

Ejercicios Tema 1

Muestreo: distribuciones de \bar{X} , S^2 , S_1^2 , p

Rabadán-Pérez, F.

Ibar-Alonso, R.

Muñoz-Céspedes, E.

19-02-2026

1. Media muestral en población normal con σ conocida (Z)

En muchos contextos económicos (ventas, tiempos de espera, costes unitarios), es razonable modelizar una variable X como normal cuando la distribución agregada es aproximadamente simétrica y la desviación típica σ puede considerarse conocida por histórico o por calibración del proceso. Si tomamos una muestra aleatoria simple $X_1, \dots, X_n \stackrel{iid}{\sim} N(\mu, \sigma)$, entonces la **media muestral** cumple

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

Para calcular probabilidades sobre \bar{X} usamos la tipificación:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

En los ejercicios de esta sección, el objetivo es traducir un enunciado económico a una probabilidad del tipo $P(\bar{X} > c)$, $P(\bar{X} < c)$ o $P(a \leq \bar{X} \leq b)$, y resolverla consultando la tabla de la normal estándar **en cola derecha** $P(Z > z)$ (con z redondeado a dos decimales).

1.1. Ejercicio 1

En un centro de atención telefónica, el tiempo de llamada (en minutos) se modeliza como una variable normal $X \sim N(8, 4)$. Se extraen muestras aleatorias simples de tamaño $n = 16$. Obtener:

- La probabilidad de que la media muestral del tiempo de llamada sea mayor que 9 minutos.
- La probabilidad de que la media muestral esté comprendida entre 8 y 9 minutos (ambos inclusive).
- El valor c tal que la probabilidad de que la media muestral sea mayor que c sea 0.05.

1.1.1. Solución

Como $X_1, \dots, X_{16} \stackrel{iid}{\sim} N(\mu, \sigma)$ con $\mu = 8$ y $\sigma = 4$,

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(8, \frac{4}{\sqrt{16}}\right) = N(8, 1).$$

Definimos

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - 8}{1} \sim N(0, 1).$$

a)

$$P(\bar{x} > 9) = P\left(z > \frac{9 - 8}{1}\right) = P(z > 1.00).$$

De la tabla oficial (cola derecha), $P(z > 1.00) = 0.1587$.

$$\boxed{0.1587}$$

b)

$$P(8 \leq \bar{x} \leq 9) = P(\bar{x} \leq 9) - P(\bar{x} < 8).$$

$$P(\bar{x} \leq 9) = 1 - P(\bar{x} > 9) = 1 - 0.1587 = 0.8413, \quad P(\bar{x} < 8) = P(z < 0) = 0.5.$$

Por tanto,

$$0.8413 - 0.5 = 0.3413.$$

$$\boxed{0.3413}$$

c) Queremos c tal que $P(\bar{x} > c) = 0.05$. Tipificamos usando que $\mu = 8$ y $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 4/\sqrt{16} = 1$:

$$P(\bar{x} > c) = 0.05 \iff P\left(\frac{\bar{x} - \mu}{\sigma_{\bar{x}}} > \frac{c - \mu}{\sigma_{\bar{x}}}\right) = 0.05 \iff P\left(z > \frac{c - 8}{1}\right) = 0.05.$$

De la tabla oficial (cola derecha), $P(z > 1.64) \approx 0.0505$, luego tomamos $z_{0.05} \approx 1.64$ y:

$$\frac{c - 8}{1} = 1.64 \implies c = 8 + 1.64 \cdot 1 = 9.64.$$

$$\boxed{c \approx 9.64}$$

(Con $z = 1.65$, $P(z > 1.65) \approx 0.0495$ y $c \approx 9.65$.)

1.2. Ejercicio 2

En una empresa de logística, el coste por envío (en euros) se modeliza como una variable normal $X \sim N(120, 20)$. Se toma una muestra aleatoria simple de tamaño $n = 25$. Obtener:

- La probabilidad de que la media muestral del coste por envío esté comprendida entre 118 € y 122 € (ambos inclusive).
- La probabilidad de que la cuasivarianza muestral del coste por envío sea menor que 500,00.
- El valor c tal que la probabilidad de que la media muestral sea mayor que c sea 0,10.

