Tamaño de muestra agropecuaria

Pablo Tadeo Cruz *

RESUMEN

En este documento, se analizan condiciones de productividad y rentabilidad de los Sistemas de Producción Agropecuaria, dónde se aplica metodología de determinación de tamaño de muestras a partir del Coeficiente de Error Típico para medias y del Error Típico de Proporciones, con productores rurales de Acayucan. Se obtuvieron resultados donde se infiere que la producción de leche por vaca al día, se encuentra entre 3.7 kg y 4.1 kg, para una probabilidad del 95.00%. Se acepta la muestra de n = 97 vacas, dado que *u* = 0.0233. Se deduce que la proporción de vacas con ubre sana, se encuentra entre 80.54% y 89.38%, para una probabilidad del 95.00%. Se acepta la muestra de n = 113 vacas, con error típico de proporciones de, *e* = 0.0223.

I. INTRODUCCIÓN

La población en estudio, de la cual se infiere la *media (μ)* y *desviación típica (σ)* de los kilogramos diarios de leche producida por unidad animal y la *proporcionalidad (P)* con ubre sana, se compone de un hato de vacas en ordeño de la comunidad rural Esperanza Malota, con régimen ejidal en la tenencia de la tierra y ubicada en el trópico húmedo, del Municipio de Acayucan, Estado de Veracruz de Ignacio de la Llave.

Para conocer los parámetros de la población en ordeño, se identificó una muestra aleatoria que representó las vacas en producción y de ella, se realizó la inferencia de sus características principales con cierto grado de confiabilidad.

Se retomó la opción de determinar el tamaño mínimo de datos de vacas productoras de leche para la conformación del trabajo, a partir de estadísticas muestrales; con un nivel de confianza del intervalo de estimación de la media de los kilogramos de leche y proporcionalidad de ubre sana de todas las vacas del hato, el valor de la desviación típica de los datos de la muestra, el coeficiente de error típico de aceptación y del error típico de proporciones. Suponiendo que el registro de pesos de los kilogramos de leche producida y la proporcionalidad de ubre sana, se distribuye en forma normal.

Se dio por hecho que, a partir de la media ($\overline{XX}$) y desviación típica (s) y la proporcionalidad ($\overline{pp}$) de una muestra (n) aleatoria extraída de una población de vacas en ordeña (N) que se distribuye normalmente, se puede estimar el promedio general, con un error de estimación, que delimita el intervalo de confianza donde se encuentra la verdadera media (μ) o proporcionalidad (P) de la población.

II. METODOLOGÍA PARA CALCULAR EL TAMAÑO DE MUESTRA

II.1 Tamaño de muestra para poblaciones infinitas

Cuando no se conoce el tamaño de la población (N), se supone que ésta es infinita y la estimación de la media (μ) a partir del *intervalo de confianza* para poblaciones de este tipo con distribución normal, se aplica la siguiente ecuación, por ser la más usual:

$$(1) \quad \mu = \bar{x} \pm t_{\left(1-\frac{\alpha}{2}\right)}\left(\frac{s}{\sqrt{n}}\right)$$

μ = Media de la población,

$\overline{XX}$ = Media de la muestra,

s = Desviación típica de la muestra,

n = Tamaño de la muestra.

$t_{\left(1-\frac{\alpha}{2}\right)}$ $t_{\left(1-\frac{\alpha}{2}\right)}$ = Índice de confianza para distribución de probabilidad t de Student

La relación del error típico con la media de la muestra, se representa por un *coeficiente de error típico* (Tadeo Cruz, Pablo. IIESCA 2005) y su ecuación para poblaciones infinitas, es:

$$(2) \quad u = \frac{s}{\bar{x}\sqrt{n}}$$

u = Coeficiente de error típico de la muestra

$\left(\frac{s}{\sqrt{n}}\right) = \left(\frac{s}{\sqrt{n}}\right)$ = Error típico de la muestra

$$s = \sqrt{s^2 s} = \sqrt{s^2}$$

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{(n-1)}$$

* Licenciado en Economía. Egresado del Instituto Politécnico Nacional y Profesor en la Facultad de Ingeniería en Sistemas de Producción Agropecuaria de la Universidad Veracruzana. México. Correo Electrónico: ptadeo@uv.mx

