MUESTREO E INFERENCIA

1. Hemos hecho un estudio de las horas que duermen los ciudadanos de un pueblo. Hemos elegido aleatoriamente a 130 de ellos, los resultados son los siguientes:

Horas de sueño	4 - 5	5 - 6	6 - 7	7 - 8	9 - 10	10 - 11
Número de personas	4	6	32	50	28	10

- a) Haz una estimación del tiempo medio que duermen los ciudadanos de dicho pueblo.
b) Comenta cómo modificarías el tipo de muestreo realizado con el fin de mejorar su representatividad.

2. Disponemos en una biblioteca de los siguientes ejemplares clasificados según su temática con los que vamos a hacer un estudio estadístico:

Psicología y Filosofía	Sociología	Matemáticas	Literatura	Historia
1 741	1 161	805	14 397	2 653

- a) Si la variable estadística a estudiar fuese el número total de páginas que tienen estos libros, ¿sería preferible un muestreo estratificado por nivel de estudios a uno aleatorio simple sin estratificar?
b) ¿Y qué muestreo preferiríamos si la variable estadística a estudiar fuese la nacionalidad del autor?
c) Si por fin se decidiera elegir a 150 libros utilizando el método de muestreo estratificado por temática, ¿cuántos libros de cada tipo deberíamos incluir en la muestra?

3. Nos dicen que el peso de los huevos de gallina que distribuyen una granja avícola siguen una distribución normal con una media de 63g y desviación típica de 16g. Si elegimos aleatoriamente tres docenas de huevos, calcula la probabilidad de que pesen menos de 2 kilos.

4. La altura de los adultos de una población numerosa se distribuye normalmente con una media de 165 cm y 10 cm de desviación típica.

- a) Elegido un individuo al azar calcula la probabilidad de que mida entre 163 cm y 166 cm.
b) Se elige al azar una muestra de 64 individuos calcula la probabilidad de que la altura media de dicha muestra esté entre 163 cm y 166 cm.

5. Dicen que el año pasado la epidemia de gripe afectó de alguna forma al 75% de las personas de una ciudad. Si se hubiesen tomado, elegidos al azar, 75 ciudadanos, ¿cuál es la probabilidad de que, de ellos, les haya afectado a menos de 60?

6. En una Comunidad Autónoma, las notas de las pruebas de acceso a la Universidad se distribuyen normalmente con media 6'9 y desviación típica 0'5, mientras que en otra Comunidad Autónoma la media es 6'7 y 0'6 la desviación típica. ¿Cuál es la probabilidad de que la nota media de 100 estudiantes elegidos al azar de la primera Comunidad Autónoma supere en más de tres décimas a la nota media de 100 estudiantes elegidos al azar de la segunda Comunidad Autónoma?

① a) $\bar{x} = \frac{45 \cdot 4 + 55 \cdot 6 + 65 \cdot 32 + 75 \cdot 50 + 85 \cdot 28 + 105 \cdot 10}{130} = \frac{1005}{130} = \boxed{773 \text{ h}}$

b) Como el nº de horas de sueño cambia mucho con la edad podríamos hacer un muestreo estratificado por segmentos de edad. Necesitaríamos conocer los datos de pirámide de población del pueblo.

② a) Es fácil que los libros de historia sean más voluminosos que los de literatura o física matemáticas. Sería entonces preferible un muestreo estratificado.

b) Total libros = $1741 + 1161 + 805 + 14397 + 2653 = 20756$

Muestra de 150 libros: $\frac{1741}{20756} \cdot 150 \approx 13$ de Psicología / Filosofía

$\frac{1161}{20756} \cdot 150 \approx 8$ de Sociología

$\frac{805}{20756} \cdot 150 \approx 6$ de Matemáticas

$\frac{14397}{20756} \cdot 150 \approx 104$ de Literatura

$\frac{2653}{20756} \cdot 150 \approx 19$ de Historia

③ Variable suma de muestras se distribuye como $N(\mu_m, \sigma_m)$:

$N(36 \cdot 63; 16 \cdot \sqrt{36}) = N(2268, 96)$

$P[X < 2000] = P\left[Z < \frac{2000 - 2268}{96}\right] = P\left[Z < -\frac{268}{96}\right] = P[Z < -2.79] = 1 - 0.9974 = \boxed{0.0026}$

④ a) $X \equiv N(165; 10)$

$P[163 < X < 166] = P\left[\frac{163-165}{10} < Z < \frac{166-165}{10}\right] = P[-0.2 < Z < 0.1] = 0.5398 - (1 - 0.5793) = \boxed{0.1191}$

b) $\bar{X} \equiv N\left(165; \frac{10}{\sqrt{64}}\right) = N(165; 1.25)$

$P[163 < \bar{X} < 166] = P\left[\frac{163-165}{1.25} < Z < \frac{166-165}{1.25}\right] = P[-1.6 < Z < 0.8] = 0.7881 - (1 - 0.9452) = \boxed{0.7333}$

⑤ $p_0 = 75\% = 0.75$

$\hat{p} \equiv N\left(0.75; \sqrt{\frac{0.75 \cdot 0.25}{75}}\right) = N(0.75; 0.05)$ ← se puede hacer porque $n > 30$, $np > 5$ y $n \cdot q > 5$

$P[\hat{p} < \frac{60}{75}] = P[\hat{p} < 0.8] = P\left[Z < \frac{0.8 - 0.75}{0.05}\right] = P[Z < 1] = \boxed{0.8413}$

⑥ $\bar{X}_1 \equiv N(6.9; 0.5)$

$\bar{X}_2 \equiv N(6.7; 0.6)$ $\Rightarrow \bar{X}_1 - \bar{X}_2 \equiv N(6.9 - 6.7; \sqrt{\frac{0.5^2}{100} + \frac{0.6^2}{121}}) = N(0.2; 0.078)$

$P[\bar{X}_1 - \bar{X}_2 > 0.3] = P\left[Z > \frac{0.3 - 0.2}{0.078}\right] = P[Z > 1.28] = 1 - 0.8997 = \boxed{0.1003}$