

Muestreo Secuencial: Clases, Ventajas y Aplicaciones

(Sequential Sampling: Types, Advantages, and Applications)

Badii, M.H., A. Guillen, E. Cerna, J. Valenzuela & J. Landeros *

Resumen. Se presentan los conceptos básicos de muestreo secuencial en sus cinco modalidades, incluyendo muestreo secuencial de tipo estándar, usando solo una densidad crítica, por una precisión fija, por medio de presencia-ausencia y finalmente de tipo temporal. Se notan un ejemplo real para cada caso. Se nota y se explica la importancia de la distribución de los datos y los pertinentes ecuaciones. Se presentan a través de ejemplos reales las ecuaciones para estimar el tamaño óptimo de la muestra para cada uno de las cinco modalidades del muestreo secuencial. Se enfatiza la ventaja económica de este tipo de muestreo.

Palabras clave: Distribución de los datos, muestreo secuencial, tamaño óptimo de la muestra

Abstract. Fundamental concepts of sequential sampling in all its types including standard sequential sampling, sequential sampling using only one critical density, fixed precision sequential sampling, presence-absence sequential sampling, and time sequential sampling are highlighted. A real case example for each of the five sequential samplings is given. The relevance of data distribution for sequential sampling along with corresponding equations is provided and explained. Equations to assess the optimal sample size for each case are noted and exemplified. The economic advantages of sequential sampling in stressed

Keywords: Data distribution, sample size, sequential sampling

Fundamento

El propósito del muestreo secuencial es agrupar los elementos en diferentes clases, segmentos o poblaciones (Morris, 1955, Kogan & Herzog, 1980, Badii et al., 2000). Este muestreo es particularmente apto para las situaciones de mucha abundancia o mucha rareza, y por tanto, determina en forma rápida, la ubicación correcta de los elementos. Como consecuencia, el muestreo secuencial ahorra recursos (tiempo, energía, material, etc.) de forma significativa y esto constituye una ventaja grande de este tipo de muestreo. El muestreo secuencial se presenta en varias modalidades (Figura 1).

A. Muestreo secuencial clásico o estándar (MSC)

Cuando el objetivo del trabajo es clasificar la población en una de diferentes clases de densidad poblacional (alta, media o baja) y no necesariamente determinar la densidad poblacional con mucha precisión, entonces el esquema del muestreo secuencial es lo

más adecuado (Morris, 1955, Kogan & Herzog, 1980). El muestreo secuencial (MS) se usa principalmente para determinar si la población ha excedido un umbral de importancia económica; por lo tanto, este tipo de muestreo es muy práctico ya que ahorra mucho tiempo en determinar rápidamente en qué clase poblacional va a pertenecer el organismo. La ventaja principal de MS es la economía o la optimización de los recursos que se emplean en el muestreo. Hay tres requisitos para el muestreo secuencial clásica:

- 1. El patrón de dispersión espacial.** Este tipo de muestreo requiere conocer estadísticamente el tipo de dispersión espacial de los datos.
- 2. Los niveles económicos.** Estos niveles son: m_1 (límite superior de la clase baja) y m_2 (límite inferior de la clase alta), estos dos niveles están relacionados con el umbral económico, es decir, el máximo nivel poblacional tolerable sin ocasionar daño económico o el nivel poblacional sobre el cual se debe emplear un método de control para evitar que la población creciente alcance el nivel de daño económico (el mínimo nivel que sí ocasiona pérdidas económicas) para que de esta manera, se puede aprovechar el recurso de forma racional (Badii, et al., 1995a, 1995b, 1996b).
- 3. Los niveles de riesgo.** El riesgo "a" que es la probabilidad de clasificar a una población como grande cuando no la es, y "b" que es lo contrario de "a". La ecuaciones generales de las líneas del muestreo secuencial son:

$$d_1 = bn - h_1 \quad \text{ecuación de la línea inferior}$$

$$d_2 = bn + h_2 \quad \text{ecuación de la línea superior}$$

Donde, d es el número acumulativo de los individuos, b es la pendiente de la línea, n es el número de muestras y h_1 y h_2 son las intersecciones con la ordenada para las líneas inferiores y superiores, respectivamente. Ahora bien, se calculan la b y la h según las ecuaciones específicas dependiendo del tipo de dispersión espacial:

Caso de distribución Poisson:

$$b = (m_2 - m_1) / (\ln m_2 - \ln m_1)$$

$$h_1 = \ln [(1 - a)/b] / (\ln m_2 - \ln m_1)$$

$$h_2 = \ln [(1 - b)/a] / (\ln m_2 - \ln m_1)$$

Donde, \ln = Logaritmo natural y los demás notaciones como arriba descritas.

Caso de distribución uniforme:

$$b = \ln [(1 - m_1) / (1 - m_2)] / \ln [(m_2/m_1)(1 - m_1) / (1 - m_2)]$$

$$h_1 = \ln [(1 - a)/b] / \ln [(m_2/m_1)(1 - m_1) / (1 - m_2)]$$

$$h_2 = \ln [(1 - b)/a] / \ln [(m_2/m_1)(1 - m_1) / (1 - m_2)]$$

Donde, todos las anotaciones como arriba mencionadas.

Caso de distribución agregada (modelo de binomial negativa):

$$b = k [\ln (q_2/q_1)] / \ln(p_2 q_1/p_1 q_2)$$

$$h_1 = \ln[b/(1-a)] / \ln(p_2 q_1/p_1 q_2)$$

$$h_2 = \ln [(1-b)/a] / \ln(p_2 q_1/p_1 q_2)$$

Donde,

k = El parámetro de distribución de binomial negativa

$$p = m/k$$

m = Media muestral

$q = p + 1$, Y los demás notaciones como antes mencionadas.

Ejemplo de MSC

Suponemos que hay impureza en un artículo, y el nivel crítico que ocasiona merma en el artículo es de 18 ppm, también suponemos que los valores de m_1 (límite superior de la clase baja) y m_2 (límite inferior de la clase alta) son 1/3 y 2/3 de nivel crítico, respectivamente. Hay que acordar que una población baja significa un nivel de impureza aceptable mientras que una población alta significa lo contrario.

