Estadística I
Tema 5: Muestreo y distribuciones muestrales

Grado en Administración de Empresas 08/09

1. X = "duración de las bujías en Km."$\sim N(36000, 4000)$. Tomamos una m.a.s. de la duración de 16 bujías: $X_1, \ldots, X_{16}$. Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:

$$\overline{X} = \frac{1}{16}\sum_{i=1}^{16} X_i \sim N\left(36000, \frac{4000}{\sqrt{16}}\right) = N(36000, 1000)$$

   Entonces, $Z = \frac{\overline{X}-36000}{1000} \sim N(0,1)$ y

$$\begin{aligned} P(\overline{X} \leq 34500) &= P\left(\tfrac{\overline{X}-36000}{1000} \leq \tfrac{34500-36000}{1000}\right) \\ &= P(Z \leq -1.5) = P(Z > 1.5) = 0.0668. \end{aligned}$$

2. X = "peso de las bolsas de azúcar en gramos". $E[X] = \mu = 500$ y $DT[X] = \sigma = 35$. Tomamos una m.a.s. del peso de 100 bolsas de azúcar: $X_1, \ldots, X_{100}$. No conocemos la distribución de X, pero como el tamaño de muestra es grande (> 30), podemos aplicar el Teorema central del límite, que dice que

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

   cuando n es suficientemente grande.
   Así pues, en este caso tendremos que $\overline{X} \sim N(500, 35/\sqrt{100}) = N(500, 3.5)$ aproximadamente.

   a) Sea $Z = \frac{\overline{X}-500}{3.5} \sim N(0,1)$, entonces

$$\begin{aligned} P(\overline{X} < 495) &= P\left(\tfrac{\overline{X}-500}{3.5} \leq \tfrac{495-500}{3.5}\right) \\ &= P(Z \leq -1.43) = P(Z > 1.43) = 0.0764. \end{aligned}$$

   b)

$$\begin{aligned} P(\textstyle\sum_{i=1}^{100} X_i > 51000) &= P(100 \cdot \overline{X} > 51000) = P(\overline{X} > 510) \\ &= P\left(\tfrac{\overline{X}-500}{3.5} > \tfrac{510-500}{3.5}\right) = P(Z > 2.86) = 0.0021. \end{aligned}$$

3. X = "peso de los soldados en Kg."$\sim N(69, 8)$. Tomamos una m.a.s. de los pesos de 12 soldados: $X_1, \ldots, X_{12}$. Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:

$$\overline{X} = \frac{1}{12}\sum_{i=1}^{12} X_i \sim N\left(69, \frac{8}{\sqrt{12}}\right) = N(68, 2.31)$$

   a) Sea $Z = \frac{\overline{X}-69}{2.31} \sim N(0,1)$, entonces

$$\begin{aligned} P(\overline{X} > 71) &= P\left(\tfrac{\overline{X}-69}{2.31} \leq \tfrac{71-69}{2.31}\right) \\ &= P(Z > 0.87) = 0.1922. \end{aligned}$$

   b)

$$\begin{aligned} P(\textstyle\sum_{i=1}^{12} X_i < 800) &= P(12 \cdot \overline{X} < 800) = P(\overline{X} < 66.67) \\ &= P\left(\tfrac{\overline{X}-69}{2.31} < \tfrac{66.67-69}{2.31}\right) = P(Z < -1.01) = P(Z > 1.01) = 0.1562. \end{aligned}$$

c) El peso de un individuo elegido al azar, es la v.a. original X. Sea $Z = \frac{X-69}{8} \sim N(0,1)$, entonces

$$\begin{aligned} P(X > 93) &= P\left(\tfrac{X-69}{8} \leq \tfrac{93-69}{2.31}\right) \\ &= P(Z > 3) = 0.0013. \end{aligned}$$

4. $X =$ "incremento porcentual del salario"$\sim N(12.2, 3.6)$. Tomamos una m.a.s. de 9 incrementos salriales: $X_1, \ldots, X_9$. Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:

$$\overline{X} = \frac{1}{9}\sum_{i=1}^{9} X_i \sim N\left(12.2, \frac{3.6}{\sqrt{9}}\right) = N(12.2, 1.2)$$

