

CURSO DE ESTADÍSTICA GENERAL

CASO 1 (continuación)

Influencia de la densidad del cultivo sobre el crecimiento de plantas de maíz (*Zea mays* L.): Estimación de la media y de la varianza de los rendimientos por planta en cultivos con diferente densidad



Introducción

Al comienzo del curso de Estadística General describimos las distribuciones de frecuencias de varias características de las plantas de maíz cultivadas en parcelas con diferente densidad. Nuestro propósito era concluir acerca de la existencia de mayor o menor plasticidad de caracteres fenotípicos de las plantas, tales como la altura, el grosor del tallo, el número de hojas, etc., frente a diferencias en el ambiente debidas a la densidad del cultivo. En aquel momento simplemente medimos y registramos los caracteres señalados en 20 plantas ubicadas en una hilera de cultivo y usamos los datos generados para ejercitar la caracterización de distribuciones de frecuencias mediante tablas, gráficos, medidas de posición y medidas de dispersión.

Ahora que ya hemos revisado las bases de la teoría de probabilidad y de variables aleatorias y que hemos tomado contacto con el procedimiento de inferencia estadística que sirve para estimar la media poblacional con un riesgo de error controlado, estamos enterados de que para aplicarlo necesitamos generar los datos mediante el procedimiento de muestreo aleatorio. Por tal motivo, regresaremos a las parcelas de maíz para obtener muestras aleatorias y generar dichos datos apropiados.

En esta oportunidad, nos concentraremos en una sola característica de las plantas, su producción de grano. Mediremos esta producción como la suma de los largos de las espigas granadas que tiene cada planta (una medida no destructiva que permitirá a los estudiantes de varias comisiones trabajar en una misma parcela). Es decir que el objetivo central de este ejercicio es estimar las medias poblacionales de los largos totales de espiga granada por planta en dos poblaciones, una integrada por las plantas de una parcela con densidad de cultivo mayor y la otra integrada por las plantas de una parcela con densidad de cultivo menor.

Procedimiento

Los estudiantes se organizarán en grupos de 3 para cumplir con el objetivo propuesto para este ejercicio. El procedimiento a ejecutar es el siguiente:

- (1) Obtener una muestra aleatoria de 15 plantas en cada parcela (ver instrucciones más abajo).

- (2) A cada planta seleccionada al azar, medirle el largo total de espigas granadas como la suma de los largos de los sectores con grano de cada uno de los choclos que tenga.
- (3) Con los datos generados, calcular estimaciones puntuales de las medias y de las varianzas poblacionales correspondientes a cada parcela.
- (4) Construir un intervalo de confianza para la media poblacional de los largos de espiga ganada de las plantas de cada parcela y compararlos.
- (5) Combinar los datos del grupo con los de otro grupo para conformar muestras de tamaño 30 y a partir de ellas construir nuevos intervalos de confianza.
- (6) Considerar críticamente la precisión de los intervalos de confianza construidos y determinar con qué tamaños de muestra se obtendría una mejora satisfactoria.

Instrucciones para generar los datos de muestras aleatorias

El procedimiento a seguir para producir las muestras aleatorias debe otorgarle a cada planta de la correspondiente población de referencia igual probabilidad de resultar incluida en la muestra. Una alternativa para esto sería la propuesta en el capítulo 4 del texto (numerar todas las plantas de dicha población y luego sortear entre ellas mediante un dispositivo adecuado). Esta alternativa es, por supuesto, válida, pero resulta demasiado engorrosa para aplicarla en este caso. Por tal motivo, utilizaremos la siguiente aproximación que es más sencilla y produce resultados más o menos similares. Para seleccionar cada planta a incluir en la muestra:

1. Sortearemos primero una de las 5 hileras centrales de la parcela (descartamos de la población de referencia las plantas que están en las hileras de los márgenes)
2. Sorteamos un número de orden (entre 1 y 40 para la parcela de alta densidad y entre 1 y 20 para la parcela de baja densidad).
3. En la hilera que salga sorteada en identificamos la planta correspondiente al número que salga sorteado en 2 contándolas a partir del borde norte de la parcela (si el número resulta mayor que el número de plantas en la hilera lo descartamos y comenzamos de vuelta en 1).

¿Cómo sortear hileras y plantas?

Una herramienta sencilla para sortear las hileras y las plantas dentro de las hileras es la que representa la función **aleatorio()** de Excel o, alternativamente la función **Ran#** de una calculadora. Para ejecutar la función **aleatorio()** hay que escribir =aleatorio() en una casilla de una planilla Excel y luego apretar Enter. La ejecución de la función **Ran#** puede variar un poco entre calculadoras pero en general consiste simplemente en apretar la tecla que dice **Ran#** o en apretar esa tecla y luego apretar la que dice =. Cada vez que ejecutamos una de estas funciones producimos una realización de una variable “aleatoria” con distribución uniforme en el intervalo (0, 1).

- Para sortear cada hilera, tomamos una nueva realización que obtengamos de la variable “aleatoria” con distribución uniforme en el intervalo (0, 1), la multiplicamos por 5 y redondeamos el resultado al primer entero consecutivo.
- Para sortear cada planta dentro de una hilera, tomamos otra nueva realización de la variable “aleatoria” con distribución uniforme en el intervalo (0, 1) y la multiplicamos por 40 (o por 20 si se trata de la parcela con baja densidad) y redondeamos el resultado al primer entero consecutivo.