Caso 1. Distribución Poisson

Dado que los riesgos de "a" y "b" son igual a 0.10, la pregunta es el diseñar un esquema de muestreo secuencial clásico (MSC) si la distribución es de tipo Poisson.

Solución

Ecuaciones para la distribución Poisson.

$$b = (m_2 - m_1) / (\ln m_2 - \ln m_1)$$

$$b = (12-6) / (\ln 12 - \ln 6) = 8.65$$

$$h_1 = \ln [(1 - a)/b] / (\ln m_2 - \ln m_1)$$

$$h_1 = \ln [1 - 0.10]/0.10] / (\ln 12 - \ln 6) = 3.16$$

$$h_2 = \ln [(1 - b)/a] / (\ln m_2 - \ln m_1)$$

$$h_2 = \ln [1 - 0.10]/0.10] / (\ln 12 - \ln 6) = 3.16$$

Las ecuaciones de las líneas inferiores (d_1) y superiores (d_2) son:

$$d_1 = bn - h_1, \quad d_1 = 8.3n - 3.16$$

$$d_2 = bn + h_2, \quad d_2 = 8.3n + 3.16$$

Ahora asignamos por ejemplo, de 1 a 10 a la “ n ” para construir el plan de muestreo para 10 muestras consecutivas (Tabla 1).

Tabla 1. Esquema de MSC para este ejemplo con la distribución Poisson*.

N	Población (zona) baja	Zona o banda intermedia (anchura = 6)	Población (Zona) alta
1	06	07 – 11	12
2	15	16 – 20	21
3	23	24 – 29	30
4	32	33 – 37	38
5	41	42 – 46	47
6	49	50 – 54	55
7	58	59 – 63	64
8	67	68 – 72	73
9	75	76 – 80	81
10	84	85 - 89	90

*: Los datos están redondeados para tener valores íntegros.

De esta tabla podemos concluir que si la primera muestra arroja un valor igual o menor del 6, entonces tenemos una población baja (aceptable), y si el valor es al menos 12 se trata de una población alta (contaminada). Sin embargo, si el valor está entre el 6 y el 12, no se sabe si se trata de una población baja o alta, y por tanto, esta incertidumbre (estar en la zona o banda intermedia o indecisión) nos obliga y a la vez justifica realizar una muestra subsecuente, es decir, tomar la según muestra. La noción de conducir el MSC es precisamente el ahorro económico, ya que este tipo de plan de muestreo, generalmente, nos ubica de manera rápida (con las primeras pocas muestras) en una u otra zona.

Caso 2. Distribución uniforme

Para los datos del ejemplo de MSC, diseñar un plan de muestreo suponiendo que la distribución de los datos es uniforme.

Solución

Las ecuaciones para la distribución uniforme en este caso son:

$$b = \ln [(1 - m_1) / (1 - m_2)] / \ln [(m_2/m_1)(1 - m_1) / (1 - m_2)]$$

$$b = \ln [(1 - 6) / (1 - 12)] / \ln[(12/6)(1 -6) / (1 -12)] = 8.27$$

$$h_1 = \ln [(1 - a)/b] / \ln [(m_2/ m_1)(1 - m_1) / (1 - m_2)]$$

$$h_1 = \ln [(1 - 0.10)/0.10] \ln[(12/6)(1 -6) / (1 -12)] = -23.05$$

$$h_2 = \ln [(1 - b)/a] / \ln [(m_2/ m_1)(1 - m_1) / (1 - m_2)]$$

$$h_2 = \ln [(1 - 0.10)/0.10] \ln[(12/6)(1 -6) / (1 -12)] = +23.05$$

Por tanto, las ecuaciones de las líneas inferiores (d_1) y superiores (d_2) son: $d_1 = 8.27n - 23.05$, $d_2 = 8.27n + 23.05$.

Se asignan de 1 a 10 para la “n” con el objetivo de construir el plan de muestreo secuencial (Tabla 2).

Tabla 2. Esquema de MSC para este ejemplo con la distribución uniforme*.

n	Población (zona) baja	Zona o banda intermedia (anchura = 46)	Población (Zona) alta
1	00	00 – 31	32
2	00	00 – 39	40
3	02	03 – 47	48
4	11	12 – 56	57
5	19	20 – 64	65
6	27	28 – 72	73
7	35	36 – 80	81
8	44	45 – 89	90
9	52	53 – 97	98
10	60	61 -105	106

*: Los datos están redondeados para tener valores íntegros.

Se usa la misma filosofía que el caso de Poisson para el procedimiento de muestreo secuencial, es decir, ubicación rápida en una de las zonas (alta o baja) o tomar más muestras (zona intermedia).

Caso 3. Distribución agregada

Para los datos del ejemplo de MSC, diseñar un plan de muestreo suponiendo que la distribución de los datos es agregada. Supongamos que el valor de K de la distribución de binomial negativa es igual a 1.023, y los valores de p_1 , p_2 , q_1 , q_2 se calculan como sigue.

$$p_1 = m_1 / K = 6 / 1.023 = 5.8651$$

$$q_1 = 1 + p_1 = 6.8651$$

$$p_2 = m_2 / K = 12 / 1.023 = 11.7302$$

$$q_2 = 1 + p_2 = 12.7302$$

Solución

Las ecuaciones para la distribución uniforme en este caso son:

$$b = k [\ln (q_2/q_1)] / \ln(p_2 q_1/p_1 q_2)$$

$$b = 1.023[\ln (12.7302/6.8651)]/\ln(11.7302*6.8651/5.8651*12.7302)=8.35$$

$$h_1 = \ln[(b/(1-a)] / \ln(p_2 q_1/ p_1 q_2)$$

$$h_1 = \ln [0.10 (1- 0.10)] / \ln(11.7302*6.8651/5.8651*12.7302) = -28.73$$

$$h_2 = \ln [(1-b)/a] / \ln(p_2 q_1/ p_1 q_2)$$

$$h_2 = \ln [81 - 0.10]/0.10/\ln(11.7302*6.8651/5.8651*12.7302) = +28.73$$

Las ecuaciones de las líneas inferiores (d_1) y superiores (d_2) son:

$$d_1 = 8.35n - 28.73$$

$$d_2 = 8.35n + 28.73$$

Se asignan de 1 a 10 para la “ n ” con el objetivo de construir el plan de muestreo secuencial (Tabla 3).