1.2.1. Solución

Como $X_1, \dots, X_{25} \stackrel{iid}{\sim} N(\mu, \sigma)$ con $\mu = 120$ y $\sigma = 20$,

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(120, \frac{20}{\sqrt{25}}\right) = N(120, 4).$$

Definimos

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - 120}{4} \sim N(0, 1).$$

a)

$$P(118 \leq \bar{x} \leq 122) = P(\bar{x} \leq 122) - P(\bar{x} < 118).$$

Primero,

$$P(\bar{x} \leq 122) = P\left(z \leq \frac{122 - 120}{4}\right) = P(z \leq 0.50) = 1 - P(z > 0.50).$$

De la tabla oficial (cola derecha), $P(z > 0.50) = 0.3085$. Por tanto,

$$P(\bar{x} \leq 122) = 1 - 0.3085 = 0.6915.$$

Segundo,

$$P(\bar{x} < 118) = P\left(z < \frac{118 - 120}{4}\right) = P(z < -0.50).$$

Por simetría, $P(z < -0.50) = P(z > 0.50) = 0.3085$. Luego,

$$P(118 \leq \bar{x} \leq 122) = 0.6915 - 0.3085 = 0.3830.$$

$$\boxed{0.3830}$$

b) Para una población normal,

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Aquí $n = 25$, luego $n - 1 = 24$, y $\sigma^2 = 20^2 = 400$. Por tanto,

$$P(S_1^2 < 500) = P\left(\chi_{24}^2 < \frac{(24) \cdot 500}{400}\right) = P(\chi_{24}^2 < 30.00).$$

Como la tabla oficial de χ^2 da probabilidades de **cola derecha**, usamos el complemento:

$$P(\chi_{24}^2 < 30.00) = 1 - P(\chi_{24}^2 > 30.00).$$

De la tabla oficial (cola derecha) para $gl = 24$, $P(\chi_{24}^2 > 30.00) = 0.2000$. Luego,

$$P(S_1^2 < 500) = 1 - 0.2000 = 0.8000.$$

$$\boxed{0.8000}$$

c) Queremos c tal que $P(\bar{x} > c) = 0.10$. Tipificamos usando $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 4$:

$$P(\bar{x} > c) = 0.10 \iff P\left(\frac{\bar{x} - \mu}{\sigma_{\bar{x}}} > \frac{c - \mu}{\sigma_{\bar{x}}}\right) = 0.10 \iff P\left(z > \frac{c - 120}{4}\right) = 0.10.$$

De la tabla oficial (cola derecha), $P(z > 1.28) = 0.1003$, luego tomamos $z_{0.10} \approx 1.28$ y:

$$\frac{c - 120}{4} = 1.28 \implies c = 120 + 1.28 \cdot 4 = 125.12.$$

$$\boxed{c \approx 125.12}$$

1.3. Ejercicio 3

En una planta de fabricación, el peso neto (en gramos) de un lote de yogures se modeliza como una variable normal $X \sim N(125, 10)$. Se analizan $n = 16$ envases elegidos aleatoriamente. Obtener:

- La probabilidad de que la media muestral del peso neto se desvíe a lo sumo 2,5 g de la media poblacional.
- La probabilidad de que la media muestral del peso neto se desvíe más de 2,5 g de la media poblacional.

1.3.1. Solución

Como $X_1, \dots, X_{16} \stackrel{iid}{\sim} N(\mu, \sigma)$ con $\mu = 125$ y $\sigma = 10$,

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(125, \frac{10}{\sqrt{16}}\right) = N(125, 2.5).$$

Definimos

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - 125}{2.5} \sim N(0, 1).$$

a)

“Desviarse a lo sumo 2,5” significa $|\bar{x} - \mu| \leq 2.5$. Por tanto,

$$P(|\bar{x} - \mu| \leq 2.5) = P(122.5 \leq \bar{x} \leq 127.5).$$

Tipificando:

$$P(122.5 \leq \bar{x} \leq 127.5) = P\left(\frac{122.5 - 125}{2.5} \leq z \leq \frac{127.5 - 125}{2.5}\right) = P(-1.00 \leq z \leq 1.00).$$

