de respuesta. En la figura 2 durante las primeras 12 semanas el incremento de fuerza logrado el linear con el tiempo. El tiempo es la variable independiente o explicativa y el % de incremento de fuerza es la variable dependiente o variable de respuesta. La línea roja representa la línea de tendencia, que describe el comportamiento de los datos cuando existe una relación lineal. Es el equivalente a trazar una línea que cruce por la distancia más cercana entre todos los puntos. Esta línea es posible calcularla por el método de los mínimos cuadrados y puede describirse con la ecuación de una línea recta. Por eso es también llamada recta de mejor ajuste o recta de mínimos cuadrados. Esta está dada por la siguiente ecuación:

$$\hat{y} = b_0 + b_1x$$

Donde:

- \hat{y} es la variable dependiente o variable de respuesta
- b_0 , es la intersección en y
- b_1 es la pendiente de la recta
- x es la variable independiente o predictiva.

Es de notar que b_0 , y b_1 , son estadísticos muestrales de los parámetros poblacionales β_0 y β_1 .

Para calcular la pendiente b_1 , usamos la siguiente ecuación:

$$b_1 = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Y para calcular el intercepto y (b_0), usamos la siguiente ecuación:

$$b_0 = \bar{y} - b_1\bar{x}$$

Donde:

- \bar{x} es la media de los valores de x
- \bar{y} es la media de los valores de y.

Resolviendo los factores en la ecuación 3 tenemos que:

$n= 7$ (tamaño de la muestra)

$\Sigma xy = 802$

$\Sigma x = 42$

$\Sigma y = 95$

$\Sigma x^2 = 364$

$$(\Sigma x)^2 = 1764$$

$$\bar{x} = 6$$

$$\bar{y} = 13.57$$

Entonces sustituyendo en la ecuación 3 tenemos que:

$$b_1 = \frac{7(802) - (42)(95)}{7(364) - 1764}$$

$$b_1 = 2.0714$$

Por lo tanto,

$$b_0 = 13.57 - 2.0714(6)$$

$$b_0 = 0.1416$$

Por lo que sustituyendo los coeficientes b_0 y b_1 en la ecuación de la recta, quedaría de la siguiente manera:

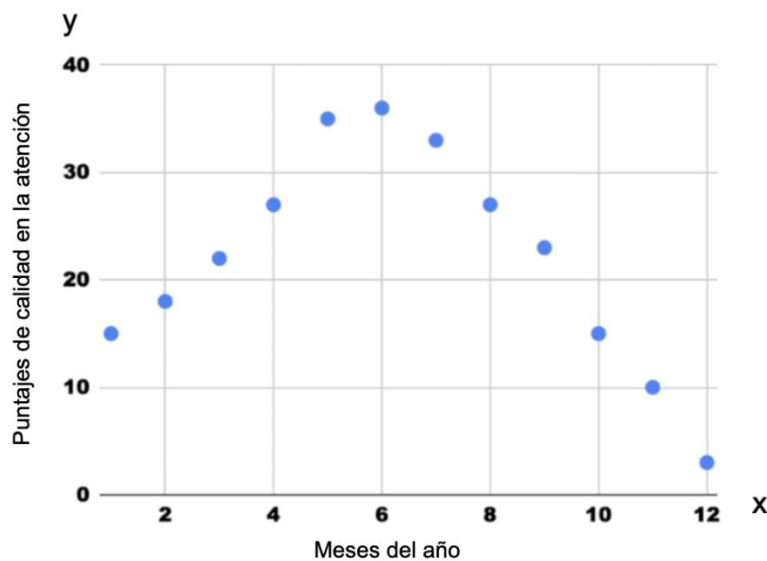
$$\hat{y} = 0.1416 + 2.0714(x)$$

Esto podría significar también que por cada semana de entrenamiento se aumentará aproximadamente un 2% de fuerza en la barra de pecho. Esto es al menos durante las primeras 12 semanas, donde la relación es aproximadamente lineal.

Otros ejemplos podrían ser, la relación entre la capacidad de enfoque y el nivel de aprendizaje, la calidad nutricional y la capacidad cognitiva, los niveles de ansiedad y la capacidad de concentración. Sin embargo, hay que recordar que no todas las regresiones son lineales, puede ser que una relación incremente, pero esta tendencia no se mantenga con el tiempo e inclusive pueda decaer. Este tipo de relación pudiera observarse al medir la asociación entre la calidad en la

atención y los meses que transcurren durante el año en una clínica de atención a pacientes con autismo, tal como puede observarse en la siguiente gráfica (Figura 5):

Figura 5: Relación entre los meses laborales y la atención al cliente.



Sabemos que la atención a personas con autismo puede al inicio ser muy recompensante y con mucho entusiasmo, sin embargo, con el paso de los meses, la atención puede decaer debido a la frustración que pudieran expresar los terapeutas. Esta relación no podría expresarse bajo una regresión lineal, sino que podría expresarse usando un modelo no lineal. Es decir, para interpretar o utilizar el método de regresión para análisis de datos debemos asumir que estamos investigando solamente relaciones lineales y que, para cada valor de x , y es una variable aleatoria con una distribución normal. Hay que recordar que hay relaciones entre variables que no se comportan de manera lineal dentro de una gráfica de dispersión de puntos y que hay otros métodos no lineales que

pueden explicar mejor la relación. Esto es cuando la tendencia de la relación es ascendente y luego descendente o viceversa.

Adicionalmente existen modelos de regresión lineal múltiple, que se utilizan para estudiar el efecto de múltiples variables independientes o variable predictoras sobre una variable dependientes o de respuesta. Por ejemplo, Pellegrini y Garces de los Fayos (2008) estudiaron diferentes variables que inciden en la conducta agresiva del niño. Se realizaron encuestas a una población muestral de 1800 estudiantes entre 11 y 16 años en distintos municipios de la provincia de Murcia, España. Los resultados del análisis de regresión múltiple de las variables predictoras con mayor efecto y sus respectivos coeficientes β de regresión incluyeron: Autocontrol social ($\beta = -0.508$, $p < 0.001$), Baja consideración a los demás ($\beta = -0.151$, $p < 0.001$), Liderazgo ($\beta = 0.062$, $p = 0.002$), Retraimiento ($\beta = 0.079$, $p < 0.001$), y Ansiedad ($\beta = -0.045$, $p = 0.037$), donde la variable dependiente fue la conducta antisocial. Los coeficientes β indican efecto de las variables, siendo la variable Autocontrol social la que tuvo mayor efecto (con una correlación negativa, es decir, menor autocontrol, mayor conducta antisocial reflejada), ya que indica la incapacidad del niño de mostrar dominio propio y actuar impulsivamente ante distintas estimulaciones, y de acuerdo al modelo esto explicaría el cerca del 50% de la conducta antisocial. La segunda variable con mayor efecto fue la baja consideración y respeto hacia a los demás, que es un reflejo de la inadaptabilidad social, que fomenta los comportamientos de falta de respeto, rebeldía a las normas establecidas y agresividad social (Pellegrini y Garces de los Fayos, 2008).

Conclusión

La correlación es una herramienta potente para establecer la asociación entre dos variables, las cuales pueden de ser de diferente tipo y estas asociaciones pueden ayudarnos a profundizar en el análisis de la conducta o la cognición humana. Puede ser muy fácilmente utilizada por profesionales de la salud para comprobar hipótesis correlacionales y puede ser proyectada usando la regresión lineal o no lineal.

En el campo de los educandos con necesidades especiales, puede ser una herramienta que arroje luz en cuanto a las variables que impactan tanto en el diagnóstico, como en la aplicación de distintas estrategias pedagógicas. Para esto, es indispensable tener muy claro, cuáles son las variables que pudieran verse relacionadas unas con otras. Esto con la finalidad de no cometer errores obteniendo correlaciones espurias, las cuales pueden tener significancia estadística, mas no valor teórico. De tal forma que el establecimiento de hipótesis correlacionales tiene un papel fundamental previo a la realización de pruebas estadísticas para obtener la correlación o regresión entre dos variables.

Referencias bibliográficas

- Al-Yagon, M. (2012). Adolescents with Learning Disabilities: Socioemotional and Behavioral Functioning and Attachment Relationships. *Journal of Youth and Adolescence*, 41, 1294-131. <https://doi:10.1007/s10964-012-9767-6>
- Croux C., y Dehon C. (2010) Influence functions of the Spearman and Kendall correlation measures. *Statistic Methods Applied*, 19, 497-515.

- García-Villamizar, D. (2017). Dolor en personas adultas con un trastorno del espectro del autismo (TEA) y comorbilidad: un análisis mediacional. *Revista Española de Discapacidad*, 5 (2), 73-86. <https://doi.org/10.5569/2340-5104.05.02.04>
- Hernandez-Sampieri R., Fernandez-Collado C., Baptista-Lucio M. del P. (2014) *Metodología de la investigación. 6a ed.* Mc GrawHill.
- Hevia-Orozco, J. y Sanz-Martin, A. (2018). EEG Characteristics of Adolescents Raised in Institutional Environments and Their Relation to Psychopathological Symptoms. *Journal of Behavioral and Brain Sciences*, 8, 519-537. <https://doi.org/10.4236/jbbs.2018.810032>
- Mondragón, M.A. (2014). Uso de la correlación de Spearman en un estudio de intervención de fisioterapia. *Movimiento científico*, 8(1), 98-104.
- Kendall M.G. (1938) A new measure of rank correlation. *Biometrika*, 30(1), 81-93.
- Pless-O'Connor, B. y Hevia-Orozco, J. (En preparación). Delving into the cognitive characteristics of gender roles in students.
- Pelegrín, A. y Garcés de los Fayos, E.J. (2008) Variables contextuales y personales que inciden en el comportamiento violento del niño. *European Journal of Education and Psychology*, (1)1, 5-20.
- Triola M. F (2004) *Estadística. 9a Ed.* Pearson Educación.

USO DE PRUEBAS NO PARAMÉTRICAS DE SIGNIFICACIÓN ESTADÍSTICA EN LA EDUCACIÓN ESPECIAL

Pedro Manuel Cortes Esparza [‡]

Citlali López Esparza

Juan Pablo García Hernández

Miguel Ángel Guevara

Resumen

Para abordar muchos de los retos que se presentan al estudiar el comportamiento a veces es necesario recurrir al análisis cualitativo. Esto adquiere especial relevancia para el educador especial, el cual necesita desarrollar métodos de intervención basados en evidencia científica que muchas veces deriva de este tipo de análisis como es el caso del estudio de la motivación, las emociones, u otros fenómenos subjetivos que tienen gran relevancia para el desarrollo integral del alumno. La estadística entonces provee herramientas para el análisis estructurado de este tipo de datos que, si bien son subjetivos por naturaleza, su interpretación se puede realizar de una manera detallada y objetiva. En el presente capítulo se buscará explicar de manera concisa la aplicación de este tipo de pruebas. Se iniciará con una breve explicación acerca del tipo de dato que resulta más apropiado para estos análisis y como discernir entre una prueba paramétrica y una no paramétrica. Entonces se hará una descripción detallada

[‡] Instituto de Neurociencias, Centro Universitario de Ciencias Biológica y Agropecuarias, Universidad de Guadalajara. Francisco de Quevedo 180, Col. Arcos Vallarta, Guadalajara, Jal. pedro.cortes@alumnos.udg.mx

de diferentes ejemplos de pruebas no paramétricas que se adaptan a diferentes paradigmas experimentales. Se espera que al final del capítulo, el educador especial sea capaz de distinguir estrategias de análisis estadístico para fenómenos subjetivos o cualitativos y promover en el lector el interés por aplicar estas metodologías para refinar las propuestas de métodos intervención y análisis de dichos resultados.

Palabras clave: Educación especial, escalas Likert, no paramétrico.

Introducción

La estadística es una herramienta poderosa a la hora de analizar información. No solo facilita su manejo, a la par nos da alternativas de análisis a partir de las cuales podemos tomar decisiones. Así, mediante la estadística podemos comprender la información que nos proveen los datos de una manera objetiva y tomar decisiones más acertadas.

Se han propuesto dos ramas principales de uso de la estadística: la estadística descriptiva y la estadística inferencial. Sí bien la estadística descriptiva facilita la representación y comprensión de un conjunto de datos, el presente capítulo abordará principalmente la estadística inferencial. Como su nombre lo sugiere, la estadística inferencial busca hacer inferencias o predicciones a partir de un conjunto de datos. Esto es útil a en el momento de tomar decisiones ya que podemos prever con cierto grado de exactitud lo que pasaría si tomamos una alternativa frente con respecto a otra. Estas predicciones están basadas en información previa, o como un conjunto de datos que son similares al que queremos predecir se ha comportado en situaciones similares.

Cuando hablamos de datos nos referimos a una representación simbólica, ya sea mediante números u otro tipo de representaciones como las letras, de

cierta información de nuestro interés. Por ejemplo, podemos representar el desempeño de los alumnos en una escuela mediante su calificación la cual, en México, habitualmente es representada con un número. El número de calificación se convierte entonces en una representación simbólica de un atributo que se busca evaluar, en este caso el desempeño escolar. Sin embargo, el desempeño escolar es algo abstracto y a veces complicado de definir, ya que dependiendo del evaluador se podría entender como diferentes cosas: un buen desempeño en la memoria, habilidades lógicas, facilidad de palabra, por mencionar algunos ejemplos. En contraste, hay algunos datos que representan atributos invariables, los cuales se pueden entender de igual manera sin importar quien sea el evaluador, por ejemplo: la estatura, la edad, el peso. Así, dependiendo del tipo de dato, difiere la cantidad de información que nos provee. Entre más cercano es al atributo que se quiere medir, más información contiene y suele ser representado de manera numérica. En el caso de los colores, podemos clasificar algo como azul. “Azul” entonces se convierte en un dato, pero con una representación abstracta ya que, dependiendo del interpretador, hay variaciones en el tono exacto de azul que se está buscando describir pudiendo ser un azul marino, cielo, rey, etc. Sin embargo, si por el contrario el dato para la representación del color es “450” nanómetros, tenemos una representación más exacta del color.

Con estos datos “exactos” se pueden hacer simulaciones de los fenómenos con mayor precisión. Estas representaciones se hacen mediante las pruebas estadísticas descritas como paramétricas, las cuales toman de supuesto principal que el conjunto de datos a analizar tiene características previamente establecidas lo cual se verifica con las pruebas de normalidad. Por ejemplo, al calcular el volumen de aire que cabe al inflar un globo, es más fácil hacer el

cálculo si suponemos que el globo forma una esfera perfecta. Así, las pruebas paramétricas hacen supuestos sobre cómo se debería de comportar el conjunto de datos a analizar con el objetivo de hacer predicciones más precisas de los fenómenos. Sin embargo, hay fenómenos que no tienen una alternativa de dato que cumpla con los supuestos necesarios de las pruebas paramétricas. Esto se vuelve cada vez más cierto cuando hablamos del comportamiento humano, el cual difícilmente tiene un comportamiento “estándar”.

Hay ciertos aspectos del comportamiento humano que son cuantificables de manera objetiva, por ejemplo: cuántas horas al día leen los alumnos, cuántas veces participan en clase, incluso cuántas faltas tienen en la materia. Sin embargo, conforme más se comienza a analizar las emociones o motivaciones que se presentan durante el desarrollo de estas conductas, más subjetivo se vuelve el fenómeno y por lo tanto se aleja de un comportamiento “normal” lo que plantea la utilización de una herramienta estadística diferente como las pruebas no paramétricas. Este conjunto de pruebas no emite suposiciones sobre el comportamiento de los datos para hacer inferencias por lo que se pueden aplicar a cualquier tipo de datos. Además, compensan la falta de precisión en sus predicciones teniendo una mayor robustez estadística. Es decir, a la hora de hacer predicciones son más estrictas que su contraparte, las pruebas estadísticas paramétricas.

En el presente capítulo se abordará el uso de las pruebas no paramétricas. Se hará una descripción detallada de qué son los datos, como se clasifican y debido a su comportamiento, qué tipos de datos son los más apropiados para ser analizados a través de las pruebas no paramétricas. Posteriormente se describirán las diferentes pruebas no paramétricas, comenzando con una descripción general de la prueba, casos en los que es apropiado utilizarla y un

ejemplo de su aplicación. Así, se espera que al final de este capítulo el lector tenga un panorama general de las pruebas no paramétricas y su aplicación.

Tipos de datos

La estadística es una rama de las matemáticas que tiene múltiples aplicaciones en distintas áreas de la vida cotidiana como el ámbito escolar, laboral, y de investigación entre otros (Mendenhall et al., 2010). Se le conoce como la ciencia de los datos, y proporciona pautas sobre su recopilación y la descripción (organización y resumen de los datos). Por lo tanto, un dato es una representación simbólica de algún atributo o característica que se está recopilando y, por tanto, contiene cierto grado de información. La información que contiene el dato suele representar ciertas características de una población que se busca evaluar, siendo la población el total de casos posibles de los cuales podemos extraer el dato en cuestión. Por ejemplo, si queremos evaluar el desempeño escolar en matemáticas de los niños en 5 de primaria, el dato sería la calificación de cada niño en matemáticas. mientras que la población serían todos los niños que asisten a ese grado. Es importante aclarar que, para sacar conclusiones sobre el desempeño en matemáticas, no es necesario evaluar a todos los niños, se pueden obtener los datos sólo de una cantidad representativa de estos, es decir se toma una muestra. Entre más grande es la muestra, la información que proporcionan los datos se vuelve más exacta, sin embargo, no suelen ser necesarios muestras grandes para hacer inferencias estadísticas adecuadas. El tamaño de la muestra depende en gran medida del tipo de dato con el que estamos trabajando, el tamaño total de la población y las características de la información que buscamos extraer. Es así como, dependiendo de la información que contiene el dato, estos se pueden clasificar

de manera jerárquica (Rumsey, 2013). De manera general, podemos clasificar los datos en cuantitativos y cualitativos.