El tamaño de la muestra definitiva para poblaciones infinitas se calcula a partir de la proporcionalidad que guarda el coeficiente de error típico respecto a la media de una muestra piloto previa, como antecedente de los datos de la población; resultado de un coeficiente esperado *u*; $\bar{x}\bar{x}$ y *s*, preliminares. Donde a partir de la ecuación del coeficiente de error típico (2), se despeja *n* y sustituyendo se obtiene:

$$(3)\quad \mathbf{n} = \frac{\mathbf{s}^2}{(\mathbf{u}\bar{\mathbf{x}})^2}$$

Cuando se trata de inferir proporciones (P) de una muestra definitiva, por intervalo de confianza, a partir de datos preliminares, generalmente se usa:

$$(4)\quad \mathbf{P} = \bar{p} \pm \mathbf{t}_{\left(1-\frac{\alpha}{2}\right)}\left(\frac{s_p}{\sqrt{\mathbf{n}}}\right)$$

P = Proporción de la población,

$\bar{p}\bar{p}$ = Proporción de la muestra,

$s_p s_p$= Desviación típica de proporciones de la muestra,

El error típico de proporciones de la muestra para poblaciones infinitas, se representa por la siguiente ecuación:

$$(5)\quad \mathbf{e} = \frac{s_p}{\sqrt{\mathbf{n}}}$$

e = Error típico de proporciones de la muestra

$$s_p = \sqrt{(s_p{}^2)}\, s_p = \sqrt{(s_p{}^2)}$$

$$s_p{}^2 = \frac{\Sigma(x-\bar{p})^2}{(n-1)}$$

La muestra definitiva de una población desconocida, calculada para un error típico de proporciones esperado *e* (5); a partir de una proporcionalidad ($\bar{p}\bar{p}$) preliminar; despejando y sustituyendo para *n*, es:

$$(6)\quad \mathbf{n} = \frac{s_p{}^2}{\mathbf{e}^2}$$

Aplicando el procedimiento para poblaciones infinitas, se infiere la media poblacional y proporcionalidad, en un intervalo de confianza con probabilidad definida. Aceptando la información de la muestra definitiva, como aproximación a la realidad, con un coeficiente de error típico (2): $u \leq 0.025$; y error típico de proporciones (5): $e \leq 0.025$.

II.2 Tamaño de muestra para poblaciones finitas

Conociendo la población de estudio (N), la estimación de la media a partir del intervalo de confianza para datos poblacionales y proporcionalidades, se puede calcular *el coeficiente de error típico (u), el error típico de proporciones* (*e*) *y el tamaño de la muestra de estudio definitiva (n)*, aplicando un factor de corrección (Freund, John E. 1994) para poblaciones finitas, [√((N - n) / (N - 1))], de la siguiente forma:

Para media poblacional y su variación:

$$(7)\quad \mu = \bar{\mathbf{x}} \pm t_{\left(1-\frac{\alpha}{2}\right)}\left(\frac{\mathbf{s}}{\sqrt{\mathbf{n}}}\right)\sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-\mathbf{1}}}$$

$$(8)\quad \mathbf{u} = \frac{\mathbf{s}}{\bar{\mathbf{x}}\sqrt{\mathbf{n}}}\sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-\mathbf{1}}}$$

La muestra definitiva extraída de una población conocida, con un coeficiente de error típico esperado *u*; a partir de $\bar{x}\bar{x}$ y *s*, preliminares, se calcula de la siguiente forma:

$$(9)\quad \mathbf{n} = \frac{\mathbf{N}\mathbf{s}^2}{[(\mathbf{u}\bar{\mathbf{x}})^2(\mathbf{N}-\mathbf{1})] + \mathbf{s}^2}$$