Usamos simetría y la tabla de cola derecha:

$$P(-1.00 \leq z \leq 1.00) = 1 - 2P(z > 1.00).$$

De la tabla oficial, $P(z > 1.00) = 0.1587$. Luego,

$$1 - 2 \cdot 0.1587 = 1 - 0.3174 = 0.6826.$$

$$\boxed{0.6826}$$

b)

$$P(|\bar{x} - \mu| > 2.5) = 1 - P(|\bar{x} - \mu| \leq 2.5) = 1 - 0.6826 = 0.3174.$$

$$\boxed{0.3174}$$

2. Varianza y cuasivarianza en población normal con σ^2 conocida (χ^2)

En problemas económicos (control de calidad, estabilidad de procesos, dispersión de costes o tiempos), no solo interesa el nivel medio, sino también **la variabilidad**. En esta sección trabajamos con una población normal $X \sim N(\mu, \sigma)$ y con muestras aleatorias simples X_1, \dots, X_n .

Usaremos las dos medidas de dispersión definidas en la asignatura:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2, \quad S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2,$$

de modo que se cumple la identidad

$$nS^2 = (n-1)S_1^2.$$

Si la población es normal (con esperanza desconocida), entonces (lema de Fisher-Cochran):

$$\frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2, \quad \frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2.$$

En los ejercicios de esta sección, el objetivo es traducir el enunciado económico a una probabilidad sobre S^2 o S_1^2 , construir el estadístico χ^2 correspondiente, y consultar la tabla oficial de **cola derecha** $P(\chi_{\nu}^2 > x)$ con

$\nu = n - 1$ grados de libertad (usando complemento cuando sea necesario).

2.1. Ejercicio 4

En una compañía eléctrica, el consumo diario de un hogar (en kWh) se modeliza como una variable normal $X \sim N(\mu, 10)$. Se observan $n = 16$ días elegidos aleatoriamente. Calcular:

- La probabilidad de que la cuasivarianza muestral del consumo sea mayor o igual que 120,00.
- La probabilidad de que la varianza muestral del consumo sea mayor o igual que 112,50.
- El valor k tal que la probabilidad de que la cuasivarianza muestral sea mayor o igual que k sea 0,05.
- El valor k tal que la probabilidad de que la cuasidesviación típica muestral del consumo sea mayor que k sea 0,10.

2.1.1. Solución

En una población normal, se cumple (lema de Fisher-Cochran):

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2, \quad \frac{nS^2}{\sigma^2} \sim \chi_{n-1}^2, \quad nS^2 = (n-1)S_1^2.$$

Aquí $n = 16$, luego $n - 1 = 15$, y $\sigma = 10 \Rightarrow \sigma^2 = 100$. Definimos el estadístico

$$\chi^2 = \frac{(n-1)S_1^2}{\sigma^2} = \frac{15 S_1^2}{100} \sim \chi_{15}^2.$$

a)

$$P(S_1^2 \geq 120) = P\left(\chi_{15}^2 \geq \frac{15 \cdot 120}{100}\right) = P(\chi_{15}^2 \geq 18.00).$$

De la tabla oficial de χ^2 (cola derecha) con $gl = 15$,

$$P(\chi_{15}^2 > 18.00) = 0.2500.$$

$$\boxed{0.2500}$$

b)

$$P(S^2 \geq 112.50) = P\left(\chi_{15}^2 \geq \frac{n S^2}{\sigma^2}\right) = P\left(\chi_{15}^2 \geq \frac{16 \cdot 112.50}{100}\right) = P(\chi_{15}^2 \geq 18.00).$$

Por tanto, es la misma probabilidad que en (a):

$$\boxed{0.2500}$$

c) Queremos k tal que $P(S_1^2 \geq k) = 0.05$. Entonces,

$$P\left(\chi_{15}^2 \geq \frac{15k}{100}\right) = 0.05.$$

De la tabla oficial (cola derecha) para $gl = 15$, el cuantil x que cumple

$$P(\chi_{15}^2 \geq x) = 0.05$$

es $x = 25.00$. Por tanto,

$$\frac{15k}{100} = 25.00 \Rightarrow k = \frac{100}{15} \cdot 25.00 = 166.67.$$

$$\boxed{k \approx 166.67}$$

d) Queremos k tal que $P(S_1 > k) = 0.10$. Como $S_1 > k \Leftrightarrow S_1^2 > k^2$,

$$P(S_1 > k) = 0.10 \Leftrightarrow P(S_1^2 > k^2) = 0.10.$$

Usamos que

$$\chi^2 = \frac{(n-1)S_1^2}{\sigma^2} = \frac{15 S_1^2}{100} \sim \chi_{15}^2.$$