Los datos cuantitativos, son aquellas representaciones numéricas de algún atributo o variable cuantificable. Al decir que podemos cuantificar algo, nos referimos a que podemos representar mediante números las diferentes magnitudes que tiene cierto atributo. Por ejemplo, podemos representar la edad de una persona mediante números, estos pueden ser años, meses, o incluso minutos. Dependiendo de la preferencia del lector, la edad se representa en diferentes unidades, la magnitud de ésta se mantiene invariable. Por ejemplo, es lo mismo describir que una persona tiene 22.5 años o 270 meses, la información relacionada a la edad permanece igual y queda adecuadamente representada de manera numérica. A diferencia de esto, existen los datos cualitativos, los cuales en lugar de representar magnitudes representan cualidades o características de ciertos atributos que se pretenden analizar. La representación de estos datos puede ser mediante números, letras, u otras representaciones simbólicas. Si bien los datos cualitativos se pueden representar mediante números, estos no son indicativos de una magnitud. Un ejemplo de esto son las escalas de Likert. Estas escalas se usan para medir el acuerdo o desacuerdo de una persona respecto a un tema. Para contestarla, al entrevistado se le hace una pregunta o afirmación y se le pide que conteste del 0 al 4 que tan de acuerdo está con el enunciado. El 0 representa totalmente en desacuerdo y el 4 un completamente de acuerdo. Con este ejemplo, se puede recopilar datos numéricos, por ejemplo, un 4 representaría estar de completamente acuerdo. Sin embargo, ese 4 solo tiene sentido en el contexto de la escala y el enunciado en particular referente al cual se contestó. Si cambiamos la escala del 1 al 10, el significado del dato “4” cambia su sentido a estar ligeramente en desacuerdo. No solo esto, dependiendo

del interpretador el valor del 1 es diferente ya que el valor de enunciados subjetivos como “estar de acuerdo” pueden variar de una persona a otra. Así, estar de acuerdo para la persona A puede ser equivalente a un ligeramente de acuerdo para la persona B. Estas diferencias interpersonales dificultan hacer una equivalencia numérica entre estos datos subjetivos recopilados de diferentes personas. Es por esto por lo que, en la escala de Likert, el número 4, en lugar de representar una magnitud, representa un enunciado que la persona en cuestión categorizó y ordenó de acuerdo con sus preferencias personales. Esta es la esencia de los datos cualitativos, que pueden ayudar a representar fenómenos abstractos para los cuales difícilmente tenemos una representación en términos de magnitudes ya que no son medibles de forma directa. Este tipo de análisis adquiere importancia en el ámbito del comportamiento humano y la educación. Actualmente es imposible medir de manera directa y objetiva conceptos como la motivación del estudiante, su esfuerzo o su desempeño escolar. Sin embargo, estos son conceptos que resulta de nuestro interés evaluar. En estos casos podemos recopilar datos cualitativos, y representar la información mediante números u otros símbolos para simplificar su uso. Ya con estos datos la estadística nos proporciona herramientas para su posterior análisis y manejo.

Estadística no paramétrica

Una vez que tenemos capturados los datos, la estadística nos da opciones para su representación. Esto lo hace mediante la elaboración de tablas o gráficas que varía de acuerdo con el tipo de dato que estamos manejando y con la información que buscamos extraer de estos. Sin embargo, la estadística no se queda ahí, también nos permite tomar el conjunto de datos y hacer predicciones sobre su comportamiento, aunque solo se observe una pequeña parte de estos

(Daniel, 1991). Así, para conocer la estatura de las mujeres no necesitamos medir a todas las mujeres del mundo. Podemos recopilar una muestra de ellas y con base en esos datos, hacer una predicción sobre el rango de estatura en los que se puede encontrar una mujer a una edad específica. Estos análisis se hacen con el apoyo de tablas de valores críticos para cada prueba estadística. Estas tablas contienen un valor que la prueba estadística debe de superar a un tamaño de muestra (n) y a un nivel de significancia (α) determinados por el analista. La significancia es un indicador del grado de certeza de las inferencias que estamos realizando y por convenio general se suele usar una confiabilidad de 0.05 que corresponde a un 95% de confianza en la predicción. Esta rama de la estadística se conoce como estadística inferencial y nos permite obtener conclusiones de un conjunto de datos con cierto grado de certeza o confiabilidad (Johnson y Kuby, 2008).

A pesar de que a través de los datos podemos obtener información, es necesario confirmar y representar la validez de ellos. Para esto es necesario el uso de análisis estadísticos los cuales deben de ser seleccionados cuidadosamente, ya que deben estar basados en las características del fenómeno y las variables estudiadas. Existen dos tipos de análisis estadísticos que pueden ser utilizados para probar las hipótesis, uno de ellos es el análisis de datos paramétricos. Este tipo de análisis presenta un alto poder predictivo y se basa en el uso de modelos de comportamiento de datos ya establecidos. Hay evidencia de que la mayoría de los datos cuantitativos tienden a una distribución “normal”, es decir, conforme la cantidad de datos aumenta, parámetros estadísticos como la distribución de los datos alrededor de la media se comparten de una manera similar no importa qué variable esté representando el dato. El tener un comportamiento estándar facilita hacer predicciones ya que,

tomando una pequeña parte de los datos, se puede inferir cómo se comporta la mayoría de estos. Sin embargo, hacer análisis basados en modelos de normalidad tiene la desventaja de que los datos utilizados deben cumplir con ciertos criterios o “parámetros” de normalidad, de ahí el nombre de análisis paramétricos. Además, no solo se requiere que un dato sea cuantitativo sino también tiene que cumplir con criterios como una distribución alrededor de la media normal y homocedasticidad.

Cuando los datos no cumplen con esos criterios, ya sea que su distribución no sea normal o que el tipo de dato sea cualitativo, se tiene que utilizar Otro tipo de pruebas denominadas no paramétricas (Flores-Ruiz et al., 2017; Hernández-Sampieri et al., 2008). Como las pruebas no paramétricas evitan hacer supuestos de normalidad en los datos, su poder predictivo es menor que las alternativas paramétricas. Lo anterior repercute en los análisis no paramétricos sean más estrictos, pero a la vez más confiables ya que no hacen supuestos respecto a la población de la cual provienen los datos. Las alternativas de análisis no paramétricos se pueden utilizar no importando el tipo de dato y su comportamiento, sin embargo, dependiendo del objetivo de análisis y la manera en que los datos fueron recolectados el emplear una prueba resulta más adecuado que otra.

Por lo general, las pruebas estadísticas se basan en la comparación entre dos o más conjuntos de datos. El objetivo de su utilización es probar si las poblaciones de las que provienen los datos son idénticas (hipótesis nula) o si por el contrario hay una diferencia significativa entre estas (hipótesis alternativa). Las diferencias entre dos poblaciones pueden ser intrínsecas (como en el caso de factores genéticos) o asociadas al efecto de un tratamiento experimental (como en el caso de un fármaco). Una primera distinción, son los

análisis para grupos correlacionados o grupos independientes. Se considera que dos conjuntos de datos están correlacionados cuando comparten características similares. Un ejemplo de esto son las medidas repetidas, es decir, cuando a un participante se le toma una medición antes y otra después de un tratamiento experimental. Una aplicación de esto sería evaluar los conocimientos en matemáticas antes y después del curso. Otro ejemplo de datos correlacionados son los grupos pareados. En estos casos las medidas se realizan a sujetos diferentes pero que comparten muchas de sus características como puede ser la edad, nivel socioeconómico, género entre otros. La aplicación más representativa de esto es hacer una medición en gemelos idénticos, en donde uno puede recibir el tratamiento experimental mientras que el otro sirve de control o referencia. En contraste, las pruebas para grupos independientes no requieren que los conjuntos de datos compartan características similares. Se puede hacer una comparación entre niñas y niños o jóvenes y adultos. Sin embargo, cabe aclarar que entre más delimitada está la muestra mayor potencia estadística tendremos. Es decir, si se van a evaluar “niños” que estos estén en un rango de edad claramente definido. Por lo tanto, tomando esta primera consideración podemos definir si los datos no paramétricos que buscamos comparar son correlacionados o independientes y luego buscar la prueba estadística que más se adapte a ellos.

Pruebas no paramétricas para grupos correlacionados

Como se mencionó anteriormente, las pruebas para grupos correlacionados consideran que el conjunto de datos a comparar comparte muchas de sus características en común, como puede ser el caso de el antes contra el después. Entre más similares son las características de los sujetos que se están

comparando, más adecuada es la aplicación de este tipo de pruebas. A diferencia de las pruebas para grupos independientes, estas no buscan una diferencia entre las poblaciones de las cuales provienen los conjuntos de datos. Las pruebas para grupos correlacionados buscan un cambio entre los dos estados a comparar.

T de Wilcoxon

La prueba de T de Wilcoxon es una prueba no paramétrica, permite comparar dos grupos pareados o correlacionados. Se utiliza para probar la hipótesis de que las dos poblaciones de la que provienen los datos son idénticas (Mendenhall et al., 2010; Triola et al., 2009). La prueba de Wilcoxon tiene su equivalente en pruebas paramétricas en la t de Student para grupos correlacionados (Bautista-Díaz et al., 2020). La prueba de Wilcoxon asume que se ha obtenido una muestra aleatoria de datos pareados o que cada parte del par de datos ha sido asignada aleatoriamente a las condiciones experimentales. La selección de la muestra para este tipo de análisis puede obtenerse por distintas técnicas, como puede ser la obtención de medidas repetidas en los mismos participantes (un solo participante pasa por dos condiciones distintas), el uso de gemelos idénticos o compañeros de camada, o la obtención de pares de participantes apareados por selección mutua, entre otras técnicas (Mendenhall et al., 2010). A continuación, se describe un ejemplo de cómo se desarrolla la prueba T de Wilcoxon en el ámbito de la educación.

Un terapeuta diseñó un programa de intervención para mejorar el cálculo en niños con problemas en dicho proceso. Para conocer la efectividad de su programa, aplicó una prueba de 10 ejercicios matemáticos antes y después de la intervención, la cual, fue estudiada en un solo grupo de 10 niños, dividido en

dos condiciones (tabla 1), el grupo “A” conformado por los niños que respondieron la prueba antes de someterse a la intervención, mientras que el grupo “B” corresponde a los mismos niños, que contestaron la prueba después de someterse al programa de intervención. A continuación, se presenta la tabla con el número de errores en la prueba antes y después de la intervención.

Tabla 1. Número de errores obtenidos por condición

Grupo A (Sin práctica)	Grupo B (Con práctica)
8	4
3	3
5	2
2	1
4	2
6	1
7	3
2	1
5	6
3	2

Para concluir si el programa de intervención fue efectivo, el terapeuta aplica la prueba de T de Wilcoxon a los datos que se van a estudiar (el número de errores en una prueba de cálculo antes y después de la intervención). La hipótesis nula es que el programa de intervención no tiene un efecto en el cálculo de los niños y por lo tanto, obtendrán la misma cantidad de errores en una prueba a pesar del tratamiento. Por otra parte, la hipótesis alternativa propone que, el programa de intervención mejora el proceso del cálculo en los niños y por consiguiente los niños presentarán una menor cantidad de errores en la prueba después de pasar por el programa de intervención.

El primer paso en este análisis es calcular la diferencia entre los datos de ambos conjuntos (d), esta se obtiene al restar el segundo valor (grupo B) al primero (grupo A), en cada fila de datos. Posteriormente, las diferencias obtenidas deben de ser ordenadas de menor a mayor, sin importar su signo (+ -), los valores a los que su diferencia sea equivalente a 0, no se tomarán en cuenta para su ordenamiento. Cuando los valores a los que se les asignará un rango se repiten, los rangos correspondientes deben promediarse y de esta forma se asigna el valor equivalente al promedio del rango de las casillas que le corresponden (Triola et al., 2009) (tabla 2).

Tabla 2. Diferencia y rango asignada a cada dato

Grupo A	Grupo B	Diferencia (d)	Rango
8	4	4	7.5
3	3	0	
5	2	3	6
2	1	1	2.5
4	2	2	5
6	1	5	9
7	3	4	7.5
2	1	1	2.5
5	6	-1	2.5
3	2	1	2.5

Una vez que se asignan los rangos, se agrega el signo (+ -) correspondiente a la diferencia que se obtuvo anteriormente, y se realiza la sumatoria de los rangos positivos y posteriormente de los negativos. El valor de la T de Wicolxon, la diferencia entre la sumatoria positiva y la negativa. Una vez que tenemos el valor de T, se busca el valor crítico en la tabla de valores

críticos de la T de Wilcoxon, en el caso de las hipótesis unilaterales o de una sola cola, se considera el valor alfa de 0.05, y se busca el valor crítico correspondiente a la cantidad de la muestra (n), si el valor de T es menor al valor crítico se rechaza la hipótesis nula (Kirk, 2008; Triola et al., 2009). En el caso del ejemplo descrito anteriormente, el valor de T es de 40, en la tabla de valores críticos, a la n de 10, corresponde el valor crítico de 8. Ya que el valor de T calculado supera el valor crítico indicado por la tabla se rechaza la hipótesis nula y se acepta la hipótesis alterna. Por tanto, en este ejemplo se concluye que los niños presentan una menor cantidad de errores en la prueba después de pasar por el programa de intervención para estimular el cálculo (Tabla 3).

Tabla 3. Diferencias y rangos ordenados

Sujeto	Grupo A	Grupo B	Diferencia (d)	Rango
2	3	3	0	N/A
4	2	1	1	2.5
8	2	1	1	2.5
9	5	6	-1	2.5
10	3	2	1	2.5
5	4	2	2	5
3	5	2	3	6
1	8	4	4	7.5
7	7	3	4	7.5
6	6	1	5	9
Suma rangos +		Suma rangos -		Suma total
42.5		2.5		40

Pruebas no paramétricas para grupos independientes

Como se mencionó anteriormente, las pruebas para grupos independientes consideran que los conjuntos de datos a comparar no comparten características en común, así que se puede comparar entre sujetos de diferentes edades o

géneros. El fundamento principal de esta prueba es que el sujeto del cual se obtuvo el dato A no es el mismo que el del cual se obtuvo el del dato B. Sin embargo, al final en estas pruebas se busca verificar si, a pesar de estas diferencias intrínsecas entre los sujetos, el dato que se está recopilando se puede asumir que proviene de una misma población (los datos son similares) o si por el contrario son de poblaciones diferentes (hay una entre los datos). Estas diferencias entre los conjuntos de datos por lo general se asocian a un tratamiento experimental.

Prueba de Ji cuadrada o X^2

La prueba de Ji cuadrada es una prueba no paramétrica que nos permite comparar más de dos grupos independientes, se utiliza para comparar las posibles diferencias en las frecuencias observadas en la distribución de una variable con respecto a las frecuencias esperadas. Es decir, se suele usar al momento de comparar dos valores únicos como por ejemplo número de niños contra el número de niñas en las escuelas. En este tipo de dato a cada observación se le asignó una categoría (niño/niña) y se cuenta cuántos casos hay de cada observación. A este tipo de datos se le llama categóricos (Gómez-Gómez et al., 2003; Siegel y Castellan, 1998). A continuación se describe un ejemplo de cómo se desarrolla la prueba de Ji cuadrada en el ámbito de la educación especial.

En una investigación se pretende reportar si existen diferencias en el rendimiento escolar reportado por los docentes (reportado como adecuado, regular e inadecuado) en niños con trastorno por déficit de atención (TDAH) con un tratamiento médico y sin tratamiento. Se estudiaron dos grupos distintos,

uno correspondiente a la condición experimental sin tratamiento y el otro corresponde a la condición experimental con tratamiento. La variable a estudiar (rendimiento escolar) se categorizó en: adecuado (cuando los niños pueden realizar sus actividades escolares, pueden seguir instrucciones y pueden iniciar y terminar las actividades de forma correcta) regular, (cuando los niños presentan algunas dificultades para seguir instrucciones, pueden iniciar las actividades pero en ocasiones no las terminan, lo cual impacta en cómo los niños realizan sus actividades escolares) e inadecuado (cuando los niños no pueden seguir instrucciones, tienen dificultad para comenzar y terminar las actividades por lo que sus actividades escolares se ven afectadas por completo). La muestra se conformó por 20 niños de 8 años, y sus profesores asignaron la categoría correspondiente a su desempeño escolar (véase tabla 4). La hipótesis nula es que no hay diferencias en el rendimiento escolar en niños con TDAH que se encuentran bajo tratamiento médico y niños con TDAH sin tratamiento, mientras que la hipótesis alterna postula que los niños con TDAH que se encuentran bajo tratamiento médico tendrán un mejor desempeño escolar respecto a los niños con TDAH sin tratamiento.

Tabla 4. Condiciones experimentales

Participante	Tratamiento		Reporte de rendimiento
	Con tratamiento	Sin tratamiento	
1		X	Regular
2	X		Adecuado
3	X		Regular
4		X	Inadecuado
5		X	Inadecuado
6	X		Regular
7	X		Adecuado
8	X		Regular
9		X	Regular
10	X		Adecuado
11		X	Inadecuado
12		X	Regular
13	X		Inadecuado
14	X		Adecuado
15		X	Adecuado
16	X		Adecuado
17		X	Regular
18		X	Inadecuado
19	X		Adecuado
20		X	Inadecuado

Los datos deben ser ordenados y presentados en una tabla de frecuencia o contingencia, en la cual, las columnas representan a los grupos que se estudian y cada fila representa una categoría de las variables por medir, se asignará un número de celda a cada una de las casillas y se realizará un conteo de la cantidad de cada participante que corresponde a las condiciones y variables descritas en la tabla, así como los totales para cada condición experimental y cada variable estudiada (véase tabla 5), que se denominan como datos observados. (Johnson y Kuby, 2008; Siegel y Castellan, 1998) Para obtener el valor de Ji cuadrada es necesario obtener el valor de las frecuencias esperadas, este se obtiene mediante la siguiente operación:

$$E = \frac{(R)(C)}{N}$$

Donde:

- E= frecuencia esperada
- R= Número total de cada renglón
- C=Número total de cada columna
- N= Número total de datos

Para calcular la frecuencia esperada de cada celda, se multiplican los totales de cada renglón o fila por los totales de cada columna, el resultado se divide entre el número total de datos (Johnson y Kuby, 2008). De acuerdo con el ejemplo mencionado anteriormente (tabla 6), para calcular la frecuencia esperada de la celda uno, la ecuación se sustituye de la siguiente forma:

$$E = \frac{(7)(10)}{20}$$

El resultado de dicha ecuación es 3.5, que corresponde al valor de la frecuencia esperada de la celda uno, el mismo procedimiento se realiza para el resto de las celdas.

