

ECOLOGÍA GENERAL

Métodos de muestreo

Introducción

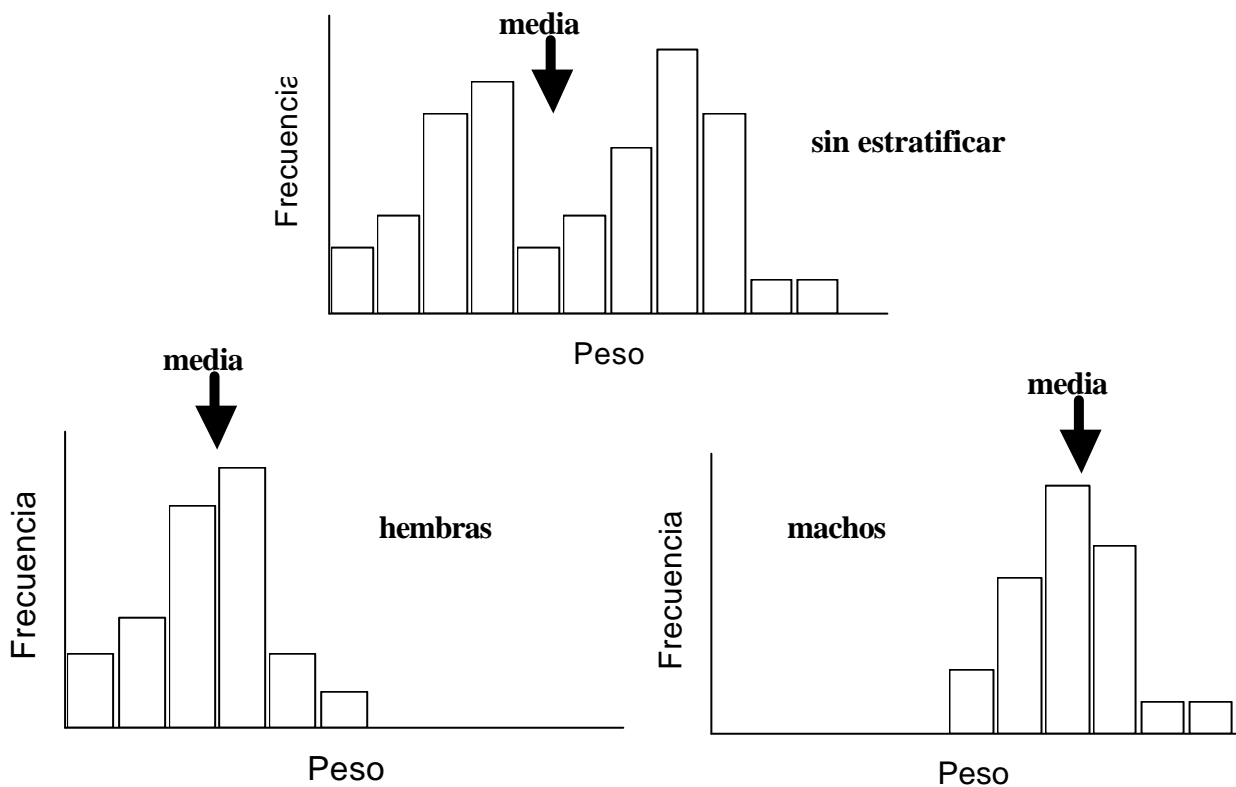
En ciencias biológicas, la mayor parte de las investigaciones son cuantitativas, con observaciones de hechos numéricos llamados datos. Las entidades biológicas son contadas o medidas y es evidente que debe existir algún método objetivo para tomar, presentar y analizar estos datos. Una distinción importante es la que se puede hacer entre *observaciones* y *experimentos*. Las observaciones son datos obtenidos sin haber manipulado el sistema y son utilizadas para describir regularidades u ordenamientos biológicos (patrones). Los datos experimentales en cambio son datos obtenidos a partir de manipulaciones de las variables que se piensa pueden estar determinando cierto patrón. Los mismos análisis estadísticos pueden realizarse a partir de observaciones o experimentos. Sin embargo, los procesos que determinan los patrones biológicos observados sólo pueden ser explicados mediante un experimento bien diseñado. En cambio, resultados similares obtenidos a partir de observaciones pueden ser causados por el efecto de cierta variable desconocida o no detectada. Esto sugiere que el conocimiento del mundo debería estar basado solamente en experimentos, sin embargo esto no es enteramente posible y hay ciertas preguntas para las cuales no es posible diseñar un experimento, pero se pueden realizar observaciones dirigidas que sirvan para contestarlas.