Tabla 3. Esquema de MSC para este ejemplo con la distribución agregada*.

N	Población (zona) baja	Zona o banda intermedia (anchura = 58)	Población (zona) alta
1	00	01 – 37	38
2	00	001 – 45	46
3	00	01 - 54	55
4	05	06 – 62	63

5	13	14 – 70	71
6	22	23 – 79	80
7	30	31 – 87	88
8	38	30 – 95	96
9	47	48 – 104	105
10	55	56 -112	113

*: Los datos están redondeados para tener valores íntegros.

B. Muestreo secuencial en base a sola una densidad crítica (MSDC)

Iwao (1975) utilizando los parámetros de su modelo de hacinamiento llegó a establecer un nuevo esquema de muestreo secuencial en donde se usa sólo una densidad crítica. Las ecuaciones de este modelo son:

$$\begin{aligned} \text{Línea inferior:} \quad d_i &= n m_0 - t \{[(a-1) m_0 + (b-1) m_0^2]^{1/2}\} \\ \text{Línea superior:} \quad d_s &= n m_0 + t \{[(a-1) m_0 + (b-1) m_0^2]^{1/2}\} \end{aligned}$$

Donde, d_i y d_s son el número acumulativo de individuos para las líneas inferior y superior, respectivamente, n es el número de muestras, m_0 es la densidad crítica, a y b son los parámetros de Iwao (1968), y la t es el valor tabulado de la distribución de t de student con los grados de libertad que se usan para estimar la varianza muestral. Basándose en este tipo de muestreo y debido al uso de sólo una densidad crítica (en lugar de dos), las líneas generadas por estas dos ecuaciones van a ser divergentes en lugar de ser paralelas (caso de muestreo secuencial clásica).

Ejemplo de MSDC

Supongamos los datos siguientes: $a = 0.621$, $b = 1.305$, la densidad crítica o $m_0 = 15$. Construir un esquema de muestreo secuencial en base de una densidad crítica (MSDC).

Solución

En base a los datos arriba mencionados, y utilizando las siguientes ecuaciones se puede construir un esquema de MSDC (Tabla 4) para diferentes valores de n de 1 hasta 10, por ejemplo.

$$\begin{aligned} \text{Línea inferior:} \quad d_i &= n m_0 - t \{[(a-1) m_0 + (b-1) m_0^2]^{1/2}\} \\ \text{Línea superior:} \quad d_s &= n m_0 + t \{[(a-1) m_0 + (b-1) m_0^2]^{1/2}\} \end{aligned}$$

Tabla 4. Esquema de MSDC*.

<i>n</i>	<i>d_i</i>	Banda intermedia	<i>d_s</i>
1	0	1 - 36	37
2	0	1 - 60	61
3	7	8 - 82	83
4	16	17 - 103	104
5	26	27 - 123	124
6	37	38 - 142	143
7	47	48 - 162	163
8	58	59 - 181	182
9	70	71 - 199	200
10	81	82 - 219	220

*: 1. Los datos estan redondeados para tener valores íntegros. 2. Notase que la anchura de la banda intermedia se expande con el aumento de la "n".

C. Muestreo secuencial en base a una precisión fija (MSPF)

El muestreo secuencial se utiliza para decidir si la población rebasa un umbral económico; sin embargo, en muestreo y monitoreo de los recursos naturales el objetivo es obtener estimaciones de densidad con un nivel dado de precisión para poder evaluar la eficiencia del uso de los recursos, por lo tanto, se debe terminar el muestreo tan pronto se consigue este nivel deseado de precisión. Los esquemas de muestreo secuencial clásico mencionados por Kuno (1969) y Green (1970), por un lado, son especialmente adecuados para los experimentos de campo cuando se requiere un alto grado de precisión; por otro lado, para los monitoreos extensivos son imprácticos, ya que el procedimiento requiere que no solamente se seleccionen las unidades muestrales de forma aleatoria, sino también, la secuencia de incluir estas unidades en la muestra debe ser estrictamente al azar. Esto exige gastar mucho tiempo en el área de monitoreo moviéndose en direcciones según el esquema aleatorio. En contraste, cuando se usa el muestreo secuencial basado en una precisión fija, se puede coleccionar las unidades muestrales en la secuencia que se les encuentran en el campo según un esquema predeterminado del muestreo.

La línea de terminación del muestreo se determina basándose en las ecuaciones siguientes según el autor.

Ecuación de Kuno (1969)

$$d_n = (a+1)/[D^2 - (b-1)/n]$$

Donde, a y b son parámetros del modelo de Iwao (1968), y las demás notaciones como antes descritas.

Ecuación de Green (1970)

$$\ln d_n = [\ln (D^2/a)/(b-2)] + [(b-1)/(b-2)] \ln n$$

Donde, \ln es logaritmo natural, a y b son los parámetros de Taylor (1961), y los demás parámetros son como antes mencionadas.

Ejemplo para MSPF

Suponiendo los siguientes datos ($a = 0.586$, $b = 1.336$, y $D = 0.25$), construir un plan de MSPF:

Tabla 5. Un esquema de MSPF*.

N	dn (línea de terminación)
1	30
2	21
3	17
4	15
5	13
6	12
7	11
8	11
9	10
10	10
11	9
12	9
13	8
14	8
15	8
16	8
17	7
18	7

*: Los datos están redondeados para tener valores íntegros.

Solución

Usar el modelo de Green (1970):

$$\ln d_n = [\ln (D^2/a)/(b-2)] + [(b-1)/(b-2)] \ln n$$

Se construye la Tabla 5 para MSPF para diferentes valores de “n” de 1 hasta 18.