Tabla 5. Asignación de celdas y totales correspondientes

Reporte de rendimiento	Tratamiento		Total
	Con tratamiento	Sin tratamiento	
Adecuado	6 celda 1	2 celda 4	7
Regular	3 celda 2	4 celda 5	7
Inadecuado	1 celda 3	5 celda 6	6
Total	10	10	20

Una vez que se han calculado las frecuencias esperadas, se obtendrá el valor de Ji cuadrada mediante la siguiente ecuación (Johnson y Kuby, 2008):

$$X^2 = \Sigma \left(\frac{O - E}{E} \right)^2$$

Donde:

- O= datos observados
- E= frecuencias esperadas

Continuando con el ejemplo descrito, la sustitución de la ecuación con los valores correspondientes (tabla 6) sería la siguiente:

$$x^2 = \frac{(6 - 3.5)^2}{3.5} + \frac{(3 - 3.5)^2}{3.5} + \frac{(1 - 0.35)^2}{0.35} + \frac{(1 - 0.35)^2}{0.35} + \frac{(4 - 1.4)^2}{1.4} + \frac{(5 - 1.5)^2}{1.5}$$

Al realizar las operaciones, se obtiene como valor de Ji cuadrada el siguiente:

$$x^2 = 17.774$$

Tabla 6. Datos esperados y frecuencias observadas por celda

Reporte de rendimiento	Tratamiento		Total
	Con tratamiento	Sin tratamiento	
Adecuado	celda 1	celda 4	7
	O = 6	O=1	
	E=3.5	E=0.35	
Regular	celda 2	celda 5	7
	O=3	O=4	
	E=3.5	E=1.4	
Inadecuado	celda 3	celda 6	6
	O=1	O=5	
	E=0.35	E=1.5	
Total	10	10	20

Una vez que se ha calculado el valor de ji cuadrada, es necesario calcular los grados de libertad de acuerdo con nuestra cantidad de datos (Johnson y Kuby, 2008). Para esto se realiza la siguiente operación:

$$gl = (c - 1)(r - 1)$$

Donde:

- gl= grados de libertad
- c= número de columnas
- r =número de renglones o filas

De acuerdo con el ejemplo que se ha desarrollado la sustitución de los valores para obtener los grados de libertad es la siguiente:

$$gl = (2 - 1)(3 - 1) = 2$$

Una vez que se tiene el valor de Ji cuadrada y los grados de libertad, se busca el valor crítico en la tabla de valores críticos de la distribución de Ji cuadrada, para poder rechazar o aceptar las hipótesis; Para rechazar la hipótesis nula, el valor crítico de acuerdo con los grados de libertad obtenidos, (2, en caso del ejemplo anterior) y el nivel de significancia de 0.05 obtenido de la tabla debe ser menor que el valor calculado de Ji cuadrada (Siegel y Castellan, 1998). El valor crítico que corresponde al ejemplo desarrollado es de 5.99, y el valor que se obtuvo de Ji cuadrada es de 17.774, por lo que, al ser mayor, podemos rechazar la hipótesis nula y aceptar la hipótesis alternativa, que postula que los niños con TDAH que se encuentran bajo tratamiento médico tienen un mejor desempeño escolar respecto a los niños con TDAH sin tratamiento.

Prueba U de Mann-Whitney

La prueba de la U de Mann Whitney forma parte del conjunto de pruebas no paramétricas, puede utilizarse para evaluar si dos conjuntos de datos no paramétricos provienen o no de la misma población (Siegel y Castellan, 1998). La prueba de la U de Mann-Whitney es la prueba equivalente a la prueba paramétrica de t de Student, que permite comparar las diferencias en poblaciones con una distribución normal (Bautista-Díaz et al., 2020). Esta prueba se basa en los rangos de las observaciones y no en sus valores numéricos, por lo que es apropiado para la mayoría de los datos de las ciencias del comportamiento y la educación (Kirk, 2008). A continuación, se describe un ejemplo de cómo se desarrolla la prueba de la U de Mann Whitney en el ámbito de la educación.

Un grupo de profesores quiere saber si recibir apoyo extracurricular mejora las calificaciones en un grupo de niños con problemas de aprendizaje respecto a un grupo de niños que no recibe apoyo extracurricular. Los profesores estudiaron dicha hipótesis en dos grupos (tabla 7), el grupo “A” correspondiente a los niños que reciben apoyo en casa y un grupo “B” que representa a los niños que no reciben apoyo en casa para estudiar y realizar sus tareas; a continuación se muestran las calificaciones obtenidas por cada grupo, conformado por una n de 7 niños por cada grupo.

Tabla 7. Calificaciones obtenidas por grupo

Grupo A (con apoyo)	Grupo B (sin apoyo)
85	65
70	67
69	71
87	73
70	82
86	66
99	62

Para determinar si existen diferencias entre ambos grupos, los profesores aplican la prueba de U de Mann Whitney a los datos que van a estudiar (las calificaciones obtenidas por cada grupo). La hipótesis nula es que no existen diferencias entre las calificaciones en niños con problemas de aprendizaje que reciben apoyo extracurricular, mientras que la hipótesis alterna es que se

observará un mejor rendimiento en las calificaciones de los niños que sí reciben apoyo extracurricular.

El primer paso en este análisis es ordenar las puntuaciones por rangos. Al asignar rangos a las puntuaciones, los datos de los grupos experimental (con apoyo) y control (sin apoyo) se tratan como una sola muestra (Kirk, 2008). Para calcular la estadística de prueba, se combinan las dos muestras y se acomodan por rangos todas las observaciones, desde las más pequeñas hasta las más grandes, teniendo presente a cuál muestra pertenece cada observación (Daniel, 1991) (Tabla 8).

Tabla 8. Datos ordenados por rangos.

Rango	Calificación	Grupo
1	62	B
2	65	B
3	66	B
4	67	B
5	69	A
6.5	70	A
6.5	70	A
8	71	B
9	73	B
10	82	B
11	85	A
12	86	A
13	87	A
14	90	A

Se asignan rangos para cada uno de los datos que se van a analizar, con la puntuación más pequeña recibiendo un rango de 1 y la más grande recibiendo un rango de n (14 para el rango en el ejemplo). Si dos o más observaciones tienen el mismo valor (puntuaciones empatadas), se les asigna la media de los rangos que hubieran ocupado, por ejemplo, en un conjunto de 5 datos (6, 3, 4, 5, 4) el dato de menor valor (3) ocupa el menor rango (1), los siguientes valores se repiten (4), por lo que el rango 2 y 3 que son sus casillas correspondientes se suman y se promedian, obteniendo un rango de 2.5 para cada uno, el siguiente rango corresponde a la casilla 4, en el cual se asigna el valor 5, y por último el rango correspondiente a la casilla 5, se asigna el dato 6 (Kirk, 2008). La prueba de la U de Mann Whitney se basa en la suma de los rangos como se indica en la tabla 9.

Tabla 9. Sumatoria de los rangos por grupo

Grupo A	Grupo B
	1 (62)
	2 (65)
	3 (66)
	4 (67)
5 (69)	
6.5 (70)	
6.5 (70)	
	8 (71)
	9 (73)
	10 (82)
11 (85)	
12 (86)	
13 (87)	
14 (90)	
68	37

Una vez que se realizó la sumatoria de los rangos, se obtendrá el valor de U de cada grupo. Esto se realizará mediante la siguiente fórmula (Kirk, 2008).

$$U1 = n1 n2 + \frac{n1 (n1 + 1)}{2} - \Sigma R1$$

$$U2 = n1 n2 + \frac{n2 (n2 + 1)}{2} - \Sigma R2$$

En cuanto al ejemplo mencionado anteriormente, para obtener el valor de U1 se multiplica el valor de la muestra del grupo A, por la muestra del grupo B, ambos con 7 participantes cada uno (tabla 7), a dicho resultado, se sumará al valor obtenido al multiplicar n1 por n1 más 1 entre 2. Por último, se restará el valor de la sumatoria total de los rangos en el grupo 1. Para obtener el valor de U2 se realiza el mismo procedimiento a excepción de la sustitución de los valores correspondientes a n2 y a R2. La fórmula que se utiliza para analizar el ejemplo descrito anteriormente sería la siguiente:

$$U1 = (7) (7) + \frac{7 (7 + 1)}{2} - 68$$

$$U2 = (7) (7) + \frac{7 (7 + 1)}{2} - 37$$

En la prueba de U de Mann Whitney se adopta el nivel de significancia del 0.05. Para que el resultado sea significativo, el valor de U obtenido debe ser menor o igual que el valor crítico que se obtiene en la tabla de valores de U. Cuando el valor calculado de U es menor que el valor crítico correspondiente a la cantidad de la muestra de cada grupo, se rechaza la hipótesis nula y se acepta la hipótesis alternativa (Kirk, 2008). Con base al ejemplo desarrollado anteriormente, el valor de U1 es de 9, mientras que el valor d U2 corresponde a

40, por lo tanto, se tomará el menor valor, que es el de U_1 , y se busca el valor correspondiente en la tabla de valores críticos correspondientes a la prueba de U , obteniendo un valor crítico de 11. Al ser menor el valor obtenido de U_1 , se puede rechazar la hipótesis nula, la cual postula que no existen diferencias en las calificaciones de los niños que reciben apoyo extracurricular y los que no la reciben. Además, se acepta la hipótesis alternativa, que propone que si existe una diferencia en la población estudiada, así como un mejor rendimiento en calificaciones ante la variable mencionada anteriormente.

Prueba H de Kruskal-Wallis

La prueba de Kruskal-Wallis es una prueba no paramétrica que utiliza rangos de muestras de tres o más poblaciones independientes, esta prueba permite confirmar o negar que existen diferencias entre los distintos grupos (Triola et al., 2009). La prueba de H de Kruskal-Wallis tiene como objetivo evaluar la diferencia entre tres o más conjuntos de datos que provienen de grupos independientes y tiene su contraparte paramétrica en el análisis de variancia para un factor (ANOVA) (Bautista-Díaz et al., 2020). A continuación se describe un ejemplo de cómo se desarrolla la prueba H de Kruskal-Wallis en el ámbito de la educación.

Una psicóloga pretende conocer si existen diferencias en la comprensión lectora en niños con problemas de aprendizaje, que difieren en la cantidad de lectura que dedican fuera de la escuela. Se estudiaron a 3 grupos con 4 participantes cada uno, el primer grupo no realizaba lectura fuera de la escuela, el segundo realizaba 15 minutos de lectura en casa diariamente y el tercer grupo realizaba 30 minutos de lectura, también en casa diariamente. Para medir el

efecto de la lectura extracurricular se realizó una prueba en la que se les presentaba un ejercicio de lecto-comprensión, el cual, consistía en un breve texto y 5 preguntas sobre el contenido de la lectura, el número de errores era un indicador de dificultad en la lecto-comprensión. La hipótesis nula es que no existen diferencias en la lecto-comprensión independientemente del tiempo que se realiza lectura fuera de la escuela, mientras que, la hipótesis alternativa postula que existen diferencias en la lecto-comprensión de acuerdo con el tiempo que se realiza lectura fuera del horario escolar (tabla 10).

Tabla 10. Número de errores en prueba de lecto-comprensión por grupo

Grupo A (sin lectura extracurricular)	Grupo B (con lectura durante 15 minutos)	Grupo C (con lectura durante 30 minutos)
1	3	4
3	2	5
2	5	5
1	3	3

El primer paso para realizar la prueba H de Kruskal-Wallis es combinar temporalmente todos los valores de cada grupo, y asignar un rango a cada valor muestral como se ha mencionado en la asignación de rangos en la prueba de U de Mann Whitney y la prueba T de Wicolxon (Tabla 11), en caso de encontrar valores muestrales repetidos, se asigna a cada observación la media de los rangos implicados (Daniel, 1991; Triola et al., 2009). Una vez que los rangos han sido asignados, se suma el total de los rangos de cada grupo o condición (Tabla 12).

Tabla 11. Rangos ordenados correspondientes a las calificaciones por cada grupo.

Rango	Número de errores	Grupo
1.5	1	A
1.5	1	A
3.5	2	A
3.5	2	B
6.5	3	A
6.5	3	B
6.5	3	B
6.5	3	C
9	4	C
11	5	B
11	5	C
11	5	C

Tabla 12. Rangos asignados

Grupo A (sin lectura extracurricular)	Grupo B (con lectura durante 15 minutos)	Grupo C (con lectura durante 30 minutos)
(1) 1.5	(3) 6.5	(4) 9
(3) 6.5	(2) 3.5	(5) 11
(2) 3.5	(5) 11	(5) 11
(1) 1.5	(3) 6.5	(3) 6.5
Suma de rangos por grupo		
13	27.5	37.5

Con la sumatoria de los rangos se prosigue a obtener el valor de H mediante la siguiente ecuación:

$$H = \left[\frac{12}{n(n+1)} \sum \frac{T^2}{n^j} \right] - 3(n+1)$$

Donde:

- n = total de participantes
- nj =total de participantes en cada grupo
- T²= Sumatoria de los rangos por columna al cuadrado

En cuanto al ejemplo descrito anteriormente, al sustituir la ecuación con los valores correspondientes, podría representarse de la siguiente forma:

$$H = \left[\frac{12}{12(13)} \left(\frac{169}{4} + \frac{756.25}{4} + \frac{1406.25}{4} \right) \right] - 3(13)$$

El valor obtenido de H:

$$H = 8.8365$$

Para rechazar la hipótesis nula y aceptar la hipótesis alternativa, es necesario buscar el valor crítico en la tabla de valores de la prueba de Kruskal-Wallis, se seleccionará el valor crítico correspondiente a el número de datos muestrales totales para cada grupo y el valor de alfa (significancia) de 0.05. Si el valor de H calculado es mayor al valor crítico que se obtuvo de la tabla, se rechaza la hipótesis nula y se acepta la hipótesis alterna. En el caso del ejemplo utilizado anteriormente, se trabajó con tres grupos cada uno con 4 participantes. El valor crítico de tabla correspondiente a nuestros datos es de 7.6538 y el valor de H calculado es de 8.8365, por lo que al ser mayor, podemos rechazar la

hipótesis nula que postula que no existen diferencias en la lecto-comprensión independientemente del tiempo que se realiza lectura fuera de la escuela, y se acepta la hipótesis alternativa que confirma que existen diferencias en la lecto-comprensión de acuerdo al tiempo que se realiza lectura fuera del horario escolar. Cuando se tienen tres grupos y cinco o menos observaciones en cada grupo, el valor de H se obtiene mediante la tabla de valores críticos de la prueba de Kruskal-Wallis, sin embargo, cuando se presentan más de cinco observaciones o más de tres grupos, H se compara con los valores tabulados de la χ^2 , con k -1 grados de libertad (Daniel, 1991).

Pruebas no paramétricas para varios factores

Existen casos en los que es necesario conocer, no solo el efecto de una variable, sino el efecto de la suma de dos variables, es decir, si los cambios entre los conjuntos de datos se deben al efecto combinado de más de un tratamiento experimental o si cada tratamiento actúa de manera independiente. Estas son las llamadas pruebas de interacción y para su cálculo en el caso de que las variables sean no paramétricas y los grupos sean independientes existe la prueba exacta de Fisher; para mediciones correlacionadas, es decir un mismo grupo evaluado en más de una ocasión (pre y post test, por ejemplo), existe la prueba de McNemar (Siegel y Castellan, 1995).

Muestras independientes: prueba exacta de Fisher

La prueba exacta de Fisher es utilizada para conocer si existe relación entre dos variables categóricas o cualitativas (aquellas que describen datos con

base a cualidades o categorías) cuando las muestras son obtenidas de grupos independientes (Freeman y Campbell, 2007).

Por ejemplo, si se busca conocer la prevalencia de dos trastornos del desarrollo (variable A): trastorno del espectro autista y trastorno de déficit de atención, y su relación con el sexo (variable B): niños y niñas, con el objetivo de conocer si la prevalencia de alguno de estos dos trastornos es afectada por el sexo del infante. De manera hipotética, podría esperarse que haya más prevalencia de trastorno del espectro autista en niños respecto a las niñas y respecto a los al trastorno de déficit de atención. Por lo que la hipótesis nula sería que hay independencia entre estos dos trastornos del desarrollo y el sexo, y la hipótesis alternativa señala que sí habrá dependencia entre las variables, es decir, que habrá mayor prevalencia del trastorno del espectro autista específicamente en los niños.

Para evaluar y comprobar lo anterior, primero se debe realizar una tabla de contingencias en donde se vacía la frecuencia real de cada variable en cada nivel, además del total de las columnas y de las filas (Freeman y Campbell, 2007) (Tabla 13).

Tabla 13. Identificación de celdas en tabla de contingencia.

Trastorno \ sexo	Niña	Niño	Total
Espectro autista	A	B	A+B
Hiperactividad	C	D	C+D
Total	A+C	B+D	N

Considerando que la variable A tiene los niveles de trastorno del espectro autista y trastorno de déficit de atención, y la variable B los niveles de niña y niño, los datos quedarían como se observa a continuación (Tabla 14).

Tabla 14. Observaciones registradas para la variable de trastornos del desarrollo y sexo

trastorno \ sexo	Niña	Niño	Total
Espectro autista	5	8	13
Hiperactividad	1	2	3
Total	6	10	16

Una vez que la tabla de contingencia está hecha se lleva a cabo la siguiente fórmula para calcular la probabilidad, la cual debe ser menor a 0.05 para indicar dependencia entre las variables:

$$P = \frac{(A + B)! (C + D)! (A + C)! (B + D)!}{N! A! B! C! D!}$$

Siguiendo el procedimiento, se sustituye la nomenclatura por los números correspondientes a las frecuencias correspondientes a cada celda de la tabla:

$$= 13! 3! 6! 10! / 16! 5! 8! 1! 2!$$

$$= 0.482$$

Así, para el ejemplo presentado, se obtiene el valor de $P = 0.482$, comprobando que hay dependencia entre las variables del trastorno del desarrollo y el sexo, es decir, se rechaza la hipótesis nula y se acepta la alternativa (Fernández y Díaz, 2004).

Muestras correlacionadas: prueba de McNemar

Por otro lado, cuando se quiere conocer un cambio en una variable correlacionada, como puede ser el caso de los grupos que derivan de la comparación del antes y después de cierto evento, se utiliza la prueba de McNemar (Coronel-Carvajal, 2020).

Para ejemplificar lo anterior, se expondrá el caso en el que se busca conocer el efecto de un entrenamiento conductual dirigido a mejorar la capacidad de sostener la atención en 18 niños con trastorno de déficit de atención e hiperactividad, donde se espera que, como efecto del tratamiento, exista una mejoría en el mantenimiento de la atención por lo que se recogerán datos previos y posteriores a esta intervención de la cantidad de niños con la capacidad de mantener la atención. La hipótesis nula es que habrá independencia entre las variables no observándose un cambio en el mantenimiento de la atención después del tratamiento, mientras que la hipótesis alternativa es que si habrá dependencia de las variables reflejada en un aumento en la proporción de niños capaces de mantener la atención posterior al tratamiento.