La estimación de algún parámetro de una población involucra tres elementos principales: (1) el parámetro de interés; (2) la población que está siendo muestreada; y (3) el método de muestreo. Los parámetros de una población son valores numéricos que resumen información de los ítems de una población. La población es la colección de todos los ítems que son de interés. Existen poblaciones finitas, como los alumnos del presente año de la cátedra de ecología (cada alumno es un ítem), y poblaciones infinitas, como el resultado que se puede obtener de tirar una moneda al aire (siempre se la puede tirar una vez más, cada tirada es un ítem). En éste último ejemplo es lógico pensar que para conocer la frecuencia con la que cae cada lado de la moneda, no se puede tener acceso a la población entera. De manera similar, en el caso de poblaciones finitas, no siempre es posible tener acceso a toda la población (en términos de tiempo, trabajo o dinero). Los métodos de muestreo permiten estimar parámetros poblacionales a partir del estudio de una fracción de la misma. Existen diferentes métodos de muestreo y la preferencia de uno sobre otro dependerá del objetivo del estudio, del parámetro a estimar y de la población (distribución espacial, temporal, movilidad, etc.). Esto quiere decir que el método a elegir depende de la pregunta que se quiera contestar.

Muestreo al azar simple: Un muestreo al azar simple es aquel en el cual todos los ítems de la población tienen la misma probabilidad de ser muestreados. Si los ítems (individuos, unidades muestrales, etc) de una población de tamaño N pueden numerarse desde 1 hasta N , una forma de llevar a cabo un muestreo aleatorio simple consiste en seleccionar n de estos números de manera aleatoria (por ejemplo utilizando una tabla de números al azar). Si se está trabajando con muestreos espaciales (la unidad muestral no es un individuo sino un espacio) se puede dividir el área en grillas y seleccionar las casillas a muestrear de manera aleatoria.

Muestreo estratificado: Consiste en subdividir a priori la población en subunidades o estratos y luego hacer un muestreo aleatorio dentro de cada estrato. La subdivisión

responde a un conocimiento a priori de la existencia de dichos estratos y de que el parámetro a medir está influido por ellos. La no estratificación llevaría a que datos de dos poblaciones distintas sean tomados como si fueran de una única población lo que llevaría a un aumento en la varianza y a un sesgo en la estimación del parámetro. Si se piensa en el peso de los elefantes marinos (los machos son mucho más grandes y pesados que las hembras) se pueden definir a priori dos estratos (machos y hembras) y hacer un muestreo de individuos al azar dentro de cada estrato. Como se observa en el gráfico, no estratificar llevaría a estimar un peso promedio intermedio que sólo algunos individuos presentan, con mucha varianza alrededor de él. Al separar entre sexos, se estiman dos pesos promedios que son representativos de las distribuciones de peso de cada sexo.