Los datos de la Tabla 5 indican una tendencia exponencial negativa, es decir, con el aumento en la “n”, se disminuye el valor de dn . El procedimiento es que el investigador comienza a muestrear y al instante que el valor de dn cruza la línea de terminación de muestreo, se debe terminar el acto de muestreo, y en este momento al dividir el valor de la eje vertical (dn) entre el valor correspondiente de la eje horizontal (n) se obtiene la densidad promedio de la población con una precisión predeterminada fija, por ejemplo, para este caso la precisión sería igual a 0.75, ya que el valor de $D = 0.25$.

D. Muestreo secuencial basado en presencia-ausencia (MSP-A)

A pesar de que el muestreo secuencial como objetivo trata de minimizar el tamaño de muestra, sin embargo, el conteo de los individuos, especialmente, cuando varias especies están involucradas, (caso de especies muy abundantes; gorgojos como contaminante de galletas, etc.) es muy laborioso; por lo tanto, se han propuesto procedimientos alternos de muestreo, donde se evita el conteo de los individuos (Nachman, 1981, 1984). Los dos métodos descritos aquí están basados en la frecuencia de las unidades muestrales con y sin individuos y ambos métodos requieren muestreo preliminar para clarificar la forma de interrelación que existe entre la media muestral y algunos atributos específicos de la dispersión espacial del organismo bajo estudio. Debido a que esta fase inicial puede involucrar bastante conteo, la inversión del tiempo sólo se justifica por los ahorros subsecuentes en muestreos poblacionales posteriores. El primer método requiere que la dispersión espacial de la binomial negativa se ajuste a los conteos de los individuos provenientes de un muestreo aleatorio, y que el parámetro de la binomial negativa “ k ” sea independiente de la media muestral para que de este modo se pueda estimar un “ k común” para todas las muestras (Bliss & Owen, 1958). En estos casos, la media poblacional “ m ” (expresada como número de individuos por unidad de muestra) se puede estimar de la ecuación siguiente: $m = k (P_o^{(-1/k)} - 1)$, donde k es el valor estimado de k común, y P_o es la proporción observada de unidades muestrales sin individuos en una muestra. Wilson & Gerrard (1971) presentan una fórmula aproximada para la varianza de m . El tamaño óptimo de la muestra (n_{opt}) y el error estándar (EE_m) se estiman en base a las siguientes ecuaciones:

$$n_{opt} = 1/D^2 m^2 (1 + m/k)^{k+2} [1 - (1 + m/k)^{-k}]$$

$$EE_m = 1/n P_i [(1 - P_i)^{(-2/k - 1)}]$$

Donde, P_i = Proporción de las UM's ocupados con elementos = $1 - P_o$.

A pesar de que se ha usado el modelo de binomial negativa para describir la frecuencia de distribución (Croft et al., 1976, Badii & McMurtry, 1990, Badii et al., 1994a, 1994b, 1996a, 1998), el uso de k común se justifica probablemente solo en aquellos casos en donde la especie demuestra poca variabilidad en la abundancia. De otra manera, es muy probable que K varíe con la densidad poblacional (Taylor et al., 1979). El segundo método de estimación (Gerrard & Chiang, 1970, Torii, 1971, Nachman, 1981, 1984) no requiere definir el tipo de dispersión espacial de los elementos, este método solo requiere que los datos derivados de una muestra piloto demuestren una relación lineal entre “ $\ln m$ ” y “ $\ln(-\ln P_o)$.” Se puede aplicar una regresión lineal para obtener los parámetros de una línea recta (línea de regresión) dada por: $\ln m = b'[\ln(-\ln P_o)] + a'$. Se insertan valores de P_o en esta ecuación y de este modo se obtienen estimaciones de m . Las fórmulas para la varianza y los intervalos de confianza están dadas por Nachman (1984). Badii et al., (1998) generan un esquema de muestreo secuencial basándose en presencia-ausencia para los ácaros sobre los cítricos en Nuevo León, México.

En práctica se usan las siguientes ecuaciones para diferentes estimaciones necesarias para este tipo de muestreo secuencial en base a presencia-ausencia para el grupo de Kono-Sujino (KS):

1. Estimar los valores de “ a ” y “ b ”: $\ln m = \ln a + \beta \ln(-\ln P_o)$
Donde, ‘ $\alpha = \text{antiln}(\ln a)$, $a = e^{(\alpha/\beta)}$, y $b = 1/\beta$.
2. $\mu = [(1/a) * \ln P_o]^{1/b}$
Donde, μ = la media poblacional estimada.
3. $n_{\text{opt}} = (1/1/D^2 a^2 b^2) \mu^{-2b} * [\exp(a\mu^b) - 1]$
4. $EEm = (P_i/n) a^{(-2/b)} * b^{(-2)} (1 - P_i)^{-1} * [-\ln(1 - P_i)]$
5. Límite de error de estimación (L) = $\pm tEEm$
Donde, t = valor de la tabal de t-student.

Cuando se trata de muestrear poblaciones de organismos que alcanzan densidades muy altas por unidad de muestra, tales como los gorgojos, ácaros, pulgones, trips, etc., y que además son especies con un grado muy alto de contagio, las estimaciones de la abundancia se asocian con gran incertidumbre, y por lo tanto, se requiere de tamaños grandes de muestra para poder reducir el error del muestreo.

Resumen de MSP-A

Los elementos por ser de tamaño pequeño su conteo es muy laborioso y tedioso, por lo que se necesita de un método que implique menos tiempo y trabajo, sin un sacrificio fuerte de la confiabilidad. Lo anterior es posible si existe una relación funcional específica entre densidad y distribución espacial de una población. En este sentido, la densidad media de una población se puede estimar a partir de la proporción de unidades de muestreo que no están ocupadas por los elementos, es decir, el grupo cero de la distribución de frecuencias. Este método fue propuesto originalmente por Pielou (1960) y después por Gerrard & Chiang (1970), y posteriormente, Nachman (1984) lo aplico exitosamente para el muestreo de ácaros en invernadero en Dinamarca.