En la Tabla 15 se observan los datos obtenidos tanto previamente al tratamiento como después, donde se registró en cada celda la cantidad de niños capaces de sostener la atención en las diferentes condiciones:

Tabla 15. Datos de las variables dicotómicas cualitativas obtenidos previo y posterior al tratamiento.

		Después		
		Sí atención	No atención	Total
Antes	Sí atención	3	1	4
	No atención	11	3	14
Total		14	4	18

A priori, se podría considerar que sí hay un aumento en la proporción de niños capaces mantener la atención después del tratamiento (celda C: 11 niños, respecto a las demás celdas) sin embargo, se lleva a cabo la prueba de McNemar con el objetivo de conocer si esta diferencia es estadísticamente significativa (Pembury Smith y Ruxton, 2020). Con los datos de la Tabla 15 se desarrolló la siguiente fórmula para calcular el valor del estadístico:

$$X_{MN}^2 = \frac{(|B - C| - 1)^2}{B + C}$$

Siguiendo el procedimiento de la fórmula, se sustituyen las letras por la cantidad de la celda correspondiente:

$$= (|11-1|-1)^2 / 11+1$$

$$= (9)^2 / 12$$

$$= 81 / 12$$

$$= 6.75$$

Así, el valor obtenido de la prueba de McNemar es de 6.75. Una vez obtenido este valor se debe comprobar que sea mayor al valor crítico en una

tabla de McNemar/Chi² considerando que siempre se considerará 1 grado de libertad para este análisis (Lachenbruch, 2014).

Por ejemplo, con probabilidad de 0.05 y un grado de libertad, el valor crítico es de 3.84. Tomando este valor, se compara con el obtenido en la prueba (6.75) para verificar que sea mayor al valor crítico obtenido (3.84), de ser así se considera que hay relación entre las variables con una significancia de 0.05; se rechaza la hipótesis nula. Confirmando en el ejemplo mencionado que la mayor proporción de niños con mejorías en el mantenimiento de la atención tras el tratamiento conductual es significativa es decir, se acepta la hipótesis alternativa (Fernández y Díaz, 2004).

Otros tipos de pruebas no paramétricas

No siempre se utiliza a la estadística inferencial para comparar las diferencias entre dos conjuntos de datos. En ocasiones, nuestro objetivo puede ser el analizar si al comparar un conjunto de datos con otro, se encuentran relacionados. Estos análisis de correlación se usan cuando tenemos dos diferentes datos de un mismo sujeto y buscamos evaluar la relación entre estos. Un ejemplo sería comparar la relación entre la edad y la comprensión lectora en niños de 7 a 12 años. Además, existen pruebas estadísticas que nos ayudan a confirmar nuestras conclusiones y evaluar si un tratamiento vale la pena ser aplicado. Estas pruebas se llaman de tamaño del efecto y son complementarias a cada tipo de prueba estadística aplicada.

Correlación de Spearman

El término de correlación hace referencia al grado de relación de dos variables o tratamientos. La identificación más directa de este concepto es numérica, en donde se busca relacionar dos conjuntos de datos: si se obtiene una correlación positiva indica que a mayores valores de un grupo (variable), mayores valores del otro. En contraste, para la correlación negativa a menores valores de un grupo, mayores del otro. Por último, cuando no hay relación entre los conjuntos de datos se considera una correlación nula.

Una medida de correlación conocida es la correlación de Pearson, la cual se usa en el caso de datos paramétricos. Sin embargo, dentro de la estadística no paramétrica, existe el coeficiente de correlación de rangos de Spearman, basada en los rangos de números ordenados (Martínez Ortega et al., 2009). Para ejemplificar y desarrollar el uso de esta prueba estadística, se explorará la relación entre dos conjuntos de datos cualitativos. Por un lado, la variable “Dinámica familiar” con los niveles muy mala (1), mala (2), regular baja (3), regular alta (4), buena (5) y muy buena(6), y por otro la variable “Nivel socioeconómico” con los niveles muy bajo (1), bajo (2), regular bajo (3), regular alto (4), alto (5) y muy alto (6). En la Tabla 16 se encuentran los datos de ambas variables recolectados de 6 estudiantes.

Tabla 16. Registro de datos de 6 estudiantes en la variable de Dinámica familiar y Nivel socioeconómico.

Estudiante	Dinámica familiar	Nivel socioeconómico
1	regular baja	alto
2	mala	muy bajo
3	regular mala	regular alto
4	buena	regular bajo
5	muy mala	bajo
6	muy buena	muy alto

Una vez que se han obtenido los datos, se codifican los conceptos a números siguiendo la nomenclatura, donde muy malo equivale a 1, muy bueno a 6 y los intermedios a sus respectivos valores; muy bajo a 1, muy alto a 6 y los intermedios a sus respectivos valores. Posteriormente, estos valores se ordenan del menor al mayor y se coloca el número de lugar correspondiente, comenzando con 1 y terminando en el número equivalente a la muestra (6 para este ejemplo). En este caso, la numeración se mantiene en relación con la nomenclatura (ya que hay números del 1 al 6 sin repeticiones), sin embargo, si el número menor obtenido hubiera sido 3, se convierte en 1 por ser el de menor orden, y así sucesivamente hasta que el número mayor corresponda con el total de la muestra evaluada. Después, se agrega una cuarta columna ("d") que corresponderá a la diferencia entre el dato de la primera variable respecto a la segunda, y una quinta columna ("d cuadrada") en donde se coloca el cuadrado de la columna anterior, así como la suma de los valores cuadrados en la última fila (Tabla 17) (Martínez Ortega et al., 2009).

Tabla 17. Datos de las dos variables de interés codificadas, diferencia entre rangos y diferencia al cuadrado.

Estudiante	Dinámica familiar	Nivel socioeconómico	d	d cuadrada
1	3	5	-2	4
2	2	1	1	1
3	4	4	0	0
4	5	3	2	4
5	1	2	-1	1
6	6	6	0	0
				Σ 10

Una vez teniendo los datos codificados y ordenados de esta manera se procede a realizar la operación siguiendo la fórmula:

$$r_s = 1 - \frac{6\sum_i d_i^2}{n(n^2 - 1)}$$

Donde d_i a la diferencia entre las dos variables y n equivale al número de observaciones en cada variable.

$$= 1 - (6 * 10) / (6(36 - 1))$$

$$= 1 - 60 / 210$$

$$= 1 - .27$$

$$= 0.71$$

Por lo tanto, el resultado de correlación de Spearman en este ejemplo de es 0.72, la cual es una correlación alta, es decir, una alta relación entre ambas variables (tabla 18). En otras palabras, con estos datos hipotéticos se podría inferir que a mejor dinámica familiar, mejor nivel socioeconómico.

Tabla 18. Magnitudes de los valores de correlación.

Valor de correlación	Magnitud de asociación
0.0 - 0.26	Baja
0.27 - 0.45	Media baja
0.46 - 0.55	Media baja
0.56 - 0.70	Media alta
0.71 - 1	Alta

Tamaño del efecto

Cuando usamos modelos estadísticos por lo general comparamos las diferencias o los cambios entre dos o más conjuntos de datos. Esta comparación se basa en que tan probable es (a un valor de probabilidad preestablecido que suele ser menor de 0.05) que las diferencias entre los conjuntos de datos se deban al azar o si efectivamente hay un efecto de las variables o tratamientos experimentales. Sin embargo, no es suficiente con solo conocer si existen estas diferencias, también hay que conocer que tan importante es el efecto de las variables evaluadas. Por ejemplo, después de administrar un curso de comprensión lectora se observa que las calificaciones sobre 100 de los alumnos subieron 0.02 puntos (de 81 a 81.02). Si bien, hay una diferencia avalada estadísticamente, el efecto es tan pequeño que puede no valer la pena tomar el curso. Así, una medida complementaria a la significancia que nos permite evaluar este concepto es el tamaño del efecto.

El tamaño del efecto es la magnitud del resultado, es decir, mediante esta medida se tiene un índice del tamaño de la diferencia evaluada. Similar al cálculo de la significancia, el tamaño del efecto se calcula mediante una prueba estadística, la cual dependerá del diseño y estadístico utilizados (Ledesma y

Macbeth, 2008). En este sentido, los posibles valores a obtener van a depender de la fórmula utilizada para calcular el tamaño del efecto, pero cada fórmula tiene puntos de corte delimitados para que su interpretación pueda ser equiparable entre distintas fórmulas. Así, la interpretación del resultado en el tamaño del efecto, independientemente del método seguido, caerá en uno de tres posibles niveles: efecto pequeño, efecto mediano o efecto grande (Dominguez-Lara, 2018).

A continuación, se presentan las fórmulas para calcular el tamaño del efecto cuando el estadístico usado es U de Mann-Whitney y cuando es una chi cuadrada.

Cuando es U de Mann-Whitney, se sigue la siguiente fórmula:

$$r = \frac{Z}{\sqrt{N}}$$

Donde Z equivale al valor crítico obtenido en el estadístico de la U de Mann-Whitney y N al total de la muestra ($N = n_1 + n_2$) (Dominguez-Lara, 2018).

Por otro lado, se puede usar la prueba Phi (φ) para calcular el tamaño del efecto al usar una chi cuadrada (Kim, 2017).

$$\varphi = \sqrt{(x^2 / n)}$$

Dónde:

- x^2 es el valor obtenido en la chi cuadrada
- n es el número total de observaciones.

Para las dos pruebas ejemplificadas, la magnitud del tamaño del efecto se podrá interpretar de manera similar (tabla 19), sin embargo, esto dependerá de los puntos de corte fijados por la prueba utilizada.

Tabla 19. Magnitudes del tamaño del efecto.

	<i>Pequeño</i>	<i>Mediano</i>	<i>Grande</i>
<i>r</i>	0.1	0.3	0.5
φ	0.1	0.3	0.5

Conclusión

Las pruebas no paramétricas ayudan a abordar el estudio de fenómenos subjetivos del comportamiento que resultan de interés para el educador especial. Componentes como la motivación, las emociones o incluso el interés de los estudiantes se pueden analizar e interpretar de una manera estructurada y objetiva. Para esto es necesario distinguir las características de los datos que se recolectan y el contexto en el que fueron recabados. Se tiene que considerar si los datos provienen de un mismo sujeto medido en diferentes situaciones (grupos correlacionados) y aplicar pruebas con la T de Wilcoxon o si bien, el objetivo es comparar dos grupos de sujetos que fueron sometidos a diferentes tratamientos experimentales (grupos independientes) y por tanto es más apropiado aplicar una U de Mann-Whitney. También están los casos cuando se buscan comparar más de dos grupos experimentales, en cuyo caso nos apoyamos en pruebas como la H de Kruskal-Wallis. Las pruebas no paramétricas no solo sirven para comparar los resultados de diferentes grupos, también pueden establecer la relación entre dos variables medidas del mismo grupo mediante la correlación de Spearman o establecer el tamaño del efecto el

cual sirve como referencia sobre la viabilidad en la aplicación de dichos tratamientos experimentales.

El conocer estas pruebas, proporciona al educador especial herramientas invaluableles en el ejercicio de su profesión y permite abordar desafíos derivados del desarrollo de nuevas metodologías de intervención-plan educativo de una manera estructurada y objetiva. La información presentada en este capítulo es breve, pero tiene como objetivo fomentar el interés y servir como punto de referencia para el desarrollo de los conocimientos del lector en la aplicación de este tipo de pruebas en su práctica profesional.

Referencias bibliográficas

- Bautista-Díaz, M. L., Victoria-Rodríguez, E., Vargas-Estrella, L. B., y Hernández-Chamosa, C. C. (2020). Pruebas estadísticas paramétricas y no paramétricas: su clasificación, objetivos y características. *Educación y Salud. Boletín Científico Instituto de Ciencias de La Salud Universidad Autónoma Del Estado de Hidalgo*, 9(17), 78–81. <https://doi.org/10.29057/icsa.v9i17.6293>
- Coronel-Carvajal, C. (2020). *Forma correcta de presentar los datos y uso de McNemar en las intervenciones educativas*. 24(1).
- Daniel, W. W. (1991). *Bioestadística. Base para el análisis de las ciencias de la salud*. En *Bioestadística: Base para el análisis de las ciencias de la salud (3ra ed.)*. Editorial Limusa
- Fernández, P., y Díaz, P. (2004). *Asociación de variables cualitativas: Test X²*. 11, 304-308.
- Flores-Ruiz, E., Miranda-Novales, G. y Villasís-Keever, M. (2017). Metodología de la investigación. *Rev Alerg Mex*, 64(3). <http://www.revistaalergia.mx>
- Freeman, J. V., y Campbell, M. J. (2007). *The analysis of categorical data: Fisher's exact test*. 3.

- Gómez-Gómez, M., Danglot-Banck, C. y Vega-Franco, L. (2003). Sinopsis de pruebas estadísticas. Cuándo usarlas. *Revista Mexicana de pediatría*, 70, 91-99.
- Hernández-Sampieri, R., Fernández-Collado, C., y Baptista-Lucio, P. (2008). *Metodología de la investigación (4ta ed.)*. McGraw-Hill.
- Johnson, R., y Kuby, Patricia. (2008). *Estadística elemental: Lo esencial (10a ed.)*. Cengage Learning.
- Kirk, R. E. (2008). *STATISTICS An Introduction (5ta ed.)*.
- Lachenbruch, P. A. (2014). *McNemar test*.
- Martínez Ortega, Péndas, L. C. T., Martínez Ortega, M. M., Abreu, A. P., y Cánovas, A. M. (2009). El coeficiente de correlación de los rangos de Spearman caracterización. 8, 2.
- Mendenhall, W., Beaver, R.J., y Beaver, B.M. (2010). *Introducción a la probabilidad y estadística (13a ed.)*. Cengage Learning. <http://latinoamerica.cengage.com>
- Pembury Smith, M. Q. R., y Ruxton, G. D. (2020). Effective use of the McNemar test. *Behavioral Ecology and Sociobiology*, 74(11), 133. <https://doi.org/10.1007/s00265-020-02916-y>
- Rumsey, D. J. (2013). *Estadística para dummies (2da ed.)*. Planeta.
- Siegel, S., y Castellan, N. J. (1995). *Estadística no paramétrica, aplicada a las ciencias de la conducta*. (4a ed.). Trillas.

USO DE PRUEBAS PARAMÉTRICAS DE SIGNIFICACIÓN ESTADÍSTICA EN LA EDUCACIÓN ESPECIAL

Manuel Alejandro Cruz Aguilar

Enrique Hernández Arteaga [‡]

Anders Ågmo

Resumen

El educador especial es un profesionalista orientado al estudio y la intervención de personas con discapacidad, trastornos del neurodesarrollo y/o aptitudes sobresalientes. Como profesionalista, es menester que desarrolle métodos de intervención basados en la evidencia científica, la cual puede ser abordada desde el enfoque cuantitativo. En este sentido, la estadística es la herramienta matemática que permite el análisis objetivo, que puede ser empleada en el contexto de la investigación en educación o bien, como un coadyuvante en la interpretación de resultados de investigaciones científicas. Es por esta razón, que el presente capítulo tiene como objetivo discutir la interpretación, su posible aplicación y uso en el contexto de la investigación en educación especial, de las pruebas de significación estadística paramétricas, como lo son la t de Student y el Análisis de Varianza (ANOVA). Por lo que, se pretende que el educador especial sea capaz de identificar cuándo se habla de una diferencia significativa en un artículo de investigación, así como la relevancia de dicho dato estadístico; para que así pueda obtener sus propias conclusiones al respecto. De tal manera que, se interese en desarrollar más investigación relacionada

[‡] Facultad de Ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Carretera Tlaxcala-Puebla Km. 1.5. Tlaxcala, Tlax. eharteaga@uatx.mx

con la educación especial, o bien, pueda plantear métodos de intervención educativa basados en la evidencia científica.

Palabras clave: Estadística paramétrica, t de Student, ANOVA.

Introducción

La educación especial es una multidisciplina que se ocupa del estudio y la intervención de los procesos de enseñanza-aprendizaje orientados al logro de aspectos cognitivos, actitudinales y procedimentales de los sujetos con necesidades educativas especiales, tales como individuos con discapacidad, trastornos del neurodesarrollo y con aptitudes sobresalientes (Sánchez Palomino, 2011).

Partiendo desde este enfoque multidisciplinario de la educación especial, se puede afirmar que el quehacer diario del educador especial debería estar enfocado en dos tareas principales: la intervención educativa en la población con necesidades educativas especiales y la investigación para brindar una intervención basada en aspectos científicos (Acle Tomasini, 2011).

En este contexto, como ya se ha mencionado en capítulos anteriores, la investigación se realiza para poder describir un fenómeno. Por su parte, la investigación puede ser cualitativa (cuando se busca describir un fenómeno a partir de descripciones detalladas de contextos, personas o sucesos), cuantitativa (cuando se busca describir un fenómeno a partir de la objetividad de datos numéricos y la comprobación de hipótesis), o mixta (cuando se emplean métodos que contemplen ambos tipos de análisis) (Hernández Sampieri, et al., 2010).

Haciendo particular énfasis en el enfoque cuantitativo, es de observarse que el razonamiento detrás de éste implica, la obtención de datos numéricos, la prueba de hipótesis, la extrapolación y la interpretación a partir de los números. Es entonces, que la estadística se convierte en el mecanismo de análisis para los estudios cuantitativos (Bunge, 2017; Molina y Rodrigo, 2009).

La estadística, se divide en estadística descriptiva (se encarga de organizar y procesar los datos, agrupándolos en tablas de frecuencia o calculando las medidas de posición, medidas de tendencia central y medidas de dispersión), e inferencial (se encarga de tomar decisiones a partir de una muestra y estimar los parámetros descriptivos de una población a partir de una muestra representativa, mediante el cálculo de intervalos de confianza, la correlación y las pruebas de significación estadística) (Runyon y Haber, 1992; Daniel, 1996).

Particularmente, las pruebas de significación estadística sirven para comparar muestras, para encontrar diferencias significativas entre ellas, esto a partir del contraste de dos hipótesis estadísticas (nula y alternativa). Si las muestras cumplen una serie de criterios (que serán descritos posteriormente), se realizarán pruebas paramétricas; mientras que, si no, se tendrán que realizar pruebas no paramétricas (Gómez-Biedma, et al., 2001).