Para proporciones y su variación:

$$(10)\quad \mathbf{P} = \bar{p} \pm \mathbf{t}_{\left(1-\frac{\alpha}{2}\right)}\left(\frac{s_p}{\sqrt{\mathbf{n}}}\right)\sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-\mathbf{1}}}$$

$$(11)\quad \mathbf{e} = \frac{s_p}{\sqrt{\mathbf{n}}}\sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-\mathbf{1}}}$$

La muestra calculada para un error típico de proporciones (11) esperado, a partir de una población finita, con datos de una muestra preliminar, despejando para *n* y sustituyendo, se obtiene la siguiente fórmula:

$$(12)\quad \mathbf{n} = \frac{\mathbf{N}s_p{}^2}{[\mathbf{e}^2(\mathbf{N}-\mathbf{1})] + s_p{}^2}$$

III. APLICACIONES EN EL SECTOR AGROPECUARIO

III.1 Media poblacional y variación de producción de leche por vaca al día

Se recabaron datos preliminares de la población de vacas en ordeño de la Comunidad Esperanza Malota, de 31 de ellas se pesaron los kg de producción de leche por vaca al día y los resultados fueron resumidos de la siguiente forma (Cuadro 1):

Producción de leche, litros/vaca/día	
Media de la muestra, $\overline{xx}$	3.8634
Error típico, $s/\sqrt{n}$	0.1713
Mediana	4
Moda	4
Desviación típica de la muestra, s	0.9536
Varianza de la muestra, s^2	0.9094
Rango	4.7037
Mínimo	1.2963
Máximo	6
Suma	119.7645
Muestra preliminar, n	31
Nivel de confianza t = 2.0423, 95.0%	0.3498
Coeficiente de error típico, u	0.0443

Cuadro 1. Fuente: F I Sistemas de Producción A

El coeficiente de error típico (2), resulta de:

$$u = \frac{s}{\bar{x}\sqrt{n}} = \frac{0.9536}{3.8634\sqrt{31}} = 0.0443$$

El coeficiente de error típico de la muestra de prueba es, *u* = 0.0443 y por lo tanto es mayor al requerimiento de que sea *u* ≤ 0.025. Para que la muestra de trabajo esperada se aproxime lo suficiente a la media de la población de las vacas en ordeño, es indispensable aumentar la cantidad de sus elementos.

Con base en lo anterior, se calculó una muestra definitiva (3) de *n* = 97, resultado de un coeficiente de error típico esperado de *u* ≤ 0.025 y de una población infinita; s = 0.9536; *u* = 0.025 y $\overline{xx}$ = 3.8634 (Cuadro 1). Donde a partir de la ecuación del *coeficiente de error típico* (2), se despeja *n* (3) y sustituyendo se obtiene:

$$n = \frac{(0.9536)^2}{(0.025 * 3.8634)^2} = 97$$

A partir de este resultado, se recabaron los 66 datos restantes para completar las 97 vacas en ordeño de la muestra de trabajo, se agruparon y resultó el siguiente resumen de estadísticas (Cuadro 2):

Producción de leche, litros/vaca/día	
Media de la muestra, $\overline{xx}$	3.9150
Error típico, $s/\sqrt{n}$	0.0911
Mediana	3.8462
Moda	3.75
Desviación típica de la muestra, s	0.8970
Varianza de la muestra, s^2	0.8046
Rango	4.8583
Mínimo	1.875
Máximo	6.7333
Suma	379.7578
Muestra de trabajo, n	97
Nivel de confianza t = 1.9850, 95.0%	0.1808
Coeficiente de error típico, u	0.0233