Si se quiere estimar la densidad de peces en la Laguna de los Padres, se puede asumir que dicha densidad no es homogénea y existen estratos que influyen dicha densidad (presencia de vegetación, profundidad), en cuyo caso sería mas adecuado hacer un muestreo aleatorio dentro de cada estrato. De manera similar, la cobertura de algas en el intermareal rocoso puede estar influida por el tiempo de exposición al aire (altura de marea) y se podría separar en diferentes estratos (submareal, intermareal bajo e intermareal alto, por ejemplo). La estratificación puede ser tanto espacial como temporal. En la laguna de Mar Chiquita, por ejemplo, se pueden encontrar especies de aves que permanecen todo el año y otras que son migratorias y sólo se encuentran durante el verano. Si se quiere estudiar la diversidad de aves en la laguna sería lógico estratificar en estaciones del año y luego tomar muestras aleatoriamente dentro de cada estación.

Muestreo sistemático: Cuando una población puede ser numerada en orden o cubre un área espacial bien definida, se puede muestrear de manera sistemática muestreando a intervalos regulares (tomando los individuos pares de la población o muestreando cada cierta distancia). Para que este tipo de muestreo permita calcular errores de muestreo o realizar comparaciones entre poblaciones, se debe cumplir el supuesto que los ítems estén distribuidos al azar. Este tipo de muestreo es generalmente más fácil de llevar a cabo que los muestreos aleatorios.

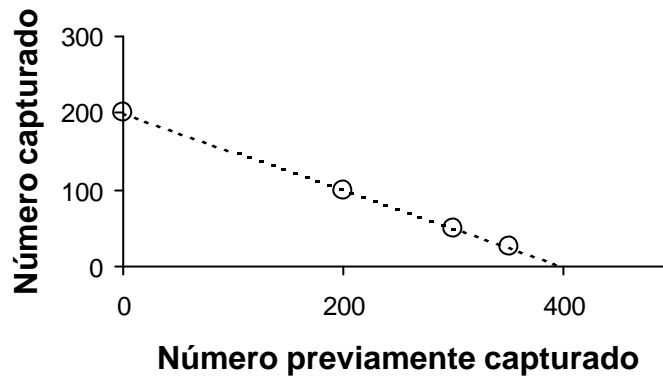
Como se puede ver a lo largo de los ejemplos, las muestras pueden ser individuos, áreas espaciales, fracciones de tiempo, etc., todo depende de la pregunta que se quiera responder y del sistema donde se esté trabajando. Existen situaciones donde son necesarias aproximaciones o diseños muestrales que son de gran utilidad y son tal vez menos intuitivos. El método de transectas es uno de ellos y presenta muchas ventajas prácticas en ciertos casos. Si se quisiera conocer la densidad de aves en campos cultivados con dos especies de cereal distintos, definir áreas muestrales cuadradas o rectangulares para luego contar los individuos dentro puede ser poco práctico. Una alternativa es caminar un número previamente definido de metros y contar los individuos que se detectan a lo largo de la transecta. Es necesario definir una distancia lateral hasta donde se considera que es posible detectar a los individuos. La transecta puede entonces ser vista como un área muestral rectangular. Cuando existe un gradiente ambiental muy marcado (como por ejemplo desde la línea de marea baja hasta la de marea alta en un intermareal) se puede realizar una transecta a lo largo del gradiente y tomar muestras a intervalos regulares dentro de cada transecta. Es importante tener en cuenta que en estos casos, la unidad muestral es la transecta (desde el punto de vista estadístico corresponde a una única observación). El método de marca y recaptura o captura y recaptura ha sido desarrollado para estimar el número de individuos en poblaciones de animales móviles y esta basada en simples reglas de proporciones y probabilidades. Si se quisiera calcular el número total de peces de la Laguna de los Padres, se podría pasar una red capturando un número n_1 de individuos, esos individuos pueden ser marcados con algún método y liberados nuevamente. Luego de un período prudencial de tiempo (días o semanas en el ejemplo) se puede pasar la red nuevamente, en este caso capturando n_2 individuos. Dentro de éstos, habrá un número m de individuos que habían sido marcados previamente. La proporción de peces marcados en toda la laguna debería ser igual a la de peces marcados en la segunda muestra así que la cantidad de peces de toda la laguna (N) será:

$$N = n_1 * n_2 / m$$

Los supuestos de este método son bastante intuitivos: (1) La población es cerrada, no hay emigración ni inmigración de individuos hacia y desde otras áreas; (2) Todos los animales tienen la misma posibilidad de caer en la primer muestra y la segunda muestra es una muestra al azar de la población; (3) La marca no afecta la posibilidad de ser capturado o la supervivencia, ni se borra o pierde.