Cuando se realizan estos procedimientos estadísticos, al profesional en educación especial le proporciona las competencias necesarias para desarrollar el pensamiento analítico y crítico, le permite realizar estudios en los que se ponen a prueba conjeturas que se plantea, para buscar la respuesta a preguntas que le surjan, por ejemplo ¿cuáles son los mejores diagnósticos en la población con discapacidad, trastornos del neurodesarrollo o aptitudes sobresalientes?, ¿cuál es la mejor manera de intervenir educativamente en tal población? Así mismo le permite entender las publicaciones (notas de prensa, artículos en

revistas especializadas, informes de investigación, etc.) acerca de temas relacionados con la educación especial, que son las que en el futuro van a permitirle especializarse y mejorar su desempeño profesional (Molina y Rodrigo, 2009).

Es entonces que el presente capítulo abordará el uso de las pruebas de significación paramétricas. Debido a la complejidad en la interpretación, primeramente, se analizarán los criterios que se deben de cumplir para poder realizar pruebas estadísticas de este tipo. Posteriormente, se realizará una descripción de las pruebas estadísticas paramétricas “t de Student” y “ANOVA”, mencionando ejemplos de su aplicación. Para finalizar con las pruebas de comparaciones múltiples, haciendo principal hincapié en la pertinencia de su uso, así como un ejemplo de su aplicación. De tal manera, que con este capítulo el lector tenga un panorama general de las pruebas de significación estadística paramétricas, su interpretación, su posible aplicación y uso en el contexto de la investigación en educación especial.

Criterios para realizar una prueba estadística paramétrica

Como ya se ha mencionado en capítulos anteriores, la estadística inferencial es una rama de la estadística cuyos cálculos y procedimientos se realizan con la finalidad de establecer inferencias estadísticas a partir de una muestra, es decir, estimar los parámetros poblacionales a partir de datos muestrales. En el caso de que, los parámetros que se buscan inferir sean medias y varianzas, se deben realizar pruebas paramétricas, las cuales están fundamentadas en la distribución normal. Para que estas inferencias tengan sentido y no se cometa un error de interpretación, es necesario que se realicen

sólo cuando los datos cumplan con una serie de criterios. Los criterios para poder hacer pruebas paramétricas son: que se realice un muestreo probabilístico, que haya validez interna en la recolección de datos, que la variable dependiente se mida en una escala de intervalos o de razón, que los datos provengan de una población con distribución normal y que los datos tengan homocedasticidad (Runyon y Haber, 1992; Daniel, 1996; Dagnino, 2014a). Si bien, varios de estos criterios ya han sido descritos en otros capítulos, se retoman de manera general a continuación con la intención de que queden más claros los conceptos, y se parte de un ejemplo en particular aplicado para la educación especial:

Muestreo probabilístico: cuando se realice el muestreo se debe optar por obtener una muestra con la técnica probabilística, la cual permite conocer la probabilidad que tiene cada individuo para ser incluido en la muestra a través de una selección al azar; las técnicas de muestreo probabilístico más empleadas son aleatorio simple, aleatorio estratificado, aleatorio sistemático, por conglomerados, etc. (Otzen y Manterola, 2017).

Supóngase que un educador especial quiere realizar un plan de intervención educativa para implementar acciones de “vida independiente”, para que niños con discapacidad intelectual aprendan a recoger sus platos después de comer. Para ello necesita probar que su plan de intervención educativa podría funcionar para diferentes niños. Por lo que decide realizar un muestreo probabilístico, de tipo estratificado. En este sentido, visita los 35 Centros de Atención Múltiple (CAM) que hay en el estado de Tlaxcala, y de cada CAM selecciona de manera aleatoria a 2 niños, por lo que tiene una muestra de 70 niños con los cuales trabajar.

Validez interna: se debe diseñar un experimento tal que posea validez interna, ésta hace referencia a las características que debe contener un estudio

cuantitativo para que el investigador siga el principio de causalidad. Es por esta razón, que deben evitarse al máximo las variables extrañas, es decir, aquellos factores que pudieran confundir los resultados y que pudieran llevar al investigador a proponer una conclusión equivocada. Entre los factores más comunes que pueden alterar la validez interna se encuentran la historia del participante (aquellos acontecimientos que pudieran influir sobre las respuestas que se evalúen), la selección (cuando se eligen grupos que no son comparables o muestras demasiado pequeñas), la maduración (cambios que ocurren de manera natural), los efectos del pretest (influencia del pretest, ya que hay pruebas que inducen el aprendizaje y que, cuando los participantes vuelven a realizarlas, las ejecutan de una mejor manera), la difusión (los participantes del experimento intercambian información que puede afectar los resultados, dicha información puede transmitirse por la interacción, si hablan o se observan entre sí mientras realizan la prueba) (Mayorga-Ponce, et al., 2021).

Retomando el ejemplo anterior, el educador especial debería pensar en que su plan de intervención educativa, es como el “experimento” que se quiere realizar con los niños con discapacidad intelectual, de tal manera que, debe diseñarlo considerando cuidadosamente las variables extrañas, por ejemplo, pensando en la historia de vida que tenga el participante con relación a la costumbre de recoger los platos después de comer, además considerar tener un pretest y un post test que no involucre el aprendizaje de la prueba en sí misma, que se trabaje de manera individual con los participantes, etc.

Escala de intervalos o de razón: Estos tipos de escala representan magnitudes, con la propiedad de igualdad de la distancia entre puntos de escala de la misma amplitud, por lo que solamente se puede obtener en mediciones cuantitativas (numéricas) de los datos. La diferencia entre ambos tipos de

escalas es que en la escala de intervalo el valor cero de la escala no es absoluto, sino un cero arbitrario, por ejemplo: mediciones realizadas con pruebas psicométricas. Mientras que en el caso de la escala de razón, ésta posee el cero absoluto, el cual representa la ausencia total de la magnitud que se está midiendo, por ejemplo: longitud, peso, distancia, ingresos, precios, tiempo en segundos y frecuencias (Orlandoni Merli, 2010).

Retomando el ejemplo que se ha considerado en el capítulo, para que el educador especial realice una prueba estadística paramétrica, la variable a medir podría ser el tiempo que tarda el niño con discapacidad intelectual en recoger sus platos o el número de veces que lo realiza en una semana. Este tipo de mediciones, corresponderían a una variable dependiente con una escala de tipo de razón que permitirían al educador especial comparar el antes de la intervención educativa contra el después, de tal manera que se evidencie la efectividad de dicho plan de intervención.

Distribución normal: Antes de continuar, resulta necesario hacer la acotación de la “normalidad” a la que se referirá este capítulo, para no causar una confusión en dicho término; ya que, en el ámbito de la educación especial, algunas personas hacen referencia a “normalidad” cuando un niño no presenta alguna discapacidad ni trastorno, es decir, presenta un desarrollo típico. Entonces, en estadística este concepto hace referencia a la distribución de probabilidad propuesta por el matemático Carl Friederich Gauss.

La distribución normal es un constructo teórico que se utiliza en estadística para modelar y analizar datos continuos. En muchos casos, esta distribución se aproxima a la distribución de la muestra obtenida de una población aleatoria. Esto permite calcular la probabilidad de obtener ciertos valores en una muestra y hacer inferencias sobre la población a partir de la

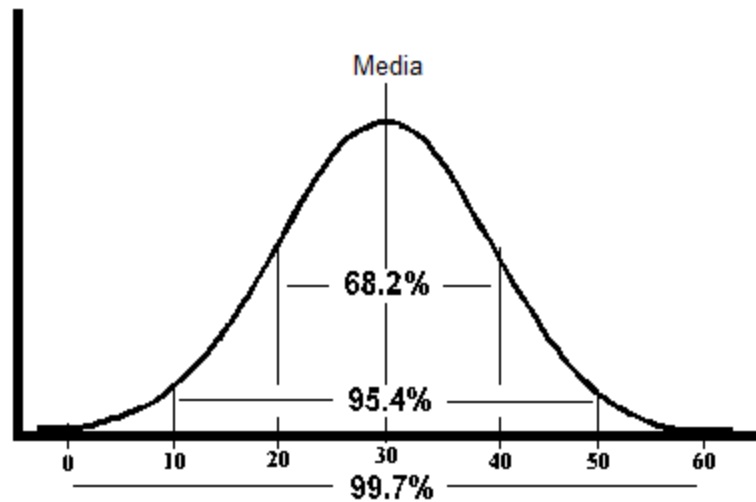
muestra. Si bien, esta distribución de probabilidad es teórica, afortunadamente muchos fenómenos biológicos, psicológicos y sociales se aproximan a esta distribución. De manera general, dicha distribución de probabilidades tiene la forma de la campana de Gauss, es simétrica, alcanza su máximo en la media (la media tiene el mismo valor que la moda y la mediana), tiene una distribución 68.2%, 95.4% y 99.7%, es decir, el 68.2% de la población se distribuye a una desviación estándar de la media, el 95.4% a dos desviaciones estándar y el 99.7% a tres desviaciones estándar (Dagnino, 2014b; Daniel, 1996; Runyon y Haber, 1992).

Para dejar más claro este punto, supóngase que la figura 1 representa la distribución de probabilidades para el tiempo (en segundos) que tardan los niños con discapacidad intelectual para recoger los platos, posterior a un plan de atención-intervención educativa propuesto por un educador especial. En este ejemplo, hipotéticamente se toman a toda la población del mundo de niños con discapacidad intelectual y se les aplicará dicho plan de intervención educativa. Si se les calcula la media, supóngase que se obtiene un valor de 30 segundos; así mismo, si se les calcula la desviación estándar, se obtendría un valor de 10 segundos. En este sentido, la gráfica representa cuántas personas existen en la población, que obtuvieron esos determinados puntajes. En dicha gráfica, se observa también que tiene forma de campana de Gauss, es simétrica, la media es el punto más alto de la campana ya que corresponde a la mediana y la moda.

Además, si consideramos la media de tiempo que tardan los niños con discapacidad intelectual en recoger sus platos, podemos observar que la mayoría de ellos (aproximadamente el 68.2%) tardarán entre 20 y 40 segundos. Esto se debe a que, si restamos y sumamos la desviación estándar a la media (que es de 10 segundos), obtenemos los valores de 20 ($30 - 10 = 20$) y 40 ($30 + 10 = 40$)

segundos, respectivamente. En otras palabras, la desviación estándar nos indica cuánto se alejan los datos individuales de la media, y al considerar un rango de ± 1 desviación estándar, podemos estimar la probabilidad de que los datos estén dentro de ese rango, que en este caso es del 68.2%.

Figura 1. Características de la Distribución Normal en el caso hipotético de estudio.



Nota: Se muestran los valores hipotéticos del ejemplo de los niños con discapacidad intelectual. La media hipotética es de 30 y la desviación estándar hipotética es de 10. En el eje de las X, se encuentran los valores hipotéticos de dispersión, mientras que los porcentajes corresponden a la probabilidad de encontrar esos valores en los datos.

Fuente: Elaboración propia.

Como ya se mencionó, la distribución normal es un constructo teórico, ya que es necesario contar con toda la población para poder construir esta representación gráfica. Lo cual, no es humanamente posible, ya que como se ha revisado en este y en otros capítulos, en estadística inferencial se recurre a obtener una muestra representativa de la población, para poder inferir los parámetros en una población a partir de dicha muestra.

En este contexto, existen pruebas estadísticas, que permiten determinar (con un cierto porcentaje de probabilidad), si la muestra que se consiguió en el estudio proviene o no de una población con una distribución normal, lo cual es bastante importante, ya que las pruebas paramétricas son muy sensibles a las desviaciones de la normalidad, por lo que, si no se satisface el criterio de normalidad y se realiza una prueba paramétrica, se podría obtener un error estadístico que conllevaría a obtener interpretaciones erróneas de los datos (Yarkoni, 2022).

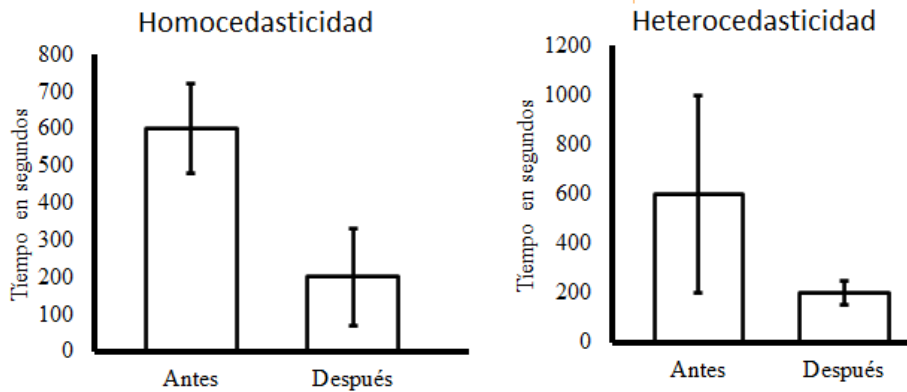
Entre las pruebas más comunes para determinar la normalidad de los datos, se encuentran las siguientes pruebas: Kolmogorov-Smirnov, Lilliefors, Anderson-Darling, la prueba de bondad de ajuste χ^2 , Cramer-Von Mises, Shapiro-Wilk, entre otros. Sin embargo, se ha descrito que esta última prueba parece tener una mayor potencia estadística respecto a las otras (Flores-Muñoz, et al., 2019). Para fines prácticos de este capítulo, no se abordarán dichas pruebas, sin embargo, debe de suponerse que en todos los ejemplos que se mencionen, se utilizó al menos alguna de estas pruebas para determinar la normalidad de los datos.

Homocedasticidad de los datos: Como ya se ha mencionado en capítulos anteriores, la variabilidad es una característica inherente a todas las variables, debido a que normalmente se estudian fenómenos que son muy complejos. Dicha variabilidad ha sido objeto de estudio de muchas disciplinas; a manera de ejemplo, se observa que los biólogos están interesados en las variaciones fisiológicas, ecológicas y evolutivas de los seres vivos; los psicólogos se interesan por las variaciones comportamentales de los individuos; por su parte, los educadores especiales se interesan por la diversidad en las características que pueden tener los usuarios con trastornos del neurodesarrollo,

con discapacidades o con aptitudes sobresalientes. En este sentido, una manera de medir la variabilidad es mediante el cálculo de la varianza, la cual cuantifica el “alejamiento” de los datos con respecto a la media (mientras más alejados, los datos presentan más diferencia entre sí, la variabilidad es mayor y, por lo tanto, la varianza es mayor) (Daniel, 1996; Runyon y Haber, 1992; Oddi, et al., 2020). Considerando lo anterior, la homocedasticidad hace referencia a cuando la varianza es semejante en los grupos evaluados, mientras que cuando ésta no es semejante, se le conoce como heterocedasticidad. Habría que tener especial cuidado con que las muestras no presenten heterocedasticidad, ya que esto involucra un problema en la interpretación, ya que los valores p , no son lo que corresponden, los errores estándar no son correctos y los estadísticos no siguen la distribución normal (Oddi, et al., 2020; Yarkoni, 2022).

En la figura 2, se ejemplifica usando gráficas de barras, en el caso hipotético del que se ha mencionado; donde la barra representa el valor de la media, mientras que las barras del error estándar representan la varianza dividida entre la raíz cuadrada del número de sujetos que conforman la muestra (es decir, es una medida obtenida a partir del valor de la varianza y a esta medida se le conoce como error estándar). En este sentido, supóngase que se evalúa el tiempo (en segundos) que tardan los niños con discapacidad intelectual para recoger los platos, antes y después de un plan de atención-intervención educativa propuesto por un educador especial. En el caso de la gráfica de la izquierda, se observa que las variaciones (representadas por la barra del error estándar), presentan homocedasticidad (la varianza permanece constante en ambos casos), mientras que, en la gráfica derecha, las variaciones presentan heterocedasticidad (la varianza es diferente en ambos casos).

Figura 2. Ejemplos de homocedasticidad y heterocedasticidad en el caso hipotético de estudio



Fuente: Elaboración propia.

Muy semejante a lo que ocurre con el criterio de normalidad, es necesario recurrir a pruebas estadísticas para determinar la presencia de homocedasticidad/heterocedasticidad en los datos. Entre estas pruebas se encuentra la prueba de Levene y la de Breush-Pagan, la cual se ha descrito como una de las más eficaces para dicho fin (Tafalla, 2014). Sin embargo, para fines prácticos de este capítulo, no se abordará cómo se realiza dicha prueba.

Si no existe dicha homocedasticidad, el investigador puede cometer errores en la interpretación si realiza pruebas paramétricas de significación. Entonces, una vez que se verifica el cumplimiento de dichos criterios, el siguiente paso es determinar cuál es la prueba estadística idónea para analizar los datos, las cuales se describen a continuación.

Pruebas paramétricas de significación estadística

Para determinar cuál es la prueba estadística que se empleará, se necesita considerar el número de variables independientes y el tipo de diseño estadístico que se planteó.

Variable independiente: También llamada tratamiento, es la variable que el investigador “manipula” esperando que tenga un efecto sobre una variable dependiente (medición) (Villasís-Keever y Miranda-Novales, 2016). Retomando el ejemplo anterior (evaluación del plan de atención-intervención educativa para niños con discapacidad intelectual), vemos claramente que el educador especial proporciona un “tratamiento” a su muestra de niños, esperando que tengan un efecto sobre una variable dependiente (medición del tiempo que tardan los niños con discapacidad intelectual en recoger sus platos después de comer).

Dependiendo del número de variables independientes que se manipulan en un estudio, los diseños experimentales se clasifican en simples (se manipula una sola variable independiente) y factoriales (se manipulan dos o más variables independientes) (Nuñez-Peña y Bono, 2020). Para fines didácticos, este capítulo sólo abordará los diseños simples.