Desarrollo del modelo

Si consideramos a un ambiente físicamente estable, que consiste de un número grande de microhábitats óptimos para una especie (parches), la cual lógicamente se agrega en ellos, a medida que aumenta el tamaño de la población, también más microhábitats son colonizados por la especie. La tasa con la cual esto ocurre, se supone que es proporcional a la tasa instantánea de incremento per capita (r), el número promedio de individuos por microhábitats (μ), la tasa de colonización per cápita $g(\mu)$ y la proporción de microhábitats vacíos (p_0). Ya que el número de microhábitats es grande, p_0 se puede considerar como una variable continua y su cambio instantáneo durante un intervalo infinitesimal de tiempo d_t , se describe por:

$$dp_0/d_t = -rp_0g(\mu)\mu$$

Donde, $g(\mu)$ es una función dependiente de densidad descrita por: $g(\mu) = \alpha\mu^B$, α y B como parámetros característicos de la especie de este ambiente. Ya que $r\mu d_t$ se refiere al cambio neto en la densidad poblacional durante d_t , es substituido por $d\mu$, y se transforma por integración a: $p_0 = e^{(-\alpha'\mu^B)}$, Donde: $\alpha' = \alpha/(b+1)$, y $B' = B+1$. La ecuación $p_0 = e^{(-\alpha'\mu^B)}$ es independiente del tiempo y describe una relación funcional entre la densidad media y la proporción de microhábitats ocupados. Asumiendo que esta ecuación es un modelo disponible para la especie bajo estudio, la densidad media poblacional (expresada como el promedio de individuos por unidad de muestreo) puede ser estimada a partir de una muestra de la proporción de unidades vacías, transformando la ecuación arriba a:

$$\ln\mu = b' \ln(-\ln p_0) + a', \quad 0 < p_0 < 1$$

Si la ecuación $[\ln\mu = b' \ln(-\ln p_0) + a']$ es un modelo apropiado de la relación entre la densidad media y la distribución espacial, al graficar $\ln \mu$ (dependiente) contra $\ln(-\ln P_0)$, aparecerá una tendencia lineal y una recta se puede ajustar por medio de regresión lineal ordinaria (debe ser significativa), para el cálculo de a' y b' . Sin embargo, para remover el error introducido al aplicar regresión lineal a los datos logarítmicos, la ecuación de estimación debe corregirse sumándole $CM/2$ a a' , donde CM = es el cuadrado medio del error de regresión; por lo tanto, la densidad media será estimada por:

$$\mu = e^{[b' \ln(-\ln p_0) + (a' + CM/2)]}$$

Además de lo anterior, la aplicación del modelo requiere que se calcule la ley de poder de Taylor, determinado a y b (con \log_{10}), con las mismas muestras de la población. En general, los parámetros del modelo son:

- a' = Intercepto de regresión de $\ln\mu$ contra $\ln(-\ln p_0)$
- b' = Pendiente de regresión de $\ln\mu$ contra $\ln(-\ln p_0)$
- CM = Cuadrado medio del error
- x^2 = Suma de cuadrados de desviaciones de x de Taylor
- x = Variable horizontal de Taylor
- y = Variable vertical de Taylor
- N = Número de observaciones totales
- n = Tamaño de muestra
- a = Intercepto de ley de poder de Taylor
- b = Pendiente de ley de poder de Taylor

Una vez que se obtienen todos los parámetros, se calcula la varianza de $\ln\mu$, como sigue:

$$V[\ln\mu] = CM(1/N + (\ln(-\ln p_0) - m)^2 / \sum x^2) + (b')^2 (1 - \ln p_0) / np_0 (\ln p_0) + (a'/n)(\mu)^{(b-2)}$$

El coeficiente de variación asociada al método es: $C\{\mu\} = \sqrt{V(\ln\mu)}$, mientras que el de la ley de poder de Taylor es: $C\{m\} = \sqrt{a/n} * (m^{(b-2)/2})$.

Se ha observado que el método, como quiera, sacrifica precisión, porque a densidades intermedias la máxima precisión es un 50% del error estándar de la media.

Ejemplo de MSP-A

Caso 1

Distribución binomial negativa

Dado que el valor de $K = 1.02459$ y los valores de P_i (proporción ocupada de las unidades muestrales) son los siguientes (Tabla 6), estimar los valores de la media (m), el tamaño óptimo de la muestra (n) y el error de estimación (error estándar).

Tabla 6. Valores de P_i (proporción de UM's ocupados).

.01	.03	.09	.12	.15	.31	.38	.40	.45	.50	.70	.75	.90	.95
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Solución

En base a las ecuaciones para el cálculo de la m , n_{opt} y EE_m se construye la siguiente tabla (Tabla 7):

Tabla 7. Valores de m , n_{opt} y EE_m para los datos de la Tabla 16.

P_i	M	EE_m	n_{opt}
0.01	0.0100	0.1009	1631
0.03	0.0309	0.1785	549
0.09	0.0987	0.3302	195
0.12	0.1361	0.3952	152
0.15	0.1761	0.4590	126
0.31	0.4471	0.8412	74
0.38	0.6081	0.0626	67
0.40	0.6622	1.1360	66
0.45	0.8117	1.3463	64
0.50	0.9907	1.6060	63
0.70	2.2924	3.7939	74
0.75	2.9397	5.0657	84
0.95	18.0400	8.0968	323

$$m = k[(1 - P_i)^{-1/k} - 1]$$

$$n_{opt} = 1/D^2 m^2 (1 + m/k)^{k+2} [1 - (1 + m/k)^{-k}]$$

$$EE_m = 1/n P_i [(1 - P_i)^{-(2/k - 1)}]$$

Caso 2

Modelos de tipo Kono-Sujino (KS)

Suponer los siguientes datos (Tabla 8) de la media (m) y la proporción de las unidades muestrales ocupada (P_i), construir un plan de muestreo secuencial en base a la presencia ausencia (MSP-A).

Tabla 8. Valores de m y P_i .

M	P_i
1.1	0.03
1.7	0.09
1.9	0.15
2.0	0.20
2.1	0.22
2.2	0.28
2.3	0.37
2.5	0.42
3.5	0.60
5.4	0.73

Solución

En base a los datos de la Tabla 8 y las ecuaciones siguientes se construye la Tabla 9.