Entonces

$$P(S_1^2 > k^2) = 0.10 \Leftrightarrow P\left(\chi_{15}^2 > \frac{15k^2}{100}\right) = 0.10.$$

De la tabla oficial (cola derecha) para $gl = 15$, el valor x que cumple

$$P(\chi_{15}^2 > x) = 0.10$$

es $x = 22.30$. Por tanto,

$$\frac{15k^2}{100} = 22.30 \Rightarrow k^2 = \frac{100}{15} \cdot 22.30 = 148.67 \Rightarrow k = \sqrt{148.67} = 12.19.$$

$$\boxed{k \approx 12.19}$$

2.2. Ejercicio 5

En una empresa de mantenimiento industrial, el coste de una intervención estándar (en euros) se modeliza como una variable normal $X \sim N(50, 10)$. Se toma una muestra aleatoria simple de tamaño $n = 16$. Calcular:

- La probabilidad de que la media muestral del coste sea mayor que 52,50 €.
- La probabilidad de que la cuasivarianza muestral del coste sea mayor o igual que 120,00.
- La probabilidad de que, simultáneamente, la media muestral sea mayor que 52,50 € y la cuasivarianza muestral sea mayor o igual que 120,00.

2.2.1. Solución

Como $X_1, \dots, X_{16} \stackrel{iid}{\sim} N(\mu, \sigma)$ con $\mu = 50$ y $\sigma = 10$,

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(50, \frac{10}{\sqrt{16}}\right) = N(50, 2.5).$$

Definimos

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - 50}{2.5} \sim N(0, 1).$$

Además, para una población normal:

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2, \quad \text{con } n-1 = 15, \sigma^2 = 100.$$

a)

$$P(\bar{x} > 52.50) = P\left(z > \frac{52.50 - 50}{2.5}\right) = P(z > 1.00).$$

De la tabla oficial (cola derecha), $P(z > 1.00) = 0.1587$.

$$\boxed{0.1587}$$

b)

$$P(S_1^2 \geq 120) = P\left(\chi_{15}^2 \geq \frac{15 \cdot 120}{100}\right) = P(\chi_{15}^2 \geq 18.00).$$

De la tabla oficial de χ^2 (cola derecha) con $gl = 15$,

$$P(\chi_{15}^2 > 18.00) = 0.2500.$$

$$\boxed{0.2500}$$

c) En población normal, \bar{x} y S_1^2 son independientes. Por tanto,

$$P(\bar{x} > 52.50, S_1^2 \geq 120) = P(\bar{x} > 52.50) P(S_1^2 \geq 120) = 0.1587 \cdot 0.2500 = 0.0397.$$

$$\boxed{0.0397}$$

3. Diferencia de medias en poblaciones normales con σ_x y σ_y conocidas (z)

En muchos problemas económicos interesa comparar dos situaciones (dos productos, dos plantas, dos regiones, dos periodos) a partir de **muestras independientes**. Si modelizamos

$$X \sim N(\mu_x, \sigma_x), \quad Y \sim N(\mu_y, \sigma_y),$$

y tomamos dos muestras aleatorias simples e independientes X_1, \dots, X_n y Y_1, \dots, Y_m , entonces

$$\bar{x} - \bar{y} \sim N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right).$$

Para calcular probabilidades sobre $\bar{x} - \bar{y}$ tipificamos:

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}} \sim N(0, 1).$$

En los ejercicios de esta sección, el objetivo es traducir el enunciado económico a probabilidades del tipo “la diferencia de medias supera un umbral”, “es menor que un umbral” o “está entre dos valores”, y resolverlas consultando la tabla oficial de la normal estándar de **cola derecha** $P(z > z_0)$ (con z_0 redondeado a dos decimales), utilizando complementos y simetría cuando sea necesario.