Diseños estadísticos: estos hacen referencia a la manera en que serán analizados estadísticamente los datos. De manera general, corresponden a la manera en que fueron obtenidos los datos, y se clasifican en tres tipos (Nuñez-Peña y Bono, 2020):

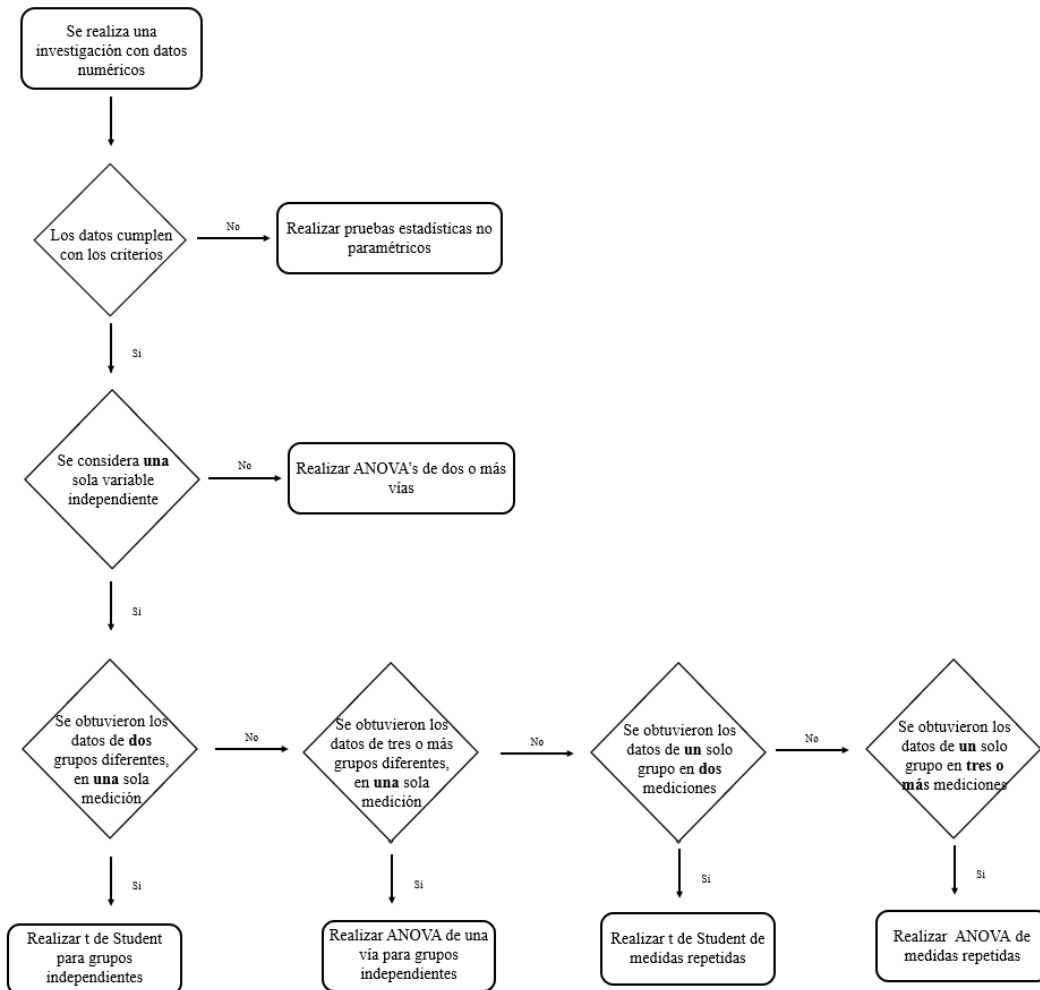
- *Diseños de grupos al azar:* Los sujetos son divididos en dos o más grupos de manera aleatoria, siendo uno de ellos el grupo de control y el otro u otros los grupos experimentales. Se procede a medir la variable

dependiente en un momento determinado. Por ejemplo, supóngase que un educador especial quiere conocer si los niños con trastorno del espectro autista (TEA) grado 3 de severidad, pueden realizar tareas de aprendizaje discriminativo (ya que tiene una propuesta de plan de atención-intervención que, para que funcione, los niños con TEA deben tener aprendizaje discriminativo). En este sentido, realiza un estudio en el que compara las curvas de aprendizaje (variable dependiente) entre niños control (neurotípicos) y niños diagnosticados con TEA grado 3.

- *Diseño de grupos homogéneos:* Los grupos experimentales no son formados mediante una aleatorización pura, sino mediante la introducción de ciertas restricciones con el fin de lograr homogeneidad entre ellos con respecto a cierta característica de interés. Posteriormente, se procede a medir la variable dependiente en un momento determinado. Por ejemplo, supóngase que un educador especial quiere probar si los niños que tienen discapacidad visual y saben Braille, muestran menores niveles de ansiedad que niños que no saben Braille. Por lo tanto, se toma una muestra aleatoria de niños con discapacidad visual y se clasifican en función de si saben este tipo de lectura o no. Posteriormente se evalúa sus niveles de ansiedad, empleando el inventario CMARS-2 (variable dependiente).
- *Diseño de medidas repetidas:* Es aquel en el que el investigador registra dos o más medidas de la variable dependiente para cada sujeto. Por lo general, se realiza una medición antes de aplicar un tratamiento y otra posterior a éste. En este tipo de diseño, el sujeto se convierte en su propio control. Un ejemplo de este diseño es el que ya se ha planteado recurrentemente en este capítulo.

Entonces, considerando precisamente el número de variables independientes, y el diseño estadístico, se observa que, a partir de la medición de la variable dependiente, se pueden obtener varios grupos de datos. De manera general, cuando se decide analizar dos grupos de datos, se emplea la prueba estadística de “t de Student”, mientras que, si se analizarán tres o más grupos de datos, se empleará una prueba de “ANOVA”. En la figura 3, se muestra un diagrama de flujo para la toma de decisión acerca de la prueba estadística que se elegirá para analizar los datos.

Figura 3. Diagrama de flujo en la toma de decisión de la prueba estadística que se utilizará.



Fuente: Elaboración propia.

t de Student

La prueba de *t* de Student fue propuesta por el químico y matemático William Sealy Gosset (1876-1937), bajo el pseudónimo de “estudiante” (*Student* en inglés). Dicha prueba fue pensada para examinar la diferencia entre dos muestras provenientes de grupos independientes, que tengan normalidad y que pueden ser pequeñas (menores a 30 sujetos por grupo) (Sánchez-Turcios, 2015).

Para la realización de una prueba de *t* de Student, es necesario seguir estos tres pasos (Runyon y Haber, 1992):

1. *Formular las hipótesis estadísticas (nula y alternativa)*: En este sentido, de manera general, la hipótesis nula plantea la igualdad entre las medias de las poblaciones de donde provinieron los datos, mientras que la alternativa plantea la diferencia. Normalmente, se realiza la prueba de significación estadística para estimar la probabilidad de que ocurra la igualdad entre las medias (hipótesis nula), de tal manera que si esta probabilidad es muy baja, el investigador pueda suponer que la hipótesis nula es falsa, y aceptar a la hipótesis alternativa.
2. *Calcular el valor de t de Student*: Se sustituyen los valores de una fórmula y se resuelve dicha fórmula.
3. *Compararlo con los valores en tablas para encontrar la región crítica*: De manera consensuada, el valor *p* máximo permitido para rechazar la hipótesis nula (sin cometer un error estadístico), es de 0.05; siendo los más comunes el de 0.05 y 0.01. Este valor de *p* representa el nivel de significación estadística, es decir, la probabilidad máxima de error que se tiene al rechazar la hipótesis nula cuando esta es verdadera. En este sentido, dependiendo los valores de significación que se buscan, existen tablas que determinan el valor mínimo que debe tener la *t* calculada para

poder rechazar la hipótesis nula. Por ejemplo, si se reporta un valor de $p \leq 0.05$, esto implica que el investigador tiene un 5% o menos de probabilidad de cometer error al rechazar la hipótesis nula (igualdad entre dos grupos de datos), lo que es lo mismo, tiene una certeza del 95% o más de aceptar la hipótesis alternativa (diferencia entre las poblaciones de donde provienen los datos). En este contexto, la tabla 1, muestran algunos de los valores mínimos que se deben obtener en el cálculo de la t de Student para rechazar la hipótesis, sólo considerado para un valor de $p \leq 0.05$ y $p \leq 0.01$. Habitualmente, cuando se realiza la prueba se usa programas computacionales, los cuales mencionan los valores de p exactos, sin embargo, cuando se reportan estos valores en los artículos científicos, deben ir acompañados del valor de t que se calculó y de los grados de libertad correspondientes, para que el lector pueda verificar en la tabla, la correspondencia entre ambos valores (t y p).

Tabla 1. Valores críticos de t para una prueba bilateral.

Grados de libertad	0.05	0.01
1	12.706	63.657
2	4.303	9.925
3	3.182	5.841
4	2.776	4.604
5	2.571	4.032
6	2.447	3.707
7	2.365	3.499
8	2.306	3.355
9	2.262	3.250
10	2.228	3.169
11	2.201	3.106
12	2.179	3.055
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921
17	2.110	2.898
18	2.101	2.878
19	2.093	2.861
20	2.086	2.845

Nota: La t obtenida es significativa a un nivel dado si es igual o mayor que el valor indicado en la tabla. La tabla es ilustrativa y solo tiene los datos de t para los grados de libertad del 1 al 20, y los niveles de significación de 0.05 y 0.01. Para revisar la tabla completa, verificarla de Runyon y Haber (1992, p. 373), de donde han sido tomados los datos. La tabla original muestra los valores de t correspondientes a varios niveles de probabilidad, para cualquier número dado de grados de libertad.

A continuación, se ejemplifican estos tres pasos, tanto para un diseño de grupos independientes, como para uno de medidas repetidas, empleando en ambos casos, el uso de la prueba de significación estadística “t de Student”.

t de Student para grupos independientes

En un CAM, hay 10 niños que tienen discapacidad motora a raíz de una distrofia muscular progresiva. Estos niños son intervenidos por dos educadores especiales, cada uno atiende a 5 niños, en un lapso de un mes. En la propuesta del educador especial 1, se tienen una intervención basada en estimulación psicomotriz, mientras que en la propuesta del educador especial 2, la estimulación psicomotriz no está contemplada. El educador especial 1 decide probar que la estimulación psicomotriz funciona mejor para que los niños mejoren la coordinación mano-ojo. En este sentido, utiliza una prueba que permite cuantificar el nivel de este proceso (en un rango de 0 a 20 puntos, siendo los puntajes mayores los que representan una mejor coordinación), con pruebas en las que observa la ejecución motora de la misma. Realiza pruebas de normalidad y heterocedasticidad y se da cuenta que los datos cumplen estos requisitos.

Paso 1: El educador especial 1 plantea las siguientes hipótesis estadísticas:

Hipótesis nula: Ambos grupos de niños proceden de poblaciones con promedios similares de ejecución motriz ($G1 = G2$) y las diferencias sólo se deben al azar. Por lo tanto: $H_0: G1=G2$.

Hipótesis alternativa: Los dos grupos provienen de poblaciones distintas ($M1 \neq M2$) y las diferencias en ejecución motriz son probablemente reales,

causadas posiblemente por la estimulación psicomotriz. Por lo tanto: $H_a: G1 \neq G2$.

Paso 2: El cálculo de la t de Student para grupos independientes se realiza con la siguiente fórmula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Donde:

- \bar{X}_1 = Promedio del grupo 1
- \bar{X}_2 = Promedio del grupo 2

- $$SC = \sum_{i=1}^n X^2 - \frac{\left(\sum_{i=1}^n X \right)^2}{n}$$

- Σx = Sumatoria de los datos
- Σx^2 = Sumatoria de los datos al cuadrado
- n = tamaño muestral (número de datos por grupo)

Entonces, para poder resolver la fórmula, el educador especial agrupó los datos en la tabla 2 (columna de “Calificación de prueba”). Calculó el promedio por grupo. Así mismo, elevó cada calificación al cuadrado y sumó los resultados para obtener las sumatorias. A simple vista, se observa que los promedios son diferentes, pero es necesario realizar la prueba de significación estadística para demostrar que esas diferencias no fueron por el azar que hizo que tuviéramos la “suerte” de evaluar justamente a esos niños en particular.

Tabla 2. Datos hipotéticos de las calificaciones de dos grupos de niños con discapacidad motora en una prueba hipotética de coordinación mano-ojo.

Grupo 1 (Experimental)			Grupo 2 (Control)		
	$\bar{x}_1 = 11.8$			$\bar{x}_2 = 7.0$	
Niños	Calificación de la prueba	Calificación al cuadrado	Niños	Calificación de la prueba	Calificación al cuadrado
Niño 1	11	121	Niño 1	10	100
Niño 2	8	64	Niño 2	8	64
Niño 3	12	144	Niño 3	5	25
Niño 4	13	169	Niño 4	3	9
Niño 5	15	225	Niño 5	9	81
Sumatoria	$\Sigma x = 59$	$\Sigma x^2 = 723$	Sumatoria	$\Sigma x = 35$	$\Sigma x^2 = 279$

Fuente: Elaboración propia.

Entonces, una vez que el educador especial reunió esos datos en la tabla, procede a sustituir los valores en las fórmulas (Runyon y Haber, 1992):

$$Sc_1 = 723 - \frac{(59)^2}{5} = 26.8 \quad Sc_2 = 279 - \frac{(35)^2}{5} = 34$$

$$t = \frac{11.8 - 7.0}{\sqrt{\frac{26.8 + 34}{5 + 5 - 2} \left(\frac{1}{5} + \frac{1}{5} \right)}}$$

Al resolver la fórmula, se obtiene un valor de $t = 2.752$.

Paso 3: En primer lugar, se calculan los grados de libertad con los que se realizó el estudio. En este caso, al número total de los datos, se le restan 2, ya que, por definición la t de Student trabaja con un grado de libertad igual a N-2 (Runyon y Haber, 1992). Entonces, si el número total de niños son 10, los grados de libertad serían: $10 - 2 = 8$.

Una vez calculados los grados de libertad, se busca en la tabla 1 la región crítica. Se busca en la columna de los grados de libertad, y se observa que en esa fila $t \geq 2.306$, corresponde a una $p \leq 0.05$; mientras que $t \geq 3.355$, corresponde a una $p \leq 0.01$. En nuestro caso, obtuvimos una $t = 2.752$, por lo que, al ser un valor mayor del que muestra la tabla en la región crítica de $p \leq 0.05$, podemos proponer que existe menos del 5% de probabilidad de que la hipótesis nula sea verdadera y por lo tanto la podemos rechazar para aceptar la hipótesis alternativa. Esto en otras palabras, significa que efectivamente la calificación de los grupos es diferente como producto de la estimulación psicomotriz y no del azar, con una probabilidad de máximo 5% de equivocarnos al afirmar esta aseveración.

Al conocer que los grupos sí presentan estas diferencias significativas, verificamos con los promedios que el grupo experimental (niños con estimulación psicomotriz) es mayor que el promedio del grupo control, por lo que podemos concluir que dicha intervención es eficaz.

Un ejemplo de un estudio real es el reportado por Mosso-Salas y Zaragoza-Toscano (2017). En dicho estudio, evaluaron con el Inventario de Estrategias de Afrontamiento (IEA) a cuidadores primarios de niños con parálisis cerebral leve o moderada ($n = 31$) y cuidadores primarios de niños con parálisis cerebral grave ($n = 29$). Todos los niños eran atendidos en el Centro de Rehabilitación e Inclusión Infantil del Teletón (CRIT), unidad Michoacán. Encontraron que los cuidadores de niños con parálisis cerebral grave (comparados con los cuidadores de niños con parálisis cerebral leve o moderado) mostraron menores puntajes de las variables de expresión emocional ($t = 2.44$, $p = 0.017$) y pensamiento desiderativo ($t = 3.12$, $p=0.003$) del IEA. Con estos resultados, los autores sugieren que el nivel de adaptación de los

progenitores de hijos con discapacidad puede estar relacionado con la gravedad e independencia funcional de la discapacidad. Entonces, este estudio pone de manifiesto que, el análisis estadístico empleando la *t* de Student para grupos independientes permite destacar la importancia de atender de una manera integral al usuario con discapacidad, pensando en un abordaje psicoterapéutico que permitan la expresión adecuada de emociones y la aceptación de la discapacidad.

t de Student de medidas repetidas

En una Unidad de Servicios de Apoyo a la Educación Regular (USAER), hay 5 niños que tienen Trastorno Específico del Lenguaje (TEL). Estos niños son intervenidos por un educador especial, quien les hace un examen diagnóstico de lecto-comprensión. Posterior a un mes de intervención educativa, vuelve a evaluar a los niños con un examen de lecto-comprensión. En ambos casos, los exámenes tienen un puntaje máximo de 20 puntos. Suponga que el educador especial realizó las pruebas de normalidad y homocedasticidad correspondientes.

Paso 1: El educador especial 1 plantea las siguientes hipótesis estadísticas:

Hipótesis nula: Posterior al programa de intervención educativa, los niños con TEL mostrarán puntajes semejantes entre la prueba diagnóstica y la prueba final, como resultado del azar. Por lo tanto: $H_0: Pre=Post$.

Hipótesis alternativa: Posterior al programa de intervención educativa, los niños mostrarán puntajes diferentes entre la prueba diagnóstica y la prueba

final, que no se deberán al azar, sino a la intervención educativa. Por lo tanto:

Ha: Pre ≠ Post.

Paso 2: El cálculo de la t de Student para grupos independientes se realiza con la siguiente fórmula (Runyon y Haber, 1992):

$$t = \frac{MD}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{n}}{n(n-1)}}$$

Donde:

- MD = Promedio de las diferencias
- $\sum D$ = Sumatoria de las diferencias
- $\sum D^2$ = Sumatoria de las diferencias al cuadrado
- n = tamaño muestral

Entonces, para poder resolver la fórmula, el educador especial agrupó los datos en la tabla 3. Calculó las diferencias entre la prueba final y diagnóstica, restando el valor de una respecto de la otra. Así mismo, elevó cada diferencia al cuadrado y sumó los resultados para obtener las sumatorias. De manera similar, calculó el promedio de las diferencias.

Tabla 3. Datos hipotéticos de las calificaciones de dos grupos de niños con Trastorno Específico del Lenguaje en una prueba hipotética de lecto-comprensión.

Niños	Calificación de la prueba diagnóstica	Calificación de la prueba final	Diferencia de calificaciones MD = -3.2	Cuadrado de las diferencias
Niño 1	7	10	-3	9
Niño 2	3	6	-3	9
Niño 3	12	13	-1	1
Niño 4	5	16	-11	121
Niño 5	14	12	2	4
Sumatoria			$\Sigma D = -16$	$\Sigma D^2 = 144$

Fuente: Elaboración propia.

Entonces, una vez que el educador especial reunió esos datos en la tabla, procede a sustituir los valores en la fórmula:

$$t = \frac{-3.2}{\sqrt{\frac{144 - \frac{(-16)^2}{5}}{5(5 - 1)}}$$

Al resolver la fórmula, se obtiene un valor de $t = -1.485$

Paso 3: En primer lugar, se calculan los grados de libertad con los que se realizó el estudio. En este caso a N, al número total de pares de datos (p. Ej. pre – post), se le resta 1, ya que, por definición la t de Student de medidas repetidas trabaja con un grado de libertad igual a N-1 (Runyon y Haber, 1992). Entonces, si el número total de niños son 5, los grados de libertad serían $5 - 1 = 4$.

