

TEMA 3. Población y muestra

Alicia Nieto Reyes

BIOESTADÍSTICA

Conceptos de población y muestra

Objetivo principal de la estadística: inferir características de un grupo de datos mediante el análisis de un pequeño subgrupo

Ejemplo: queremos saber si el uso del teléfono móvil en niños/as menores de 10 años es perjudicial en su desarrollo

- 1 El grupo serían los niños/as menores de 10 años
 - 2 Como no se puede hacer el estudio sobre todos los niños/as de 10 años tomamos un subgrupo
- denominamos **población** al grupo
 - denominamos **muestra** al subgrupo

Nota: existen poblaciones pequeñas, como la de las personas que han viajado a la luna

Elección de la muestra

Una muestra se podría tomar como quisiésemos; pero no si queremos que sea:

- representativa de la población
- que nos permita obtener conclusiones válidas sobre la población

Tamaño de la muestra = n° de individuos en la muestra

La utilidad de una muestra no depende tanto de su tamaño como de que la hayamos elegido correctamente

Métodos más usuales:

- Muestreo Aleatorio Simple
- Muestreo Sistemático
- Muestreo Estratificado
- ...

Muestreo Aleatorio Simple (m.a.s.)

Tenemos una población de formada por diabéticos de tamaño 5000 y queremos elegir una muestra de tamaño 100

Teóricamente:

- 1 Tomamos 5000 tarjetas, cada una asociada a uno de los elementos de la población
- 2 Barajamos correctamente las 5000 tarjetas
- 3 Tomamos las 100 primeras, que nos dirán cuales son los elementos de la muestra

En la práctica:

- Utilizamos un generador de números aleatorios en el ordenador
- Utilizamos una tabla de números aleatorios (antiguamente o si no tenemos un ordenador a mano)

Muestreo Sistemático

m.a.s. es el método por excelencia

Pero existen otros que en ciertos casos lo aventaja:

En el **Muestreo Sistemático** se toma la muestra con mayor comodidad:

- 1 Elegimos un número entre 1 y 50 aleatoriamente;
Si el tamaño poblacional es 5000 y el muestral es 100 entonces $5000/100=50$
- 2 Supongamos que el número entre 1 y 50 salió 34
Entonces la muestra es 34, $34+50$, $34+2*50$, $34+3*50$,..., $34+99*50$

Nota que la muestra obtenida está ordenada

Problema del método: Que vayamos a realizar un estudio determinado, p.ej. de diabetes, y que la lista con los individuos que forman la población esté ordenada de forma que las personas que ocupan las posiciones 34, 84, 134, 184,..., 4984 tengan Diabetes Gestacional

Muestreo Estratificado

Representa bien a la población en el caso en que esta se divida en grupos

- 1 Tenemos un conocimiento previo de la población; por la que esta se divide en tres grupos con diferentes tipos de diabetes
 - **Tipo 1 autoinmune** Con 50 individuos (es el 1% de una población de 5000 individuos)
 - **Tipo 2** Con 4100 individuos (es el 82%)
 - **Gestacional** Con 650 individuos (es el 13%)
 - **Otros tipos** Con 200 individuos (es el 4%)
- 2 Con uno de los dos métodos anteriores elegimos:
 - una muestra de tamaño 1 de entre los 50 individuos que forman el grupo **Tipo 1 autoinmune**
 - una muestra de tamaño 82 de entre los 4100 individuos que forman el grupo **Tipo 2**
 - una muestra de tamaño 13 de entre los 650 individuos que forman el grupo **Gestacional**
 - una muestra de tamaño 4 de entre los 200 individuos que forman el grupo **Otros tipos**
- 3 Unimos las cuatro muestras para obtener una muestra de tamaño 100

Valores Atípicos o Outliers

Dada una muestra, ocasionalmente esta puede contener algún/os dato/s que dudemos que formen parte de la misma por lo diferente que es/son del resto.

Pesos en kg de niños con 10 años:

29,43 35,18 3,44 30,81 33,66 29,58 65,30 35,49 34,53 32,69
26,15 31,12 20,38 29,07 34,21 37,39 33,34 31,07 42,03 28,59

- 3,44 Se realizó una escritura incorrecta del dato. Siendo la correcta 34,40
- 65,30 La persona que pesó al niño midió el peso en libras en vez de en kg. Son 29,48 kg
- 20,38 Se incluyó por equivocación el peso de un niño de 6 años
- 26,15 y 42,03 son datos correctos porque el 90% de niños de 10 años pesan menos de 42 kg y el 10% menos de 26 kg

¿Podrían haber tenido el niño de 20,38 kg un niño con problemas de crecimiento y el de 65,30 kg con obesidad?

Valores Atípicos o Outliers

Dada una muestra, ocasionalmente esta puede contener algún/os dato/s que dudemos que formen parte de la misma por lo diferente que es/son del resto.

Pesos en kg de niños con 10 años:

29,43 35,18 3,44 30,81 33,66 29,58 65,30 35,49 34,53 32,69
26,15 31,12 20,38 29,07 34,21 37,39 33,34 31,07 42,03 28,59

- 3,44 Se realizó una escritura incorrecta del dato. Siendo la correcta 34,40
- 65,30 La persona que pesó al niño midió el peso en libras en vez de en kg. Son 29,48 kg
- 20,38 Se incluyó por equivocación el peso de un niño de 6 años
- 26,15 y 42,03 son datos correctos porque el 90% de niños de 10 años pesan menos de 42 kg y el 10% menos de 26 kg

¿Podrían haber tenido el niño de 20,38 kg un niño con problemas de crecimiento y el de 65,30 kg con obesidad?