Una alternativa a la estimación del tamaño poblacional por captura y recaptura es el método de remoción (método de Hayne o de Leslie y Davis) y puede ser aplicado cuando no se pueden cumplir los supuestos del método anterior o cuando no es posible capturar vivos a los individuos. El método de remoción consiste en capturas sucesivas de miembros de la población. Si la población es cerrada, a medida que capturamos y removemos individuos, ésta ira disminuyendo y ante un mismo esfuerzo de captura se atraparán menos individuos cada vez. Suponga que realiza este procedimiento para estimar el número de renacuajos en un pequeño estanque, pasa una red y captura 200

individuos, la segunda pasada de red captura 100 individuos, la tercera 50 y la cuarta 25 individuos. Si se grafica el número de individuos capturados en cada pasada vs el número capturado en todas las pasadas anteriores se obtendrá el siguiente gráfico:



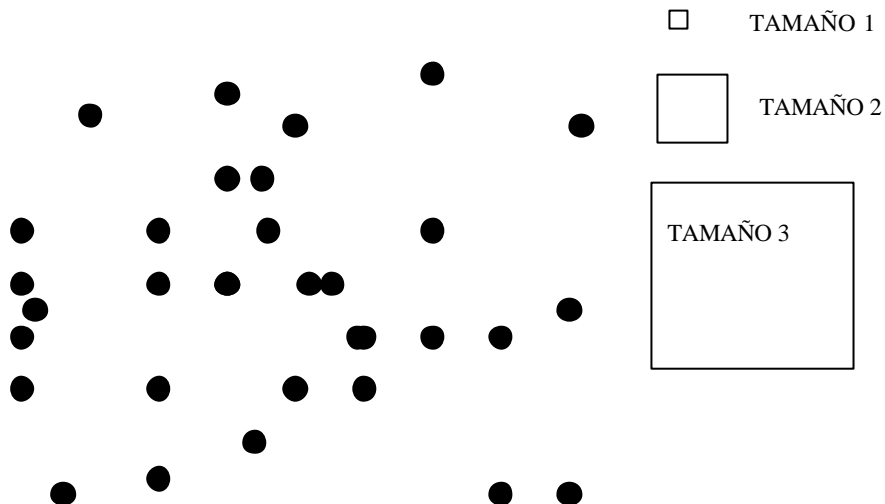
Teóricamente, cuando se hayan extraído todos los individuos, la captura siguiente será 0 y el número previamente calculado será precisamente el número total de renacuajos en el estanque. Si realizamos una regresión lineal con los datos, tendremos una estimación de los parámetros a y b de la recta $Y = a + bX_i$, como al capturar a todos $Y = 0$, entonces (reagrupando) $X_{total} = -a/b$. Los supuestos de este método son: (1) todos los individuos tienen la misma probabilidad de ser capturados; (2) Excepto por las capturas, la población se mantiene estable; (3) el esfuerzo de captura es igual en cada captura; y (3) La probabilidad de capturar a un individuo es la misma para cada período de captura.