- Estimar los valores de “ a ” y “ b ”: $\ln m = \ln \alpha + \beta \ln (- \ln P_0)$

$\ln \alpha = 1.2840$
 $\beta = 0.3621$
 $r = 0.9448$
 $\hat{\alpha} = \text{antiln}(\ln \alpha)$
 $\hat{\alpha} = \text{antiln}(1,2840) = 3.6110$
 $a = e^{(\hat{\alpha} / \beta)}$
 $a = e^{(3.6110 / 0.3621)} = 0.0000471$
 $b = 1 / \beta$
 $b = 1 / 0.3621 = 2.7586$
- $\mu = [(1/a) * \ln P_0]^{1/b}$
 $\mu = [1 / 0.000471] * \ln P_0]^{1/2.7586}$
- $n_{\text{opt}} = (1/D^2 a^2 b^2) \mu^{-2b} [\exp(a\mu^b) - 1]$
 $n_{\text{opt}} = [1/1/0.25^2(0.000471)^2(2.7586)^2] \mu^{2(2.7586)} [\exp(0.0000471\mu^{2.758}) - 1]$
- $EEm = (P_i / n) a^{(-2/b)} b^{(-2)} (1 - P_i)^{-1} [- \ln(1 - P_i)]$
 $EEm = (P_i / n) 0.000471^{(-2/2.7586)} 2.7586^{(-2)} (1 - P_i)^{-1} [-\ln(1 - P_i)^{-1}]$
- Límite de error de estimación (L) = $\pm tEEm$
 Donde, t = valor de la tabal de t-student

De los resultados de la Tabla 9 se concluye que, en base a las proporciones de unidades muestrales ocupadas (P_i) o vacías (P_o) uno puede estimar la media poblacional con su grado de error de estimación.

Tabla 9. Datos y parámetros de MSP-A para modelo de tipo KS.

M	P_i	P_o	μ	EEm	n
1.1	0.03	0.97	10.46	6.85	71
1.7	0.09	0.91	15.73	5.95	24
1.9	0.15	0.85	19.16	5.62	14
2.0	0.20	0.80	21.50	5.46	11
2.1	0.22	0.78	22.35	5.42	10
2.2	0.28	0.72	24.73	5.32	8
2.3	0.37	0.63	27.99	5.26	6
2.5	0.42	0.58	29.71	5.26	6
3.5	0.60	0.40	35.87	5.43	4
5.4	0.73	0.27	40.83	5.82	4

E. Muestreo secuencial en función del tiempo (MSFT)

El punto crucial en el manejo integral de plagas agrícolas es la detección “*a tiempo*” de brotes de plagas y subsecuentemente, la supresión poblacional de las mismas antes de que ocurra un daño económico. Para detectar estos brotes, la costumbre es la toma de muestras del campo de forma frecuente y regular (usualmente semanal) durante toda la estación del año. Sin embargo, para algunas plagas (ej. insectos), este procedimiento desde el punto de vista de la precisión y el costo, posiblemente no sea lo más eficiente para poder basar las decisiones de manejo de la plaga. En el caso de muchas especies de plagas, la población se incrementa de forma rápida y declina antes del fin de la estación. En estos casos, el muestreo muy frecuente (tres veces a la semana) es necesario y en algunas ocasiones se puede suspender completamente el muestreo en otros períodos. Consecuentemente, el diseño adecuado y la eficiencia precisa de un sistema de manejo integral de plagas dependen crucialmente en detectar con máxima precisión el tiempo de hacer el muestreo (frecuencia y terminación precisa de muestreo).

Debido a que la abundancia numérica de las plagas tiene distribuciones características en el tiempo, de la misma manera que en el espacio, se pueden aplicar métodos similares a los que emplean para hacer decisiones de muestreo secuencial en la escala espacial (MSE), en un sentido cronológico y de este manera producir un esquema de muestreo secuencial en base de tiempo (MST).

Los requisitos de MST son los mismos tres que se usan para MSE; la diferencia sería que: **1.** En el caso de MST, se establece el tipo de distribución por cada fecha, comparado con una sola distribución durante todo el muestreo para el caso de MSE. **2.** En lugar de un solo límite de clase para el caso de MSE (m_1 para clase endémica y m_2 para clase epidémica), a lo largo de la estación se establecen estos límites por cada fecha de muestreo para MST.

En MST, las fórmulas generales para las líneas de hacer decisión, después de la "t" ésima muestra están dadas por las siguientes ecuaciones (Pedigo & Van Schaik, 1984).

$$d_{i,t} = h_1 + b_t \quad \text{línea inferior}$$

$$d_{s,t} = h_2 + b_t \quad \text{línea superior}$$

donde, $d_{i,t}$ y $d_{s,t}$ son los números acumulativos ponderados de plaga observada en fecha t para las líneas inferior ($d_{i,t}$) y superior ($d_{s,t}$) respectivamente, comparado con el d_n convencional de MSE que es el número acumulativo no ponderado de plaga para la muestra n , b_t es la pendiente que varía en función de fecha de muestreo, h_1 y h_2 son las intersecciones con la ordenada y se calculan según: $h_1 = \log[b/(1-a)]$, y $h_2 = \log [(1-b)/a]$, donde \log es el logaritmo decimal y a y b son los riesgos mencionados como para el caso de MSE. Las ecuaciones de las líneas de decisión para diferentes tipos de distribución se encuentran en Pedigo & Van Schaik (1984).

En la práctica se usan los siguientes pasos y las ecuaciones para las estimaciones correspondientes:

1. Estimación de h_1 , h_2 y la "k".

$$h_1 = \log [\beta / (1 - \alpha)]$$

$$h_2 = \log [(1 - \beta) / \alpha]$$

Se calcula K por cualquier de los 3 métodos para estimar el parámetro K de binomial negativa.

2. Estimación de los valores de $q_{\text{end}(i)}$ y $q_{\text{epi}(i)}$

$$q_{\text{end}} = p_{\text{end}(i)} + 1$$

$$q_{\text{epi}} = p_{\text{epi}(i)} + 1$$

Donde, $q_{\text{end}(i)}$ y $p_{\text{end}(i)}$ son los valores de q y p para la población baja o endémica en la muestra "i"

q_{epi} y p_{epi} son los valores de q y p para la población alta o epidémica en la muestra "i".
 $P = m / k$, y $q = 1 + p$.