Entonces, $Z = \frac{\overline{X}-12.2}{1.2} \sim N(0,1)$ y

$$\begin{aligned} P(\overline{X} \leq 10) &= P\left(\tfrac{\overline{X}-12.2}{1.2} \leq \tfrac{10-12.2}{1.2}\right) \\ &= P(Z \leq -1.83) = P(Z > 1.83) = 0.0336. \end{aligned}$$

5. $X =$ "edad en años de los lectores de esa revista". $E[X] = \mu = 17.2$ y $DT[X] = \sigma = 2.3$. Tomamos una m.a.s. de la edad de 100 lectores: $X_1, \ldots, X_{100}$. No conocemos la distribución de X, pero como el tamaño de muestra es grande (> 30), podemos aplicar el Teorema central del límite, que dice que

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

cuando n es suficientemente grande.
Así pues, en este caso tendremos que $\overline{X} \sim N(17.2, 2.3/\sqrt{100}) = N(17.2, 0.23)$ aproximadamente.

a) Sea $Z = \frac{\overline{X}-17.2}{0.23} \sim N(0,1)$, entonces

$$\begin{aligned} P(16.7 \leq \overline{X} < 17.5) &= P\left(\tfrac{16.7-17.2}{0.23} \leq \tfrac{\overline{X}-17.2}{0.23} \leq \tfrac{17.5-17.2}{0.23}\right) \\ &= P(-2.17 \leq Z \leq 1.30) = 1 - P(Z > 1.30) - P(Z < -2.17) \\ &= 1 - P(Z > 1.30) - P(Z > 2.17) = 1 - 0.0968 - 0.0150 = 0.8882. \end{aligned}$$

b) Como ya hemos dicho en el apartado anterior, por el TCL, $\overline{X} \sim N(17.2, 0.23)$ aproximadamente.

6. $X \sim N(35, 6)$. Tomamos una m.a.s. de tamaño 49: $X_1, \ldots, X_{49}$.

a) Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

donde μ y σ son respectivamente la esperanza y la desviación típica de X.
Por tanto en nuestro caso:

$$\overline{X} = \frac{1}{49}\sum_{i=1}^{49} X_i \sim N\left(35, \frac{6}{\sqrt{49}}\right) = N(35, 6/7).$$

b) Sea $Z = \frac{\overline{X}-35}{6/7} \sim N(0,1)$, entonces

$$\begin{aligned} P(33 \leq \overline{X} \leq 36) &= P\left(\tfrac{33-35}{6/7} \leq \tfrac{\overline{X}-35}{6/7} \leq \tfrac{36-35}{6/7}\right) \\ &= P(-2.33 \leq Z \leq 1.17) = 1 - P(Z > 1.17) - P(Z < -2.33) \\ &= 1 - P(Z > 1.17) - P(Z > 2.33) = 1 - 0.1210 - 0.0099 = 0.8691. \end{aligned}$$

7. $X =$ "peso de los individuos de la población"$\sim N(65,7)$. Tomamos una m.a.s. del peso de 64 individuos: $X_1,\dots,X_{64}$.

   a) Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:
   $$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$
   donde μ y σ son respectivamente la esperanza y la desviación típica de X.
   Por tanto en nuestro caso:
   $$\overline{X} = \frac{1}{64}\sum_{i=1}^{64} X_i \sim N\left(65, \frac{7}{\sqrt{64}}\right) = N(35, 7/8).$$

   b) Sea $Z = \frac{\overline{X}-65}{7/8} \sim N(0,1)$, entonces
   $$\begin{aligned} P(\overline{X} > 66.5) &= P\left(\tfrac{\overline{X}-65}{7/8} > \tfrac{66.5-65}{7/8}\right) \\ &= P(Z > 1.71) = 0.0436. \end{aligned}$$

8. $X =$ "duración en horas de ese tipo de pilas". $E[X] = \mu = 50$ y $DT[X] = \sigma = 5$. Tomamos una m.a.s. de la duración de 16 pilas: $X_1,\dots,X_{16}$.