Número de unidades de muestreo:

Si muchas muestras al azar de tamaño n son tomadas independientemente de una población y la media muestral es calculada en cada muestra, probablemente tendrán diferentes valores. Si el nivel de variación es "pequeño", entonces diferentes muestreos darán mas o menos el mismo resultado. Si es "grande", las diferentes muestras arrojaran valores muy diferentes. Para realizar inferencias estadísticas a partir de muestras (para estimar parámetros poblacionales o comparar diferentes situaciones), el grado de confianza aumenta al aumentar el número de réplicas. Esto es especialmente cierto para estimadores insesgados (la ley de los grandes números dice que cuando el número de muestras tiende a infinito, el valor estimado tiende al valor del parámetro real por lo que disminuye la varianza) de manera similar, el poder estadístico (1- la probabilidad de aceptar la hipótesis nula cuando es realmente falsa) de los tests aumenta. Esto genera un compromiso entre la precisión con la que queremos realizar las inferencias y el trabajo o costo que lleva cada muestra. Existen métodos para calcular el tamaño muestral necesario para alcanzar determinado grado de precisión o poder estadístico que dependen del parámetro a estimar o del test a realizar y se pueden encontrar fácilmente en los libros de texto de estadística.

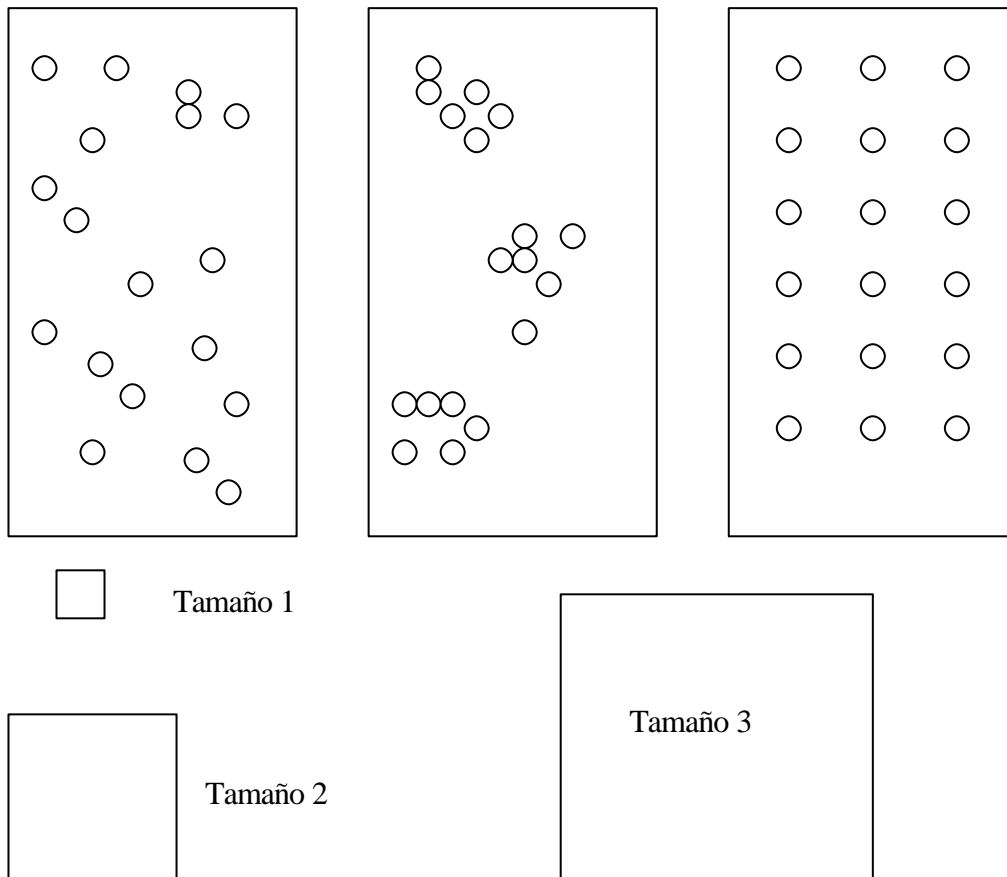
Tamaño de las unidades de muestreo:

Cuando se están utilizando áreas muestrales es importante tener en cuenta el tamaño de los organismos y tener una noción de su densidad y su distribución espacial. Si se observa el siguiente gráfico, donde cada círculo representa a un individuo, obviamente el área de muestreo no puede ser nunca menor al tamaño de los círculos (tamaño 1). Con el tamaño 2, probablemente obtengamos muchos valores de 0 y algunos de 1 o a lo sumo 2 individuos por muestra mientras que el tamaño 3 parece más adecuado. Una vez más, a pesar de que existen ciertas reglas lógicas que ayudan a elegir el tamaño muestral, éste depende de la pregunta que queramos contestar. Una forma práctica es aumentar el tamaño del área muestral progresivamente y graficar la varianza vs el tamaño de la muestra. A tamaños pequeños la varianza será muy grande y pequeños cambios en el tamaño tendrán grandes efectos en ésta mientras que a partir de cierto tamaño, la disminución en la varianza al aumentar el área será muy chica. Esto se nota en la figura cuando la curva empieza a llegar a un "plateau".



Lo mismo se aplica para la distribución espacial de los individuos, si están agrupados y el área de muestreo elegida es similar o más pequeña que los agrupamientos entonces se obtendrán datos de densidades nulas y altas ya que las muestras pueden caer dentro de un agrupamiento o fuera. Si la distribución espacial se presenta a la misma escala que las unidades de muestreo, es posible detectarla. Las disposiciones al azar pueden ser descriptas con una distribución de Poisson ($s^2 = \mu$); la disposición agrupada con una binomial negativa ($s^2 > \mu$ ya que las muestras pueden caer en zonas con alta o nula densidad y la varianza será muy grande); y la regular con una binomial positiva ($s^2 < \mu$ ya que al estar distribuidos regularmente la varianza será mucho menor). Con tamaños de unidades de muestreo a escalas mayores que el patrón de distribución, se tendrá una buena estimación de la densidad pero no se podrá detectar dicho patrón.

En la figura siguiente se puede ver que los patrones de distribución pueden ser detectados con un tamaño de unidad de muestreo intermedio. El tamaño mayor será bueno para estimar la densidad pero no permite detectar dichos patrones.



Cuando se está trabajando con la diversidad o riqueza de especies de cierta comunidad, a medida que aumenta el tamaño de la muestra aumenta el número de especies detectadas en ésta hasta un tamaño donde la comunidad está bien representada, luego del cual el número de especies se mantiene constante. Un método común para determinar el tamaño óptimo es aumentar progresivamente el tamaño de muestreo y graficar número de especies vs tamaño de muestreo y determinar visualmente el tamaño a partir del cual la variación en número de especies es muy pequeña.

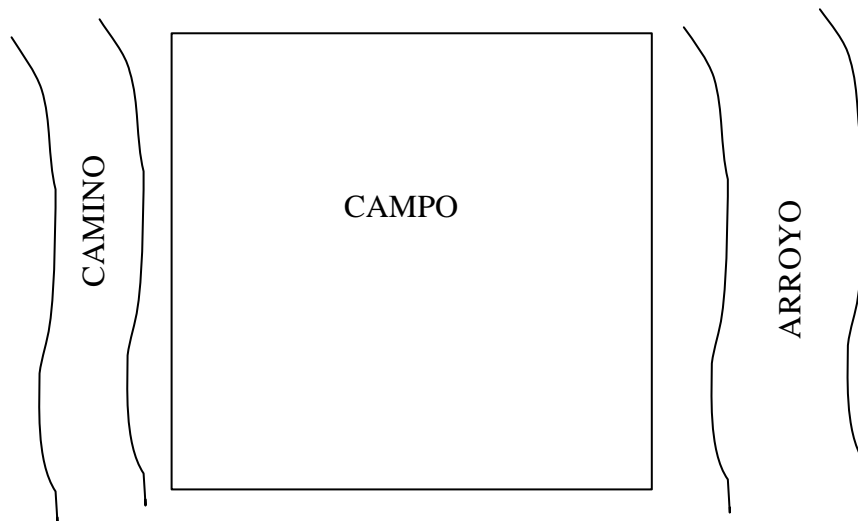
Pseudoreplicas:

Tanto cuando se trabaja con muestreos como con experimentos, es muy importante tener en claro cuáles son las unidades de muestreo ya que de esto dependerá que las inferencias que se hagan sean correctas o incorrectas. Un supuesto básico para hacer inferencias es que las unidades de muestreo (datos) deben ser totalmente independientes entre sí, esto significa que el valor de un dato no es influido ni influye el

valor de otro. Sin embargo este supuesto es frecuentemente violado. Si se quiere determinar la cantidad de insectos por hoja de cierta especie de planta, por ejemplo, y se cuenta la cantidad de insectos en 3 hojas de una planta, 5 de otra y 7 de otra, será un error pensar que se tienen 15 ($3+5+7$) datos ya que las hojas de una misma planta tienen más en común entre ellas que con las de otra planta (si una planta está infectada o ha sido colonizada, es probable que presente altas densidades en todas sus hojas). Lo correcto sería realizar un promedio por planta y luego trabajar con esos promedios como dato (o sea 3 datos). De manera similar, al realizar experimentos, los datos independientes provendrán de cada unidad experimental. Si se hacen exclusiones de herbívoros para estudiar su efecto en la producción o supervivencia de cierta especie de plantas, se pueden construir cajas de alambre tejido que impidan el ingreso de los herbívoros. Dentro de cada caja, se podrían tomar datos de varias plantas individuales, pero utilizar esto como réplicas reales es incorrecto, ya que la unidad experimental es la caja. Los valores medidos sobre las plantas de una misma caja no son independientes.

EJERCICIOS:

- 1) Un productor rural ha sembrado trigo en una hectárea de campo que está delimitada a un lado por un camino y al otro por un arroyo.



Luego de un tiempo, empieza a preocuparse ya que al ver su campo desde el camino le parece que la producción no es muy buena y que no vale la pena gastar dinero en los tractores para cosecharlo. Para estar seguro decide subdividir su campo en 100 plots de 10×10 m y cosechar los 10 plots más cercanos al camino ya que son los más fáciles y accesibles. Con los datos obtenidos calcula una producción promedio por plot, lo que le permite estimar la producción total de su campo. Luego de realizar todo esto decide abandonar su campo y dedicarse a otra cosa pero luego de un tiempo se pone a pensar si su método fue correcto. El productor pasa sus noches pensando métodos alternativos y luego de un tiempo decide que tenía 3 alternativas más: (1) Cosechar totalmente al azar 10 plots; (2) Estratificar horizontalmente (con respecto al dibujo) y muestrear al azar 1 plot de cada fila y; (3) estratificar verticalmente y muestrear al azar 1 plot de cada columna.

En la siguiente grilla se encuentran los datos de la producción de cada plot (kg de grano):

7	9	12	15	15	18	19	21	22	25
8	10	12	17	18	18	20	20	23	28
6	8	10	15	19	15	20	23	22	25
8	8	9	15	14	21	21	22	24	26
7	10	13	16	18	20	19	22	23	26
9	11	11	13	18	17	17	21	24	27
8	8	9	15	15	19	19	20	24	24
6	9	11	14	16	16	22	20	22	26
6	9	13	12	16	19	20	22	23	28
9	10	12	16	18	18	21	19	24	26

Estime la producción total del campo por cada método (inclusive el inicial) y compárelo con el valor real de producción del campo. Discuta que método considera más adecuado para este sistema. Piense y discuta otro método de muestreo adecuado para esta problemática.