Una vez calculados los grados de libertad, se busca en la tabla 1 la región crítica. Se busca en la columna de los grados de libertad, y se observa que en esa fila $t \geq 2.776$, corresponde a una $p \leq 0.05$; mientras que a $t \geq 4.604$, corresponde a una $p \leq 0.01$. En nuestro caso, obtuvimos una $t = -1.485$, por lo que el valor absoluto de la t calculada, al ser un valor menor del que muestra la tabla en la región crítica de $p \leq 0.05$, no nos permite rechazar la hipótesis nula

y tenemos que aceptarla. Esto en otras palabras significa que, aunque a simple vista pareciera que sí hay un aumento en los puntajes de la prueba posterior a la intervención, estos incrementos no fueron significativos ya que se debieron principalmente al azar y no a la intervención educativa. Por lo que podríamos concluir que la intervención educativa propuesta por el educador especial no está funcionando y debería refinarla para poder obtener mejores resultados.

Un ejemplo de un estudio real es el reportado por Reyes-López (2023) quien, en su tesis de Licenciatura en Educación Especial, evaluó el efecto de un plan de atención-intervención educativa, utilizando el Inventario de Habilidades Básicas a 6 niños de edad preescolar, todos diagnosticados con Trastorno del Espectro Autista (TEA). Realizó una evaluación pretratamiento, en la que identificó las áreas de oportunidad que tenía para trabajar con sus niños. Una vez identificado este aspecto, realizó el plan de atención que involucraba actividades para realizar en el aula y actividades por parte de los padres. Posterior a la realización de este plan, realizó otra evaluación con el mismo instrumento. Cabe aclarar que dicho instrumento no genera aprendizaje en el niño evaluado y se basa en la observación de ciertas habilidades, las cuales se engloban en áreas: básica, comunicativa, personal-social y psicomotricidad. A su vez, cada área tiene sub-áreas. Particularmente, una de las áreas que estaban afectadas en estos niños fue la de psicomotricidad. En la cual, tras el plan de atención-intervención, la educadora especial encontró que los niños mejoraban en las sub-áreas de motricidad fina ($t=-3.19$, $p=0.049$) y motricidad gruesa ($t=-7.57$, $p=0.004$). Entonces, este estudio pone de manifiesto que, dicho plan de atención-intervención educativa es eficaz, y que incluso, dependiendo los contextos, pudiera emplearse un plan semejante para niños con semejantes problemas psicomotores.

Como se observa, la prueba “t de Student” está pensada para comparar dos grupos de datos, ya sean éstos independientes o sean resultado de medidas repetidas. Sin embargo, existen diferentes estudios en los que se plantean más de dos grupos de datos. Entonces, la pregunta obvia sería: ¿se puede usar repetidas veces la t de Student, comparando cada par de grupo de datos? De hecho, esta pregunta no es trivial, ya que en bastante de la literatura se suele cometer este error. En este sentido, como ya se mencionó anteriormente, la prueba de t de Student trabaja bajo supuestos de probabilidad de error para rechazar la hipótesis nula (valor de p). Esto es correcto, siempre y cuando sólo haya dos grupos para comparar. Sin embargo, en el caso de estudios en los que hay más de dos grupos de datos, esta probabilidad (valor de p), es diferente (Dagnino, 2014c; Runyon y Haber, 1992). Por lo que, algunos autores optan por emplear una corrección del valor de p , llamada “la corrección de Bonferroni”, aunque lo correcto sería elaborar una prueba estadística llamada ANOVA. A continuación, en el capítulo se discutirá acerca de los posibles usos de ambas.

Análisis de Varianza (ANOVA)

El ANOVA (por el término en inglés *Analysis of Variance*, “Análisis de Varianza”), es un conjunto de técnicas estadísticas diseñadas por el biólogo y estadístico Ronald Aylmer Fisher (1890-1962). En el ANOVA, la variable independiente (factor) es la que determina los grupos del estudio; mientras que el número de grupos definido por un factor se conoce como el número de niveles del factor. Entonces, el ANOVA permite evaluar la variabilidad en una variable dependiente, medida en circunstancias definidas por la variable independiente (que contiene más de un nivel de clasificación) (Dagnino, 2014a; Runyon y Haber, 1996).

Existen estudios que solo tienen un factor o variable independiente, o hay otros que tienen dos o más variables independientes. Por ejemplo:

Estudios con una sola variable independiente o factor

- Cuando hay más de dos grupos que necesitan ser comparados. Por ejemplo, si se quisiera evaluar la frecuencia de aleteo (estereotipia conductual) en niños con trastorno del espectro autista (TEA), comparando los niveles de severidad (grado 1 vs 2 vs 3). El factor en este estudio es “el grado de severidad del TEA”.
- Cuando hay más de dos ocasiones en que se realizaron mediciones repetidas. Por ejemplo, si se quisiera evaluar los niveles de atención de niños con trastorno por déficit de atención (TDA), antes, durante y después de haber tomado metilfenidato. El factor en este estudio es “el momento de tomar metilfenidato”.

Estudios con más de una variable independiente o factor

- Cuando se desea analizar simultáneamente el efecto de dos tratamientos diferentes, evaluando el efecto de cada uno por separado y su posible interacción. Por ejemplo, si se quisiera evaluar el efecto conjunto de la administración de atomoxetina y métodos conductuales para disminuir la hiperactividad de niños con TDAH. Los factores en este estudio son “la administración de atomoxetina” y “la administración de métodos conductuales”.
- Cuando hay dos o más grupos en quienes se realizaron mediciones repetidas en dos o más ocasiones. Por ejemplo, si se quisiera evaluar la ansiedad de niños con discapacidad visual y niños con discapacidad auditiva, antes y después de una intervención educativa que promueva la

inclusión de la discapacidad en el aula regular. Los factores en este estudio son “el tipo de discapacidad” y “el momento en que se evalúa”.

Considerando lo anterior, cuando el estudio solamente tiene un factor se emplea un ANOVA de una vía. Mientras que, si el estudio tiene más de un factor, se busca la interacción entre estos, y se emplea un ANOVA de dos (o más) vías (Dagnino, 2014a).

Para fines prácticos, en este capítulo se abordará solamente el tema del ANOVA de una vía, ya que este tipo de ANOVA suele ser más común (respecto al ANOVA de dos o más vías) en los estudios de las ciencias relacionadas a la Educación Especial. Es por esta razón, que, a partir de este momento, cada vez que se refiera la palabra “ANOVA”, el lector deberá contemplar solamente al ANOVA de una vía.

Entonces, al igual que otras pruebas de significación estadística, el análisis de varianza (ANOVA) se basa en la comparación de una hipótesis nula con una hipótesis alternativa. La hipótesis nula establece que la variabilidad de cada grupo es similar a la variabilidad global, y que las medias de los grupos no difieren entre sí. Por otro lado, la hipótesis alternativa plantea que la variabilidad global es diferente de la variabilidad de cada grupo, debido al efecto del factor que se estudia, lo que implica que al menos una de las medias de los grupos es diferente a las demás (Dagnino, 2014a).

Para poder determinar estas diferencias en la variabilidad, se emplea el análisis de varianza, donde se descompone la variabilidad total (varianza total) en dos variabilidades: la variabilidad producida por el tratamiento (también llamado variabilidad entre grupos) y la variabilidad residual (también llamada variabilidad dentro de grupos o del error). Posteriormente, se realiza el cálculo

del estadístico F de Fisher a partir de la siguiente fórmula (Dagnino, 2014a; Runyon y Haber, 1992):

$$F = \frac{\textit{Variabilidad del tratamiento}}{\textit{Variabilidad residual}}$$

Para calcular dichas varianzas, se utilizan unas tablas, denominadas “tablas de ANOVA” (tabla 4), donde se especifican las fuentes de variación, los grados de libertad, las sumas de cuadrados, las medias de cuadrados (varianzas), y la F calculada.

Es importante destacar que, si el valor de la F calculada es menor que 1, significa que la variabilidad del error es mayor que la del tratamiento. En cambio, si el valor de la F calculada es mayor que 1, la variabilidad del tratamiento es mayor que la del error en el experimento (Dagnino, 2014a). Una vez que se calcula la F de Fisher, al igual que la t de Student, se contrasta el valor calculado de la F contra un valor de tablas, en la que se muestra el valor mínimo para poder rechazar la hipótesis nula. En este contexto, en la tabla 5, se muestran algunos de los valores mínimos que se deben obtener en el cálculo de la F de Fisher para rechazar la hipótesis nula, sólo considerado para un valor de $p \leq 0.05$. De manera semejante a las demás pruebas de significación estadística, cuando se realiza la prueba se usan programas computacionales, los cuales mencionan los valores de p exactos, sin embargo, cuando se reportan estos valores en los artículos científicos suelen ir acompañados del valor de F que se calculó para que el lector pueda verificar en la tabla, la correspondencia entre ambos valores (F y p).

Tabla 4. Tabla de ANOVA, en la que se especifican las fórmulas que se realizan en cada celda.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F de Fisher
Entre grupos (tratamiento)	$SC_t = \sum \frac{(\sum xi)^2}{Ni} - \frac{(\sum x_{tot})^2}{N_{tot}}$			
	Donde:	$gl_t = k-1$		
	$\sum x_i$ = Sumatoria de los datos en cada grupo			
	N_i = número de datos en cada grupo	Donde:	$MC_t = \frac{SC_t}{gl_t}$	$F = \frac{MC_t}{MC_e}$
	$\sum x_{tot}$ = Sumatoria de todos los datos	k = número de grupos		
	N_{tot} = número de datos totales			
Dentro de grupos (error o residual)	$SC_e = \sum x_i^2 - \frac{(\sum xi)^2}{Ni}$	$gl_e = N-k$		
	Donde:			
	$\sum x_i$ = Sumatoria de los datos en cada grupo	Donde:	$MC_e = \frac{SC_e}{gl_e}$	
	N_i = número de datos en cada grupo	N = número total de datos k = número de grupos		
	$\sum x_i^2$ = Sumatoria de los datos elevados al cuadrado en cada grupo			
Total	$SC_{tot} = SC_t + SC_e$			

Fuente: Elaboración propia. Las fórmulas fueron tomadas de Runyon y Haber, (1992).

Tabla 5. Valores críticos de F.

		<i>Grados de libertad del tratamiento</i>									
		1	2	3	4	5	6	7	8	9	10
<i>Grados de libertad del error</i>	1	161	200	216	255	230	234	237	239	241	242
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.23	3.18
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97

Nota: La F obtenida es significativa a un nivel de $p \leq 0.05$, si es igual o mayor que el valor indicado en la tabla. La tabla es ilustrativa y sólo tiene los datos de F para los grados de libertad del 1 al 10, y el nivel de significación de 0.05. Para revisar la tabla completa, ver Runyon y Haber (1992, pp. 374-379), de donde han sido tomados los datos. La tabla original muestra los valores de F correspondientes a varios niveles de probabilidad (0.01 y 0.05), para cualquier número dado de grados de libertad.

Para ejemplificar el ANOVA, relacionado con la Educación Especial, se revisará el siguiente caso hipotético de estudio. Suponga que un educador especial evalúa la memoria de trabajo, usando los cubos de Corsi, a 12 niños diagnosticados con Trastorno del Espectro Autista, todos asisten regularmente a un CAM. Estos niños, son clasificados en tres grupos, dependiendo el grado de severidad (de acuerdo con el DSM-V), teniendo cuatro niños con grado uno, cuatro con grado dos y los últimos cuatro con grado tres. A estos datos de los

niños, se les realizaron las pruebas de normalidad y homocedasticidad, concluyendo que sí se pueden realizar pruebas paramétricas.

Planteamiento formal de las hipótesis estadísticas: El educador especial postula las siguientes dos hipótesis:

Hipótesis nula: Todos los grupos de niños proceden de poblaciones con promedios similares de ejecución en la tarea de memoria de trabajo, sin importar el grado de severidad y las diferencias sólo se deben al azar. Por lo tanto: $H_0: G1=G2=G3$.

Hipótesis alternativa: Los grupos provienen de poblaciones distintas y las diferencias en ejecución de la tarea de memoria de trabajo son probablemente reales, como resultado del diferente grado de severidad, siendo al menos un grupo de niños diferente del resto. Por lo tanto: $H_a: G1 \neq G2 \neq G3$.

En la tabla 6 se muestran los datos obtenidos por el educador especial, quien decidió reportar el número de respuestas correctas de un total de 20 ensayos de la ejecución de cubos de Corsi. Posteriormente, dichos datos fueron elevados al cuadrado.

Tabla 6. Datos hipotéticos del número de respuestas correctas de tres grupos de niños con TEA en la prueba de cubos de Corsi.

	Niños	Calificación de la prueba	Calificación al cuadrado
Grado 1	Niño 1	15	225
	Niño 2	14	196
	Niño 3	14	196
	Niño 4	13	169
	Sumatoria	$\Sigma x = 56$	$\Sigma x^2 = 786$
Grado 2	Niño 1	12	144
	Niño 2	12	144
	Niño 3	11	121
	Niño 4	10	100
	Sumatoria	$\Sigma x = 45$	$\Sigma x^2 = 509$
Grado 3	Niño 1	5	25
	Niño 2	9	81
	Niño 3	6	36
	Niño 4	5	25
	Sumatoria	$\Sigma x = 25$	$\Sigma x^2 = 167$
Sumatoria total: 126			

Fuente: Elaboración propia.

Cálculo de la suma de cuadrados del tratamiento: Una vez que se calculan las sumatorias de los datos por cada grupo y se tiene la sumatoria total, se procede a sustituir los valores en la fórmula. Habrá que recordar que el símbolo Σ representa una suma, por lo que, en la sustitución de la fórmula se incorpora la suma correspondiente (revisar tabla 6 para verificar los datos).

$$SC_t = \sum \frac{(\sum x_i)^2}{N_i} - \frac{(\sum x_{tot})^2}{N_{tot}}$$

$$SC_t = \frac{(56)^2}{4} + \frac{(45)^2}{4} + \frac{(25)^2}{4} - \frac{(126)^2}{12} = 784 + 506.25 + 156.25 - 1323$$

$$SC_t = 123.5$$

Cálculo de la suma de cuadrados del error: Para poder realizar este procedimiento, se deberá primeramente restar el valor de la sumatoria de cuadrados (revisar tabla 6 para verificar los datos) menos su valor respectivo de la sumatoria por grupo al cuadrado, entre el número de datos por grupo. Una vez realizado este procedimiento, se suman todos los resultados.

$$SC_e = \sum x_i^2 - \frac{(\sum x_i)^2}{N_i}$$

$$SC_e = (786 - 784) + (509 - 506.25) + (167 - 156.25)$$

$$SC_e = 15.5$$

Cálculo de los grados de libertad: Se deben de calcular los grados de libertad tanto para el tratamiento como para el error. En el caso de los grados de libertad del tratamiento, al número de grupos se le resta uno. En este caso, como se cuenta con 3 grupos, los grados de libertad serían: $3-1=2$.

Mientras que, para el caso de los grados de libertad del error, al tamaño muestral total, se le restan el número de grupos. En este caso, como se cuenta con 12 niños divididos en 3 grupos, los grados de libertad serían: $12-3=9$.

Cálculo de la media de cuadrados: En ambos casos (Media de cuadrados del tratamiento y del error), se divide la suma de cuadrados respectiva, entre los grados de libertad.

En el caso de la media de cuadrados del tratamiento, se realiza el siguiente procedimiento:

$$MC_t = \frac{SC_t}{gl_t}$$
$$MC_t = \frac{123.5}{2} = 61.75$$

En el caso de la media de cuadrados del error, se realiza el siguiente procedimiento:

$$MC_e = \frac{SC_e}{gl_e}$$
$$MC_e = \frac{15.5}{9} = 1.72$$

Cálculo de la F de Fisher: Una vez calculadas las variabilidades producidas por el tratamiento y por el error, respectivamente, se divide la media de cuadrados del tratamiento entre la del error:

$$F = \frac{MC_t}{MC_e}$$
$$F = \frac{61.75}{1.72} = 35.9$$

Determinación de la región crítica: Al revisar la tabla de valores de F (tabla 5), justo en el cruce de los grados de libertad del error (en este caso 9), con los grados de libertad del tratamiento (en este caso 2), se observa que la tabla especifica un valor de $F = 4.26$. Como en nuestro caso, se obtuvo una $F = 35.9$, al ser este valor mayor que la F de tablas, se puede rechazar la hipótesis nula y por lo tanto, aceptar la hipótesis alternativa con una $p \leq 0.05$. Esto permite concluir al educador especial que, al menos uno de los grupos evaluados muestra una diferente memoria de trabajo, evaluada con los cubos de Corsi.

De manera general, cuando se realiza esta prueba con programas computacionales, suele obtenerse la tabla 7. En dicha tabla, se observan justamente los valores calculados paso a paso en el presente capítulo. Normalmente, se reporta el valor exacto de p .

Tabla 7. Ejemplo de Tabla de ANOVA.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F de Fisher	Valor de p
Entre grupos (tratamiento)	123.5	2	61.75	35.9	≤ 0.05
Dentro de grupos (error o residual)	15.5	9	1.72		
Total	139				

Fuente: Elaboración propia.

Las pruebas de comparaciones múltiples

Como se mencionó anteriormente, el ANOVA (cuando se obtiene un valor de $p \leq 0.05$) indica que al menos una de las medias es diferente de las otras, sin embargo, para conocer cuáles medias son diferentes entre sí, se proceden a realizar pruebas de comparaciones múltiples (Dagnino, 2014c; Fallas, 2012). Dichas pruebas están pensadas para ajustar el valor de p cuando se comparen dos grupos de datos en un diseño pensado para más de dos grupos de datos. La necesidad de este ajuste deriva de que cada vez que se hace una comparación de grupos de datos se está aceptando una probabilidad de error determinado (generalmente máximo 0.05); sin embargo, cada comparación va multiplicando este error (Dagnino, 2014c).

Entre las diferentes pruebas que existen, se encuentra la prueba de Bonferroni, de Tukey, de Dunnet, de Scheffé, de Student-Newman-Keuls, la diferencia mínima significativa irrestricta, la diferencia mínima significativa restringida de Fisher, la prueba de rangos múltiples de Duncan, entre otras (Dagnino, 2014c; Fallas, 2012). El elegir una de estas pruebas no es una tarea trivial, sino que dependerá de la situación, del objetivo y planificación del trabajo. En este contexto, el usar indiscriminadamente cualquiera de las pruebas de comparaciones múltiples produce un aumento de la probabilidad de cometer cualquiera de los dos errores estadísticos (Dagnino, 2014c):

- Error tipo I o α : rechazar la hipótesis nula cuando es verdadera, es decir, declarar como significativa una diferencia cuando en realidad los datos no permiten hacer tal afirmación,
- Error de tipo II o β : aceptar la hipótesis nula cuando es falsa, es decir, declarar no significativa la diferencia cuando en realidad sí lo es.