Cuadro 2. Fuente: F I Sistemas de Producción A

Dónde:

$$u = \frac{s}{\bar{x}\sqrt{n}} = \frac{0.8970}{3.9150\sqrt{97}} = 0.0233$$

El error de estimación de la media de la población con un nivel de confianza de 95% para una t = 1.9850, se aproxima al 4.62% respecto a la media de la muestra y el coeficiente de error típico (2) tiene un valor de 2.33%. Donde el intervalo de confianza (1) para la media de la población es de:

$$\mu = 3.9150 \pm (1.9850 * 0.0911)$$

Se infiere que la media poblacional de la producción de leche por vaca al día (1), es de aproximadamente: μ = 3.9150 kg, y puede fluctuar entre 3.7342 kg y 4.0958 kg, como intervalo de confianza para una probabilidad del 95.00%.

Se acepta la información de la muestra de las 97 vacas en ordeño, como una aproximación a la realidad considerando la probabilidad del 95% (Cuadro 2), dado que el resultado de coeficiente de error típico *u* = 0.0233, cumple con la restricción inicial considerada de, *u* ≤ 0.025.

Cuando se conoce la población de vacas en estudio (N), que son un total de 200 en el Ejido Esperanza Malota, la estimación de la media a partir del intervalo de confianza (7), se calcula aplicando un factor de corrección para poblaciones finitas. Sustituyendo los valores de las 31 vacas en ordeño iniciales (Cuadro 1), se tiene un *coeficiente de error típico* corregido (8), como sigue:

$$u = \frac{0.9536}{3.8634\sqrt{31}}\sqrt{\frac{200-31}{200-1}} = 0.0409$$

Retomando los datos de la muestra de las 31 vacas en ordeño productoras de leche diario (Cuadro 1), se calculó una muestra de trabajo definitiva (9) de n = 66, con un *coeficiente de error típico* esperado de $u \leq 0.025$; a partir de N = 200; s = 0.9536; u = 0.025 y $\overline{xx}$ = 3.8634:

$$n = \frac{200 * (0.9536)^2}{[(0.025 * 3.8634)^2(200-1)] + (0.9536)^2} = 66$$

Si desde el inicio del estudio se conociera la cantidad de elementos de la población en ordeño, no se hubieran recopilado los datos de las 97 vacas; solamente se recabarían de 66, que es el resultado al que se llegó como tamaño de la muestra para una población de N = 200 vacas.

III.2 Proporción poblacional y variación de vacas con ubre sana

De los 31 datos preliminares de la población infinita de vacas en ordeño, se detectó que el 19.35% de ellas tienen la ubre dañada con algún tipo de inflamación. Se tomó la decisión de saber cuál es el porcentaje de vacas con ubre sana del total del hato no conocido, de la Comunidad Ejidal Esperanza Malota; con un grado de confiabilidad del 95.00%. El resumen de datos se presenta a continuación (Cuadro 3):

Vacas con ubre sana	
Media de la muestra, $\bar{p}\bar{p}$	0.8065
Error típico de proporciones, e = $s_p/\sqrt{n}$	0.0721
Desviación típica de la muestra, s_p	0.4016
Varianza de la muestra, s_p^2	0.1613
Vacas con ubre sana	25
Muestra preliminar, n	31
Nivel de confianza t = 2.0423, 95.0%	0.1473

Cuadro 3. Fuente: F I Sistemas de Producción A

Dónde el error típico de proporciones e (5), resultado de los datos de la muestra preliminar, dónde p = 0.8065, es:

$$e = \frac{s_p}{\sqrt{n}} = \frac{0.4016}{\sqrt{31}} = 0.0721$$

Dado el error típico preliminar, e = 0.0721, entonces fue indispensable aumentar la cantidad de elementos de la muestra para que se cumpla el requisito de, $e \leq 0.025$.

La muestra calculada (6) para un error típico de proporciones esperado; a partir de una población infinita, e = 0.025, $s_p s_p$ = 0.4016; despejando y sustituyendo para n, resulta de:

$$n = \frac{{s_p}^2}{e^2} = \frac{(0.4016)^2}{(0.025)^2} = 258$$

Los datos de las 258 vacas como tamaño de muestra definitiva resultado de la aplicación de la fórmula para poblaciones infinitas, no se recabaron, ya que los productores manifestaron que solamente tenían 200 vacas en total y se optó por aplicar la metodología de determinación del tamaño de la muestra para poblaciones finitas.