3.1. Ejercicio 6

Se supone que el gasto por cliente (en euros) en una cadena de supermercados sigue una distribución normal. En compras de clientes con tarjeta premium, el gasto se distribuye como $X \sim N(50, 6)$. En compras de clientes sin tarjeta premium, el gasto se distribuye como $Y \sim N(30, 10)$. Se seleccionan aleatoriamente $n = 16$ clientes premium y $m = 25$ clientes no premium (muestras independientes). Calcular:

- La probabilidad de que el gasto medio muestral de los clientes premium sea mayor que 52,25 €.
- La probabilidad de que el gasto medio muestral de los clientes no premium sea menor que 31,00 €.
- La probabilidad de que la suma de los gastos medios muestrales de clientes premium y no premium sea mayor que 84,00 €.
- La probabilidad de que la diferencia entre los gastos medios muestrales (premium menos no premium) sea mayor que 24,00 €.

3.1.1. Solución

Como las muestras son independientes:

$$\bar{x} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right) = N\left(50, \frac{6}{\sqrt{16}}\right) = N(50, 1.5), \quad \bar{y} \sim N\left(\mu_y, \frac{\sigma_y}{\sqrt{m}}\right) = N\left(30, \frac{10}{\sqrt{25}}\right) = N(30, 2).$$

Definimos la tipificación z en cada caso usando la tabla oficial de cola derecha $P(z > z_0)$.

a)

$$P(\bar{x} > 52.25) = P\left(z > \frac{52.25 - 50}{1.5}\right) = P(z > 1.50).$$

De la tabla oficial, $P(z > 1.50) = 0.0668$.

$$\boxed{0.0668}$$

b)

$$P(\bar{y} < 31.00) = P\left(z < \frac{31.00 - 30}{2}\right) = P(z < 0.50) = 1 - P(z > 0.50).$$

De la tabla oficial, $P(z > 0.50) = 0.3085$. Luego,

$$1 - 0.3085 = 0.6915.$$

$$\boxed{0.6915}$$

c) Sea $W = \bar{x} + \bar{y}$. Entonces

$$W \sim N\left(\mu_x + \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right) = N\left(80, \sqrt{\frac{6^2}{16} + \frac{10^2}{25}}\right) = N(80, 2.50).$$

Por tanto,

$$P(\bar{x} + \bar{y} > 84.00) = P\left(z > \frac{84.00 - 80}{2.50}\right) = P(z > 1.60).$$

De la tabla oficial, $P(z > 1.60) = 0.0548$.

$$\boxed{0.0548}$$

d) Sea $D = \bar{x} - \bar{y}$. Entonces

$$D \sim N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right) = N\left(20, \sqrt{\frac{6^2}{16} + \frac{10^2}{25}}\right) = N(20, 2.50).$$

Luego,

$$P(\bar{x} - \bar{y} > 24.00) = P\left(z > \frac{24.00 - 20}{2.50}\right) = P(z > 1.60).$$

De la tabla oficial, $P(z > 1.60) = 0.0548$.

$$\boxed{0.0548}$$

3.2. Ejercicio 7

En una empresa de reparto, el tiempo de entrega (en minutos) se modeliza como normal. Para envíos estándar, el tiempo se distribuye como $X \sim N(45, 8)$. Para envíos urgentes, el tiempo se distribuye como $Y \sim N(40, 6)$. Se toman muestras independientes de tamaños $n = 16$ (estándar) y $m = 9$ (urgentes). Calcular:

- La probabilidad de que la diferencia entre los tiempos medios muestrales (estándar menos urgente) sea mayor que 8 minutos.
- La probabilidad de que la diferencia entre los tiempos medios muestrales (estándar menos urgente) sea

menor que 8 minutos.

- c) La probabilidad de que la diferencia entre los tiempos medios muestrales (estándar menos urgente) sea mayor que 10 minutos.