3. Cálculo de a_i :

Distribución al azar (Poisson): $a_i = m_{\text{epi}(i)} / m_{\text{end}(i)}$

Distribución binomial positiva: $a_i = n_i \log(q_{\text{end}(i)} / q_{\text{epi}(i)})$

Distribución binomial negativa: $a_i = \log(q_{\text{end}(i)} / q_{\text{epi}(i)})$

4. Cálculo de b_t según el tipo de distribución:

Distribución al azar (Poisson): $b_t = \sum a_i$

Distribución binomial positiva: $b_t = \sum a_i$

Distribución binomial negativa: $b_t = -K \sum a_i$

Donde,

b_t es la pendiente de la ecuación de la línea inferior (d_{it}) y línea superior (d_{st}).

5. Las ecuaciones para la línea inferior (d_{it}) y superior (d_{st}) son:

$$d_{it} = b_t + h_1$$

$$d_{st} = b_t + h_2$$

Ejemplo para MSFT

Los datos siguientes (Tabla 10) indican las densidades a nivel endémica (m_{end}) y epidémicas (m_{epi}) en las fechas de muestreo.

Tabla 10. Datos de las densidades m_{end} y m_{epi} en siguientes muestras.

Periodo de muestreo	m_{end}	m_{epi}
1	2	4
2	3	5
3	4	16
4	7	18
5	8	23
6	6	38
7	3	24
8	2	26
9	1	25

En base a los datos de la Tabla 20 y las ecuaciones arriba, diseñar un esquema de MSFT (Tabla 11).

Solución

Los valores de las siguientes variables: q_{end} , q_{epi} , a_i , b_i , d_{it} , y d_{st} se encuentran en la Tabla 11.

Tabla 11. Cálculo de las variables para MSFT.

i	m_{end}	m_{epi}	q_{end}	q_{epi}	a_i	d_{it}	d_{st}
1	2	4	2.51	4.03	-0.205	0	1.22
2	3	5	3.27	4.79	-0.165	0	1.44
3	4	16	4.03	13.12	-0.513	0.20	2.11
4	7	18	6.30	14.64	-0.366	0.68	2.59
5	8	23	7.06	18.42	-0.416	1.23	3.14
6	6	38	5.54	29.79	-0.730	2.19	4.10
7	3	24	3.27	26.76	-0.913	3.38	5.29
8	2	26	2.51	20.70	-0.915	4.58	6.99
9	1	25	1.76	19.94	-1.055	5.97	7.88

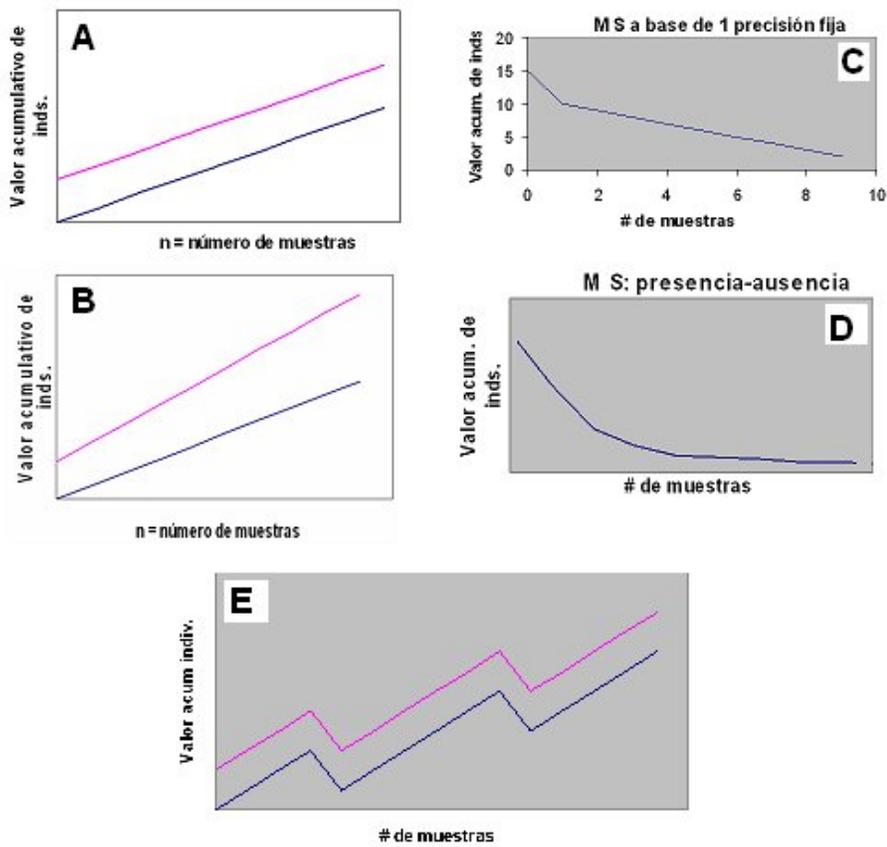


Figura 1. Modalidades de MS (A. MSC, B. MSDC, C. MSPF, D. MSP-A y E. MSBT).

Conclusiones

Si la distribución normal no se ajusta a los datos, como es el caso de la mayoría de los procesos o fenómenos en el área de ciencias naturales, sociales, psicológicas, etc., entonces, se debe acudir a los métodos alternos del muestreo. Ahora bien, si el costo del muestreo es verdaderamente un punto crítico (limitante) en el desarrollo del trabajo y el objetivo primordial de la investigación es clasificar rápidamente los elementos de la población (procesos, eventos, objetos, etc.) en diferentes grupos de interés, el muestreo secuencial en sus diferentes modalidades (cinco) es el método a seleccionar.