Considerando lo anterior, las pruebas más empleadas en ciencias aplicadas a la educación especial son la corrección de Bonferroni, la prueba de Tukey y la de Duncan.

La corrección de Bonferroni: Este método fue propuesto por el matemático Carlo Emilio Bonferroni (1892-1960), y consiste en compensar el aumento de la probabilidad de obtener significación por azar, disminuyendo el nivel de p requerido para dar por significativo el resultado. Para ello, se divide el valor de p por el número de comparaciones planeado (Dagnino, 2014c; Fallas, 2012). Por ejemplo, si se quiere comparar 3 grupos, las comparaciones serían 3: G1 vs G2, G1 vs G3 y G2 vs G3. Entonces el valor de 0.05 (máximo para poder afirmar que existe una diferencia significativa), se divide entre 3, obteniéndose así $0.05/3 = 0.016$. Por lo que, se tendría que obtener dicho valor para poder hablar de una diferencia significativa. El problema con esta prueba es que entre más comparaciones se tengan, el nivel de significación se vuelve muy estricto, por lo que puede aumentar la probabilidad de obtener un error tipo II.

La prueba HSD (diferencia honestamente significativa) de Tukey: Esta prueba fue propuesta por el químico y matemático John Wilder Tukey (1915-2000), y es aplicable sólo cuando las n 's son iguales en las muestras, permitiendo la comparación entre todos los pares de medias. Sería el procedimiento más potente y exacto para usar en la comparación múltiple y permite el cálculo de intervalos de confianza. Es por esta razón que es la prueba más empleada en la literatura (Dagnino, 2014c; Fallas, 2012). Una de las desventajas que posee, es que es una de las pruebas más conservadoras, es decir, detecta menos diferencias entre las medias ya que cuando esta diferencia es

pequeña, no es capaz la prueba de declararla una diferencia significativa (Fallas, 2012).

La prueba de rangos múltiples de Duncan: Esta prueba fue desarrollada por el estadístico David B. Duncan (1916-2006), se basa en acomodar las medias en orden, y posteriormente se ajusta la diferencia crítica considerando si las dos medias que se comparan son adyacentes o si existe otra media entre las dos medias que se están comparando. Este método suele controlar bastante bien la probabilidad de cometer un error tipo I, lo que en la práctica hace que esta prueba sea más consistente que otras pruebas de comparación múltiple. Sin embargo, una desventaja que tiene es que en varios casos suele ser muy liberal, es decir, detecta más diferencias, ya que las diferencias pequeñas las declara como significativas (Fallas, 2012).

Un ejemplo real del uso de estas pruebas en el contexto de la educación especial es el reportado por Castro-Durán, et al. (2016), quienes realizaron un perfil comparativo de calidad de vida de personas con discapacidad intelectual que asisten a centros de formación laboral. Para ello, aplicaron varios instrumentos, entre los cuales aplicaron la Escala de Evaluación Objetiva de Calidad de Vida. Las personas con discapacidad intelectual fueron clasificadas de acuerdo con el paradigma que realizan en el centro de formación laboral al que asisten; dichos centros realizan el paradigma clínico, social y en transición de la educación especial. Al analizar con un ANOVA los datos, los investigadores encontraron diferencias significativas en la calidad de vida ($F = 10.153$; $p \leq 0.05$). La prueba de Tukey demostró que los usuarios con discapacidad intelectual que asistían a centros de formación laboral con paradigma social mostraron una mayor calidad de vida comparado con los del

paradigma clínico y de transición, pero no había diferencias entre estos dos últimos.

Conclusiones

La educación especial es una multidisciplina que retoma elementos de las ciencias factuales (basadas en hechos verificables y sustentados en evidencia científica), para identificar, evaluar y atender a población con necesidades especiales (individuos con discapacidad, trastornos del neurodesarrollo o con aptitudes sobresalientes).

La obtención del conocimiento se basa en el método científico, el cual, desde el enfoque cuantitativo plantea la necesidad del escrutinio estadístico para obtener conclusiones objetivas. Es por esta razón que la estadística es la herramienta que permite a la educación especial el analizar los datos numéricos en una investigación.

Para poder obtener conclusiones objetivas, es necesario que se verifiquen una serie de criterios, los cuales permiten realizar pruebas de significación estadística paramétricas. En este sentido, es deber del educador especial verificar la pertinencia de realizar estas pruebas (basadas en el cumplimiento de estos criterios), cuando se desea realizar investigación; o bien, cuando se desea llevar a la práctica, asegurarse que los métodos de intervención estén basados en evidencia científica. Es decir, en caso de que dichos métodos hayan nacido como resultado del análisis paramétrico, es deber del educador especial verificar que en los artículos que se consulten, se describa el cumplimiento de estos criterios.

Por otra parte, cuando se desea comparar dos grupos de datos, se deberá emplear una *t* de Student. Mientras que, si se desean comparar más de dos grupos de datos, la prueba idónea será el ANOVA, seguido de una prueba de comparación múltiple (Bonferroni., Tukey o Duncan). En este contexto, es menester que el educador especial identifique cuando se habla de una diferencia significativa en un artículo de investigación ($p \leq 0.05$), para que así pueda obtener sus propias conclusiones. De tal manera que, pueda desarrollar más investigación al respecto, o bien pueda plantear métodos de intervención educativa basados en el escrutinio científico.

Referencias bibliográficas

- Acle Tomasini, G. (2011). ¿Por qué la educación especial es especial? *Revista electrónica de educación y familia*, 2(2), 5-18. https://www.fcdh.uatx.mx:80/media/integra2/numero_completo/numero_completo_Volumen_2_Numero_2_julio-diciembre_2011.pdf
- Bunge, M.A. (2017). El planteamiento científico. *Revista cubana de salud pública*, 43(3), 470-498. <http://scielo.sld.cu/pdf/rcsp/v43n3/spu16317.pdf>
- Castro-Durán, L., Cerda-Etchepare, G., Vallejos-Garcías, V., Zúñiga-Vásquez, D., y Cano-González, R. (2016). Calidad de vida de personas con discapacidad intelectual en centros de formación laboral. *Avances en Psicología Latinoamericana*, 34(1), 175-186. <https://revistas.urosario.edu.co/xml/799/79943294013/index.html>
- Dagnino, J. (2014a). Análisis de Varianza. *Revista Chilena de Anestesiología*, 43, 306-310. <https://revistachilenadeanestesia.cl/PII/revchilanestv43n04.07.pdf>
- Dagnino, J. (2014b). La distribución normal. *Revista Chilena de Anestesiología*, 43, 116-121. <https://revistachilenadeanestesia.cl/PII/revchilanestv43n02.08.pdf>
- Dagnino, J. (2014c). Comparaciones múltiples. *Revista Chilena de Anestesiología*, 43, 311-312. <https://revistachilenadeanestesia.cl/PII/revchilanestv43n04.08.pdf>

- Daniel, W. (1996). *Bioestadística: Base para el análisis de las ciencias de la salud*. Editorial Limusa.
- Fallas, J. (2012). Análisis de Varianza: comparando tres o más medias. [PDF]. <https://www.ucipfg.com/Repositorio/MGAP/MGAP-05/BLOQUE-ACADEMICO/Unidad-2/complementarias/analisis de varianza 2012.pdf>
- Flores-Muñoz, P., Muñoz Escobar, L., y Sánchez Acalo, T. (2019). Estudio de potencia de pruebas de normalidad usando distribuciones desconocidas con distintos niveles de no normalidad. *Perfiles*, 21(1), 4-11. <http://ceaa.esPOCH.edu.ec:8080/revista.perfiles/faces/Articulos/Perfiles21Art1.pdf>
- Gómez-Biedma, S., Vivó, M., y Soria, E. (2001). Pruebas de significación en Bioestadística. *Revista de Diagnóstico Biológico*, 50(4), 207-218. http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0034-79732001000400008&lng=es&tlng=es
- Hernández Sampieri, R., Fernández Collado, C., y Baptista Lucio, M.P. (2010). *Metodología de la Investigación, quinta edición*. Mc Graw Hill.
- Mayorga-Ponce, R.B., Plata-Balderas, D.Z., Martínez-Alamilla, A., y Salazar-Valdez, D. (2021). Validez interna. *Educación y Salud Boletín Científico Instituto de Ciencias de la Salud Universidad Autónoma del Estado de Hidalgo*, 9(18), 68-70. <https://repository.uaeh.edu.mx/revistas/index.php/ICSA/article/download/6554/7968/>
- Molina, J.G. y Rodrigo, M.F. (2009). Estadística descriptiva en psicología. En: J.G. Molina y M.F. Rodrigo. *Open Course Ware* (pp. 1- 22). Universidad de Valencia.
- Mosso-Salas, D.B. y Zaragoza-Toscano, R.T. (2017). Estrategias de afrontamiento y resiliencia en cuidadores primarios de niños y adolescentes con parálisis cerebral. *Uricha*, 14(34), 82-90. <http://www.revistauaricha.umich.mx/index.php/urp/article/view/138/157>
- Núñez-Peña, M.I., y Bono, R. (2020). *Diseños de investigación en psicología- Universitat de Barcelona*. [PDF].

[http://diposit.ub.edu/dspace/bitstream/2445/164539/1/Dise%C3%B1os%20de%20Investigacion Diapositivas%2019-20 FINAL.pdf](http://diposit.ub.edu/dspace/bitstream/2445/164539/1/Dise%C3%B1os%20de%20Investigacion%20Diapositivas%2019-20%20FINAL.pdf)

Oddi, F.J., Miguez, F.E., Benedetti, G.C., y Garibaldi, L.A. (2020). Cuando la variabilidad varía: Heterocedasticidad y funciones de varianza. *Ecología Austral*, 30, 438-453.

<https://rid.unrn.edu.ar/bitstream/20.500.12049/6396/1/438-453%20diciembre20-final.pdf>

Orlandoni Merli, G., (2010). Escalas de medición en Estadística. *Telos*, 12(2), 243-247.

<https://www.redalyc.org/pdf/993/99315569009.pdf>

Otzen, T., y Manterola, C. (2017). Técnicas de muestreo sobre una población a estudio. *International Journal of Morphology*, 35(1), 227-232.

<https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>

Reyes-López, T.T. (2023). Desarrollo de Habilidades básicas, Adaptativas y Sociales en Infantes de Preescolar con Trastorno del Espectro Autista. [Tesis de Licenciatura, Universidad Autónoma de Tlaxcala].

Runyon, R.P. y Haber, A. (1992). *Estadística para las ciencias sociales*. Addison-Wesley Iberoamericana.

Sánchez Palomino, A. (2011). Comentarios a: ¿Por qué la educación especial es especial? *Revista electrónica de educación y familia*, 2(2), 19-56.

https://www.fcdh.uatx.mx:80/media/integra2/numero_completo/numero_completo_Volumen_2_Numero_2_julio-diciembre_2011.pdf

Sánchez-Turcios, R.A. (2015). t-Student: Usos y abusos. *Revista Mexicana de Cardiología*, 26(1), 59-61. <https://www.scielo.org.mx/pdf/rmc/v26n1/v26n1a9.pdf>

Tafalla, A. (2014). Estudio de tamaño y potencia de algunos contrastes de Heterocedasticidad. [Tesis de licenciatura, Universidad de Zaragoza].

<https://zaguan.unizar.es/record/15689/files/TAZ-TFG-2014-1329.pdf>

Villasís-Keever, M.A. y Miranda-Navales, M.G. (2016). El protocolo de investigación IV: las variables de estudio. *Revista Alergia México*, 63(3), 303-310.

<https://www.redalyc.org/pdf/4867/486755025003.pdf>

Yarkoni, T. (2022). The generalizability crisis. *Behavioral and Brain Sciences*, 45.
<https://doi.org/10.1017/s0140525x20001685>

Autores(as) y co-autores(as) de la obra

Lic. Abril Zagnitte Gómez Méndez

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
zagnitte.gomez@alumnos.udg.mx

Dr. Anders Ågmo

Institute of Psychology. University of Trømsø. Trømsø, Noruega.
anders.agmo@uit.no

Dra. Andrea Saldívar Reyes

Facultad de ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. asaldivarr_fcdh@uatx.mx

Dra. Carolina Sotelo Tapia

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
csotelot@hotmail.com

Dr. Cid González

Tecnológico Nacional de México, campus Acayucan. Acayucan, Veracruz.
Centro Educativo New Hope. Mérida, Yucatán. cidramon@gmail.com

M. en C. Citlali López Esparza

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
citlalli.lopez2043@alumnos.udg.mx

Dra. Claudia Amezcua Gutiérrez

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
delcarmen.amezcua@academicos.udg.mx

Dra. Delia de Jesús Domínguez Morales

Facultad de ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. djdominguezm@uatx.mx

Dr. Edgar Oswaldo Zamora González

Departamento de Bienestar y Desarrollo Sustentable, Centro Universitario del Norte de la Universidad de Guadalajara. Colotlán, Jalisco.
edgar.zamora8148@academicos.udg.mx

Dr. Enrique Hernández Arteaga

Facultad de Ciencias para el Desarrollo Humano. Universidad Autónoma de Tlaxcala. Tlaxcala. eharteaga@uatx.mx

M. en C. Fabiola Alejandra Iribe Burgos

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco. fabiola.iriibe@alumnos.udg.mx

Dr. Francisco Javier Aguilar Guevara

Facultad de ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. fjaguilarg_fcdh@uatx.mx

Dr. Jahaziel Molina del Río

Laboratorio de Neuropsicología, Centro Universitario de los Valles, Universidad de Guadalajara. Ameca, Jalisco. jahaziel.mdelrio@academicos.udg.mx

Dr. Jorge Carlos Hevia Orozco

Escuela de Psicología, Universidad Anáhuac Mayab. Centro Educativo New Hope. Mérida, Yucatán. jorge.hevia@anahuac.mx

Dr. Josué Antonio Camacho Candia

Centro de Investigación e Innovación en Ciencias Aplicadas a la Educación Especial, Facultad de Ciencias para el Desarrollo Humano. Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. jacamachoc_fcdh@uatx.mx

Dr. Juan Pablo García Hernández

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco. ju.pablo93@hotmail.com

Dra. Luz Berenice López Hernández

Departamento de Ciclo de Vida, Unidad Académica de Ciencias de la Salud, Universidad Autónoma de Guadalajara. Zapopan, Jalisco. luzb.lopez@edu.uag.mx

Dra. M. del Consuelo Pedraza Aguilar

Unidad de Investigación en Neurodesarrollo “Dr. Augusto Fernández Guardiola”, Instituto de Neurobiología, Universidad Nacional Autónoma de México. Querétaro, Querétaro. mconsuelopa@gmail.com

Dr. Manuel Alejandro Cruz Aguilar

Instituto Nacional de Psiquiatría “Ramón de la Fuente Muñiz”. Ciudad de México, México. macrag@gmail.com

Dra. Marai Pérez Hernández

Laboratorio de Neurociencias Cognitivas, Departamento de Bienestar y Desarrollo Sustentable, Centro Universitario del Norte de la Universidad de Guadalajara. Colotlán, Jalisco. marai.perez@academicos.udg.mx

Lic. María Isabel Solís Meza

Doctorado en Ciencias Aplicadas a la Educación Especial. Facultad de Ciencias para el Desarrollo Humano. Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. isasolm3@gmail.com

Dra. Marisela Hernández González

Instituto de Neurociencias, Centro Universitario de Ciencias Biológica y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco. marisela.hgonzalez@academicos.udg.mx

Dr. Miguel Ángel Guevara

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
miguel.guevara@academicos.udg.mx

Mtra. Myriam Nayeli Villafuerte Vega

Facultad de ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala. mnvillafuert@uatx.mx

Dr. Pedro Manuel Cortes Esparza

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.
pedro.cortes@alumnos.udg.mx

Dra. Rosa María Hidalgo Aguirre

Laboratorio de Neuropsicología, Centro Universitario de los Valles, Universidad de Guadalajara. Ameca, Jalisco.
rosa.hidalgo@academicos.udg.mx

Dra. Susana Angelica Castro Chavira

Unidad de Investigación en Neurodesarrollo “Dr. Augusto Fernández Guardiola”, Instituto de Neurobiología, Universidad Nacional Autónoma de México. Querétaro, Querétaro. castrochavirasa@inb.unam.mx

Comité de Revisión y Dictaminación Técnica y Científica de la Obra

Dra. Andrea Cristina Medina Fragoso

Instituto de Neurobiología, Universidad Nacional Autónoma de México, campus Juriquilla. Querétaro, Querétaro.

Dr. Daniel Zarabozo Enríquez de Rivera

Instituto de Neurociencias, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara. Guadalajara, Jalisco.

Dra. Esther Olvera Cortés

Centro de Investigación Biomédica de Michoacán, Instituto Mexicano del Seguro Social. Morelia, Michoacán.

Dr. Felipe Cabrera González

Centro Universitario de la Ciénega, Universidad de Guadalajara. Ocotlán, Jalisco.

Dr. Luis Quintanar Rojas

Facultad de Ciencias para el Desarrollo Humano, Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala.

Dra. Marcela Arteaga Silva

División de Ciencias Biológicas y de la Salud. Universidad Autónoma Metropolitana, unidad Iztapalapa. Ciudad de México, México.

Dr. Óscar González Flores

Centro de Investigación en Reproducción Animal. Universidad Autónoma de Tlaxcala. Panotla, Tlaxcala.

Dra. Yolanda Cruz Gómez

Centro Tlaxcala de Biología de la Conducta. Universidad Autónoma de Tlaxcala. Tlaxcala, Tlaxcala.

Principios de estadística para el educador especial

Esta obra se terminó de editar en noviembre de 2023, en Tlaxcala de Xicohtécatl, Tlaxcala.

Tiene 362 páginas incluyendo la portada y contraportada, y se encuentra en formato PDF con un tamaño de 8.33 MB (8,741,304 bytes).

Se empleó el procesador Word para la conformación de la obra y las tipografías son Times New Roman, Alata, Californian FB y Pinyon Script.

Tiraje: Descargas ilimitadas

Diseño editorial:
Myriam Nayeli Villafuerte Vega