3.2.1. Solución

Como las muestras son independientes,

$$\bar{x} - \bar{y} \sim N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}\right).$$

Aquí $\mu_x - \mu_y = 45 - 40 = 5$ y

$$\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}} = \sqrt{\frac{8^2}{16} + \frac{6^2}{9}} = \sqrt{4 + 4} = \sqrt{8} = 2.83.$$

Definimos

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\sigma_x^2/n + \sigma_y^2/m}} = \frac{(\bar{x} - \bar{y}) - 5}{2.83} \sim N(0, 1).$$

a)

$$P(\bar{x} - \bar{y} > 8) = P\left(z > \frac{8 - 5}{2.83}\right) = P(z > 1.06).$$

De la tabla oficial (cola derecha), $P(z > 1.06) = 0.1446$.

0.1446

b)

$$P(\bar{x} - \bar{y} < 8) = 1 - P(\bar{x} - \bar{y} > 8) = 1 - 0.1446 = 0.8554.$$

0.8554

c)

$$P(\bar{x} - \bar{y} > 10) = P\left(z > \frac{10 - 5}{2.83}\right) = P(z > 1.77).$$

De la tabla oficial (cola derecha), $P(z > 1.77) = 0.0384$.

0.0384

4. Media muestral en población normal con σ desconocida (t)

En muchos problemas reales σ no es conocida y debe estimarse a partir de la muestra. Si $X \sim N(\mu, \sigma)$ y tomamos una muestra aleatoria simple X_1, \dots, X_n , la media muestral \bar{x} sigue siendo un estimador natural de μ , pero la tipificación ya no puede hacerse con σ .

En la asignatura se utilizan dos medidas de dispersión muestral:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2, \quad S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2, \quad nS^2 = (n-1)S_1^2.$$

A partir de S_1^2 se define la cuasidesviación típica muestral $S_1 = \sqrt{S_1^2}$. El estadístico de tipificación para la

media es

$$t = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} \sim t_{n-1},$$

donde los grados de libertad son $n - 1$. En los ejercicios de esta sección, el objetivo es traducir el enunciado económico a probabilidades sobre \bar{x} , tipificar mediante t y consultar la tabla oficial de la t de Student de **cola derecha** $P(t_\nu > t_0)$ (con $\nu = n - 1$), utilizando simetría y complementos cuando sea necesario.

4.1. Ejercicio 8

El ingreso diario (en miles de euros) de una tienda online se modeliza como $X \sim N(20, \sigma)$. Se registran $n = 9$ días elegidos aleatoriamente. Sabiendo que la cuasivarianza muestral vale $S_1^2 = 9$, calcular:

- La probabilidad de que la media muestral del ingreso diario sea menor que 21,86.
- El valor c tal que la probabilidad de que la media muestral sea mayor que c sea 0,10.

4.1.1. Solución

Como $X_1, \dots, X_9 \stackrel{iid}{\sim} N(\mu, \sigma)$ con $\mu = 20$ y σ desconocida, tipificamos con la t de Student usando S_1 .

A partir de $S_1^2 = 9$ se tiene $S_1 = \sqrt{9} = 3$ y

$$t = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} = \frac{\bar{x} - 20}{3/\sqrt{9}} = \frac{\bar{x} - 20}{1} = \bar{x} - 20 \sim t_{n-1} = t_8.$$

a)

$$P(\bar{x} < 21.86) = P(t < 1.86).$$

Por simetría, $P(t_8 < 1.86) = 1 - P(t_8 > 1.86)$. De la tabla oficial (cola derecha) para $gl = 8$, $P(t_8 > 1.86) = 0.0500$. Luego,

$$P(\bar{x} < 21.86) = 1 - 0.0500 = 0.9500.$$

$$\boxed{0.9500}$$

b) Queremos c tal que $P(\bar{x} > c) = 0.10$. Tipificamos:

$$P(\bar{x} > c) = 0.10 \iff P\left(t > \frac{c - 20}{S_1/\sqrt{n}}\right) = 0.10.$$

Aquí $S_1/\sqrt{n} = 3/\sqrt{9} = 1$, por tanto

$$P(t_8 > c - 20) = 0.10.$$

De la tabla oficial (cola derecha) para $gl = 8$, el valor que cumple $P(t_8 > t_0) = 0.10$ es $t_0 = 1.40$. Entonces

$$c - 20 = 1.40 \implies c = 21.40.$$

$$\boxed{c \approx 21.40}$$

4.2. Ejercicio 9

En una empresa de movilidad urbana, el retraso medio diario (en minutos) de una línea se modeliza como una variable normal $X \sim N(\mu, \sigma)$, con $\mu = 2$ y σ desconocida. Se selecciona una muestra aleatoria simple de $n = 10$ días.