Referencias

- Badii, M.H. & J.A. McMurtry. 1990. Field experiments on predation, dispersión, regulation and population changes. *Publ. Biol.* 4(1-2): 43-48.
- Badii, M. H., A. E. Flores, S. Flores y S. Varela. 1994_a. Statistical description of population distribution and fluctuation of citrus rust mite (Acari: Eriophyidae) on orange fruit in Nuevo Leon, Mexico. *Biotam*, 6(1): 1-8.
- Badii, M.H., A.E. Flores, S. Flores & S. Varela. 1994_b. Comparative estimation of distribution statistics of citrus rust mite (Acari: Eriophyidae) on leaves of three different orange orchards in Nuevo León, Mexico. *Biotam*, 6(1): 9-16.
- Badii, M.H., A.E. Flores, R. Foroughbakhch & H. Quiroz. 1995_a. Análisis conceptual de muestreo. Pp. 123-136. En: W. Rosa (ed.). VI Curso Nacional de Control Biológico. SMCB, Tapachula, Chiapas. México.
- Badii, M.H., A.E. Flores, R. Torres & H. Quiroz. 1995_b. Muestreo y evaluación económica de las plagas. Pp. 1-13. En: H. Fuente (ed.). Curso Internacional sobre Manejo de Huertas de Cítricos. SAGAR, INIFAP, CIRNE., General Allende, N. L., México.
- Badii, M.H., A.E. Flores, S. Varela, S. Flores & R. Foroughbakhch. 1996a. Dispersion indices of citrus rust mite (Acari: Eriophyidae) on orange in Tamaulipas, Mexico. Pp. 17-20. En: R. Mitchel, D. Horn, G. R. Needham y W. C. Welbourn (eds.). *Acarology IX: Volume 1, Proceedings. Section I: Behavior and Physiological Ecology.* Ohio Biological Survey, Columbus, Ohio.
- Badii, M.H., A.E. Flore, R. Foroughbakhch, H. Quiroz & R. Torres. 1996_b. Ecología de manejo integrado de plagas (mip) con observaciones sobre control microbiano de insectos. Pp. 21-49. En: L.J. Galan Wong, C. Rodreiguez-Padilla y H. Luna-Olvera (eds.). *Avances Recientes en la Biotecnología en Bacillus Thuringiensis.* Ciencia, Universitaria no. 2. UANL.
- Badii, M.H., A.E. Flores, S. Flores & R. Foroughbakhch. 1998. Population dynamics of citrus mites in northeastern Mexico. Pp. 275-280. En: G. Needham, R. Mitchell, D. Horn and W. C. Welbourn (eds.). *Acarology IX: Vol. 2, Symposia.* Ohio Biological Survey, Columbus, Ohio.
- Badii, M.H., A.E. Flores, R. Foroughbakhch & H. Quiroz. 2000. Fundamentos de muestreo. Pp: 129-144. In: M.H. Badii, A.E. Flores & J. Galán (eds.). *Fundamentos y Perspectivas de Control Biológico.* UANL, Monterrey.
- Bliss, C.I. & A.R.G. Owen. 1958. Negative binomial distribution with a common K. *Biometrika*, 45: 37-58.
- Cochran, W.G. 1977. *Sampling techniques.* Second Ed. Wiley, New York.

- Croft, B.A., S.M. Welch & M.J. Dover. 1976. Dispersion statistics and sample size estimates for populations of the mite species *Panonychus ulmi* and *Amblyseius fallacis* on apple. Environ. Entomol. 5: 227-234.
- Gerard, D.J. & H.C. Chiang. 1970. Density estimation of corn rootworm egg populations based upon frequency of occurrence. Ecology, 51: 237-245.
- Green, R.H. 1970. On Fixed precision level sequential sampling. Res. Popul. Ecol. 12: 249-251.
- Iwao, S. 1968. A new regression method for analyzing the aggregation pattern of animal populations. Res. Popul. Ecol. 10: 1-20.
- Iwao, S. 1975. A new method of sequential sampling to classify populations relative to a critical density. Res. Popul. Ecol. 16: 281-288.
- Kogan, M. & D.C. Herzog. 1980. Sampling methods on soybean entomology. Springer-Verlag, N.Y.
- Kuno, E. 1969. A new method of sequential sampling to obtain the population estimates with a fixed level of precision. Res. Popul. Ecol. 11: 127-136.
- Morris, R.F. 1955. The development of sampling techniques for forest insect defoliators, with particular reference to the spruce budworm. Can. J. Zool. 33: 225-294.
- Nachman, G. 1981. A mathematical model on the functional relationships between density and spatial distribution of a population. J. Anim. Ecol. 50: 453-460.
- Nachman, G. 1984. Estimates of mean population density and spatial distribution of *Tetranychus urticae* (Acarina: Tetranychidae) based on the proportion of empty sampling units. J. Appl. Ecol. 21: 903-913.
- Pedigo, L.P. & J.W. van Schaik. 1984. Time-sequential sampling: a new use of the sequential probability ratio test for pest management decisions. Bull. Entomol. Soc. Am. 30: 32-36.
- Southwood, T.R.E. 1966. Ecological methods with particular reference to the study of insect populations. Chapman y Hall, N. Y.
- Southwood, T.R.E. 1978. Ecological methods with particular reference to the study of insect populations. 2^a ed. Chapman y Hall, N.Y.
- Taylor, L.R. 1961. Aggregation, variance and the mean. Nature, 189:732-735.
- Taylor, L.R., I.P. Woiwood & N. Perry. 1979. The density dependence of spatial behaviour and the rarity of randomness. J. Anim. Ecol. 47: 383-406.
- Troii, T. 1971. The development of quantitative occurrence prediction of infestation by the rice stem-borer, *Chilo suppressalis* in Japan. Entomophaga, 16: 193-207.
- Wilson, L.F. & D.J. Gerard. 1971. A new procedure of rapidly estimating pine sawfly (Hymenoptera: Diprionidae) population levels in young pine plantations. Can. Ent. 103: 1315-1322.

*** Acerca de los Autores**

Badii M.H. Es profesor investigador del área de posgrado, UANL, México, mhbadiiiz@gmail.com

Guillen A. Es profesora investigadora del área de posgrado, UANL, México

Valenzuela J. Es profesor investigador del área de posgrado, UAAAN, Saltillo, Coah., México

Cerna E. Es profesor investigador del área de posgrado, UAAAN, Saltillo, Coah., México
UANL, San Nicolás, N.L., México, mhbadiiiz@gmail.com, UAAAN, Saltillo, Coah., México