   a) No conocemos la distribución de X, y puesto que el tamaño de muestra no es lo suficientemente grande (> 30) para poder aplicar el Teorema central del límite, hemos de suponer que la distribución de X es normal. Bajo esta hipótesis tenemos que $\overline{X} \sim N(50, 5/\sqrt{16}) = N(50, 1.25)$.
   Entonces $Z = \frac{\overline{X}-50}{1.25} \sim N(0,1)$ y
   $$\begin{aligned} P(\overline{X} < 48) &= P\left(\tfrac{\overline{X}-50}{1.25} < \tfrac{48-50}{1.25}\right) \\ &= P(Z < -1.6) = P(Z > 1.6) = 0.0548. \end{aligned}$$

   b) Si ahora el tamaño de la muestra fuera 100, podemos aplicar el teorema central del límite, y **sin necesidad de asumir ninguna hipótesis sobre la distribución de** X, tenemos que la distribución **aproximada** de X es:
   $$\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i \sim N\left(50, \frac{5}{\sqrt{100}}\right) = N(50, 0.5).$$

9. $X =$ "edad de los miembros de cierta asociación"$\sim N(\mu,\sigma)$. Tomamos m.a.s. de las edades de 36 individuos: $X_1,\dots,X_{36}$.
   Se sabe que $E[\overline{X}] = 52$ y $DT[\overline{X}] = 0.5$.

   a) Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es $\overline{X} \sim N(\mu, \sigma/\sqrt{n})$. Por tanto
   $$52 = \mu \qquad \text{y} \qquad 0.5 = \frac{\sigma}{\sqrt{n}} = \frac{\sigma}{\sqrt{36}} \Rightarrow \sigma = 0.5 \cdot 6 = 3.$$

   b) La edad de un miembro elegido al azar, es la v.a. original X. Sea $Z = \frac{X-52}{3} \sim N(0,1)$, entonces
   $$\begin{aligned} P(X > 60) &= P\left(\tfrac{X-52}{3} \leq \tfrac{60-52}{3}\right) \\ &= P(Z > 2.67) = 0.0038. \end{aligned}$$

10. $X =$ "resistencia de los componentes"$\sim N(\mu, 3.6)$. Tomamos una m.a.s. de la resistencia de 4 componentes: $X_1, \ldots, X_4$.
Al tener X distribución normal, sabemos que

$$\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$$

de forma exacta (independientemente del valor de n), donde $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ es el estadístico varianza muestral.
Sea $W = \frac{3}{3.6^2} S^2 \sim \chi^2_3$, entonces

$$\begin{aligned} P(S^2 > 30) &= P\left(\tfrac{3}{3.6^2} S^2 > \tfrac{3}{3.6^2} 30\right) \\ &= P(W > 6.94) = 0.075. \end{aligned}$$

(interpolando, ya que los dos valores que aparecen en la tabla son $P(W > 6.25) = 0.1$ y $P(W > 7.81) = 0.05$).

11. La longitud, en cm, de las piezas fabricadas por una cierta máquina se distribuye según una $N(10, 0.25)$. Para muestras aleatorias simples de tamaño 25, calcula:

    $X =$ "longitud en cm de las piezas"$\sim N(10, 0.25)$. Tomamos una m.a.s. de la longitud de 25 piezas: $X_1, \ldots, X_{25}$.

    a) Por tener X distribución normal, sabemos que la distribución exacta de la media muestral es:

    $$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

    donde μ y σ son respectivamente la esperanza y la desviación típica de X.
    Por tanto en nuestro caso:

    $$\overline{X} = \frac{1}{25} \sum_{i=1}^{25} X_i \sim N\left(10, \frac{0.25}{\sqrt{25}}\right) = N(10, 0.05).$$

    Sea $Z = \frac{\overline{X}-10}{0.05} \sim N(0,1)$, entonces

    $$\begin{aligned} P(9.68 < \overline{X} < 10.1) &= P\left(\tfrac{9.68-10}{0.05} < \tfrac{\overline{X}-10}{0.05} < \tfrac{10.1-10}{0.05}\right) \\ &= P(-6.4 < Z < 2) = P(Z < 2) \\ &= 1 - P(Z > 2) = 1 - 0.0228 = 0.9772. \end{aligned}$$

    b) Al tener X distribución normal, sabemos que

    $$\frac{n-1}{\sigma^2} S^2 \sim \chi^2_{n-1}$$

    de forma exacta (independientemente del valor de n), donde $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$ es el estadístico varianza muestral.
    Sea $W = \frac{24}{0.25^2} S^2 \sim \chi^2_{24}$, entonces

    $$\begin{aligned} P(S^2 \leq 0.19) &= P\left(\tfrac{24}{0.0625} S^2 \leq \tfrac{24}{0.0625} 0.19\right) \\ &= P(W \leq 72.96) = 1. \end{aligned}$$