Sabiendo que la varianza muestral vale $S^2 = 0.09$, calcular:

- a) La probabilidad de que la media muestral sea mayor o igual que 2.18 minutos.
 b) El valor c tal que la probabilidad de que la media muestral sea mayor o igual que c sea 0.10.
-

4.2.1. Solución

Primero pasamos de S^2 a S_1^2 (convención del curso):

$$nS^2 = (n-1)S_1^2 \Rightarrow S_1^2 = \frac{n}{n-1}S^2 = \frac{10}{9} \cdot 0.09 = 0.10, \quad S_1 = \sqrt{0.10} = 0.3162.$$

El error típico de \bar{x} estimado con S_1 es

$$\frac{S_1}{\sqrt{n}} = \frac{0.3162}{\sqrt{10}} = 0.10.$$

Como σ es desconocida,

$$t = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} = \frac{\bar{x} - 2}{0.10} \sim t_{n-1} = t_9.$$

a)

$$P(\bar{x} \geq 2.18) = P\left(t \geq \frac{2.18 - 2}{0.10}\right) = P(t \geq 1.80).$$

Con la tabla oficial de cola derecha para $gl = 9$, se tiene $t_{0.05,9} \approx 1.83$. Como $1.80 \approx 1.83$,

$$P(t_9 \geq 1.80) \approx 0.0500.$$

$$\boxed{0.0500 \text{ (aprox.)}}$$

b) Queremos c tal que $P(\bar{x} \geq c) = 0.10$. Tipificamos:

$$P(\bar{x} \geq c) = 0.10 \Leftrightarrow P\left(t_9 \geq \frac{c-2}{0.10}\right) = 0.10.$$

De la tabla oficial (cola derecha) para $gl = 9$, $t_{0.10,9} \approx 1.38$. Entonces

$$\frac{c-2}{0.10} = 1.38 \Rightarrow c = 2 + 1.38 \cdot 0.10 = 2.138 \approx 2.14.$$

$$\boxed{c \approx 2.14}$$

4.3. Ejercicio 10

En una empresa de comercio electrónico, el gasto medio por pedido (en euros) se modeliza como una variable normal $X \sim N(\mu, \sigma)$ con $\mu = 80$ y σ desconocida. Se toma una muestra aleatoria simple de tamaño $n = 9$.

Sabiendo que la cuasivarianza muestral vale $S_1^2 = 25$, calcular la probabilidad de que la media muestral difiera de μ a lo sumo 2 euros.

4.3.1. Solución

“Difera a lo sumo 2” significa $|\bar{x} - \mu| \leq 2$.

Como σ es desconocida,

$$t = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} \sim t_{n-1} = t_8.$$

A partir de $S_1^2 = 25$ se tiene $S_1 = 5$ y

$$\frac{S_1}{\sqrt{n}} = \frac{5}{\sqrt{9}} = \frac{5}{3} = 1.67.$$

Entonces

$$P(|\bar{x} - \mu| \leq 2) = P\left(|t| \leq \frac{2}{1.67}\right) = P(|t| \leq 1.20).$$

Usamos simetría y la tabla oficial de cola derecha:

$$P(|t| \leq 1.20) = 1 - 2P(t_8 > 1.20).$$

De la tabla oficial (cola derecha) para $gl = 8$, $P(t_8 > 1.20) = 0.1330$. Luego,

$$1 - 2 \cdot 0.1330 = 1 - 0.2660 = 0.7340.$$

0.7340

5. Diferencia de medias en normal con varianzas desconocidas pero iguales (t pooled)

En comparaciones entre dos grupos (dos turnos, dos procesos, dos perfiles de cliente), puede asumirse que ambas poblaciones son normales con **la misma varianza**, pero desconocida:

$$X \sim N(\mu_x, \sigma), \quad Y \sim N(\mu_y, \sigma),$$

con muestras aleatorias simples e independientes X_1, \dots, X_n y Y_1, \dots, Y_m .