Si partimos de que la población de las vacas en ordeño en Esperanza Malota es de N = 200 y de los datos de las 31 vacas iniciales (Cuadro 1), se puede calcular el porcentaje de vacas con ubre sana del total del hato.

Aplicando la fórmula correspondiente del error típico de proporciones para poblaciones finitas (11), se tiene:

$$e = \frac{0.4016}{\sqrt{31}}\sqrt{\frac{200-31}{200-1}} = 0.0665$$

Dado que el error típico calculado fue de, e = 0.0665, se aumentó la cantidad de elementos de la muestra para que se cumpliera con el requisito de, $e \leq 0.025$.

La muestra calculada (12) para un coeficiente de error típico esperado; a partir de N = 200, e = 0.025, $s_p s_p$ = 0.4016; despejando para n y sustituyendo, es de:

$$n = \frac{200 * (0.4016)^2}{(0.025^2 * (200-1)) + (0.4016)^2} = 113$$

Posteriormente se recabaron 82 datos faltantes para completar los 113 de la muestra definitiva de vacas en ordeño; se agruparon, ordenaron y analizaron, de donde resultó el siguiente resumen de estadísticas (Cuadro 4):

Vacas con ubre sana	
Media de la muestra, $\bar{p}\bar{p}$	0.8496
Error típico de proporciones, e = $s_p/\sqrt{n}$	0.0338
Desviación típica de la muestra, s_p	0.3591
Varianza de la muestra, s_p^2	0.1290
Vacas con ubre sana	96
Muestra de trabajo, n	113
Nivel de confianza t = 1.9814, 95.0%	0.0669

Cuadro 4. Fuente: F I Sistemas de Producción A

Siendo el error típico de proporciones para poblaciones finitas (11), de:

$$e = \frac{0.3591}{\sqrt{113}}\sqrt{\frac{200-113}{200-1}} = 0.0223$$

El error de estimación de la media de la población conocida, con un nivel de confianza de 95% para una t = 1.9814, se aproxima al 4.43% y el error típico de proporciones para poblaciones finitas (11) tiene un valor de 2.23%. Donde el intervalo de confianza (10) para la media de la población es de:

$$P = 0.8496 \pm (1.9814 * 0.0223)$$

De los datos recabados, se deduce que la proporción de las vacas con ubre sana (10) del hato de Esperanza Malota, es de 84.96%, con una fluctuación entre 80.54% y 89.38%, como intervalo de confianza para una probabilidad de ocurrencia del 95.00%.

Se acepta la información de la muestra de 113 vacas tomadas al azar del hato de las 200 productoras de leche, como aproximación a la realidad, con probabilidad del 95% (Cuadro 4) y error típico de proporciones de, *e* = 0.0223; cumpliéndose la condición inicial para, *e* ≤ 0.025.

V. REFERENCIAS BIBLIOGRÁFICAS

Douglas A.Lind (2008) *Estadística aplicada a los negocios y economía*. McGraw Hill. México.

Freund, John E. (1994) *Estadística matemática con aplicaciones*. Prentice-Hall. México.

Rodríguez Franco, Jesús. (2008) *Estadística para administración*. Grupo Editorial Patria. México.

Sánchez Corona, Octavio (2010) *Probabilidad y estadística*. McGraw Hill. México.

Scheaffer, Richard L (2000) *Elementos de muestreo*. Grupo Editorial Iberoamericano. México.

Tadeo Cruz, Pablo (2005) Coeficiente de error típico. IIESCA. Universidad Veracruzana. México.

Wackerly, Dennis D. (2010) *Estadística matemática con aplicaciones*. Editorial Cengage Learning. México.