- 2) Unos ecólogos han planteado la hipótesis de que la presencia de herbívoros aumentan la diversidad de especies vegetales. Para poner a prueba su hipótesis deciden realizar muestreos y experimentos. Uno de los muestreos consiste en muestrear al azar cuadrados de un tamaño a determinar en zonas con y sin presencia de herbívoros donde se contará el número de especies vegetales por muestra. Para determinar el área de muestreo deciden realizar un muestreo previo de donde se obtienen los siguientes datos:

Área de muestreo (m ²)	Nº especies
0.1	2
0.2	9
0.4	17
0.8	23
1.6	22
3.2	23

6.4
12.8

22
24

¿Qué área de muestreo es la recomendable? ¿Cuáles pueden ser las desventajas prácticas de tomar tamaños grandes?

3) En los siguientes casos, ¿cuántas réplicas se tiene?

3.1- Se midió el ancho de las 2 quelas de 10 cangrejos macho

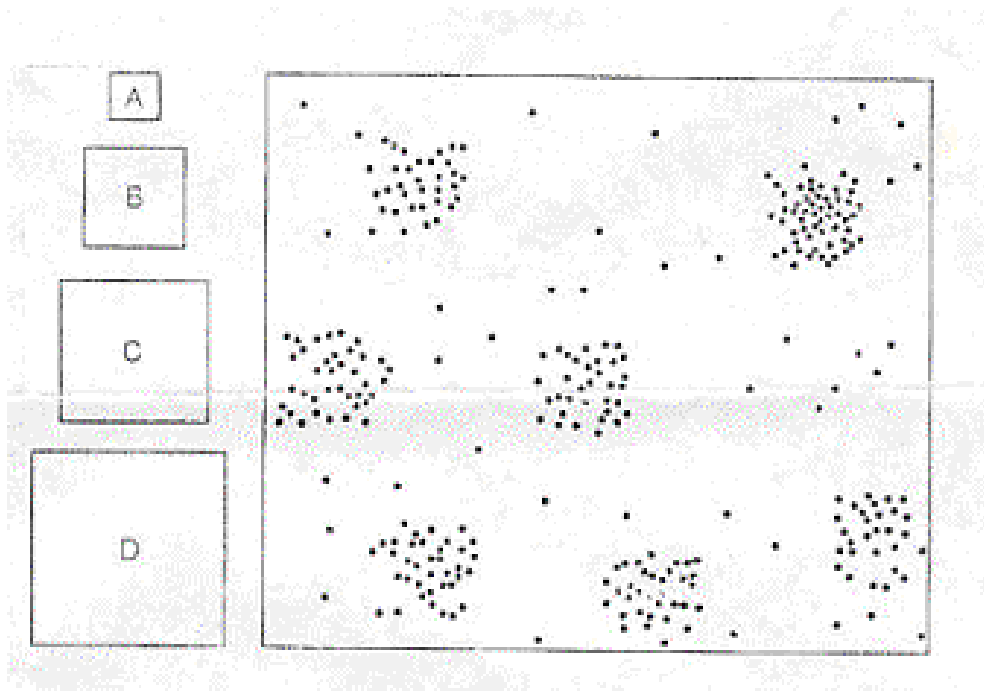
3.2- Se realizaron 20 transectas con 5 estaciones de muestreo cada una

3.3- Se realizaron 10 exclusiones de herbívoros de 10 x 10 m y se tomaron 10 plantas por exclusión a las cuales se le midió la biomasa.

4) Qué tipo de patrón de disposición espacial detectaría luego de muestrear la población usando como unidades los cuadrados de tamaño A, B, C y D?

Discuta la relación entre detección de patrones de disposición espacial y escala de análisis.

Ejemplifique que tipo de factores pueden estar generando estos patrones en distintas escalas.



Bibliografía sugerida:

Brower, J.E., Zar, J.H., von Ende, C.N., 1998. Field and laboratory methods for general ecology. McGraw-Hill, Boston, MA

Hulbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187-211

- Manly, B.F.J., 1992. The design and analysis of research studies. Cambridge University Press, UK
- Underwood, A.J., 1997. Experiments in ecology: Their logical design and interpretation using analysis of variance. Cambridge University Press, UK