En este caso, la varianza común σ^2 se estima combinando la información de ambos grupos mediante la **varianza agrupada**:

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}, \quad s_p = \sqrt{s_p^2}.$$

El estadístico de tipificación para la diferencia de medias es

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}.$$

En los ejercicios de esta sección se trabaja siempre con el supuesto de **igualdad de varianzas** (no se usa Welch). El objetivo es: (i) calcular s_p^2 , (ii) tipificar la diferencia de medias, y (iii) consultar la tabla oficial de la t de Student de **cola derecha** $P(t_\nu > t_0)$, usando simetría y complementos cuando sea necesario.

5.1. Ejercicio 11

En una empresa de consultoría, se comparan los **honorarios medios por proyecto** (en miles de euros) entre dos equipos. Se asume que los honorarios siguen distribuciones normales con **varianzas desconocidas pero iguales**.

- Equipo A: $n = 10$, $\bar{x} = 12.40$, $S_x^2 = 1.44$.
- Equipo B: $m = 12$, $\bar{y} = 11.50$, $S_y^2 = 1.00$.

Calcular:

- a) La varianza agrupada s_p^2 .
- b) El valor del estadístico t asociado al umbral $(\mu_x - \mu_y) > 0.20$.
- c) La probabilidad de que la diferencia de medias muestrales sea mayor que 0.20, es decir, $P(\bar{x} - \bar{y} > 0.20)$.

5.1.1. Solución

Se asume $X \sim N(\mu_x, \sigma)$ y $Y \sim N(\mu_y, \sigma)$ con σ común desconocida. En la asignatura:

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2, \quad S_y^2 = \frac{1}{m} \sum_{j=1}^m (Y_j - \bar{y})^2.$$

Como $nS_x^2 = \sum (X_i - \bar{x})^2$ y $mS_y^2 = \sum (Y_j - \bar{y})^2$, la varianza agrupada es

$$s_p^2 = \frac{nS_x^2 + mS_y^2}{n + m - 2}.$$

a) Con $n = 10$, $m = 12$:

$$s_p^2 = \frac{10 \cdot 1.44 + 12 \cdot 1.00}{10 + 12 - 2} = \frac{14.40 + 12.00}{20} = \frac{26.40}{20} = 1.32.$$

$$\boxed{s_p^2 = 1.32}$$

Luego $s_p = \sqrt{1.32} = 1.15$.

b) Bajo normalidad e igualdad de varianzas:

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}.$$

Aquí $df = n + m - 2 = 20$ y $\bar{x} - \bar{y} = 12.40 - 11.50 = 0.90$. Para el umbral $(\mu_x - \mu_y) = 0.20$,

$$t^* = \frac{0.90 - 0.20}{1.15 \sqrt{\frac{1}{10} + \frac{1}{12}}}.$$

Calculamos el denominador:

$$\sqrt{\frac{1}{10} + \frac{1}{12}} = \sqrt{0.10 + 0.0833} = \sqrt{0.1833} = 0.43, \quad 1.15 \cdot 0.43 = 0.49.$$

Por tanto,

$$t^* = \frac{0.70}{0.49} = 1.43.$$

$$\boxed{t^* = 1.43}$$

c)

$$P(\bar{x} - \bar{y} > 0.20) = P(t > t^*), \quad t \sim t_{20}.$$

De la tabla oficial (cola derecha) para $df = 20$, $P(t_{20} > 1.43) = 0.0830$.

$$\boxed{0.0830}$$

5.2. Ejercicio 12

En una empresa de call center, se compara la duración media de llamadas (en minutos) entre dos turnos. Se asume normalidad y varianzas desconocidas pero iguales.

- Turno A: $n = 10$, $\bar{x} = 8.20$, $S_{1x}^2 = 1.44$.
- Turno B: $m = 12$, $\bar{y} = 8.80$, $S_{1y}^2 = 1.00$.

Calcular:

- La probabilidad de que la duración media del Turno A sea menor que la del Turno B, es decir, $P(\bar{x} - \bar{y} < 0)$.
- La probabilidad de que la diferencia entre las duraciones medias muestrales se desvíe a lo sumo 1 minuto de la diferencia poblacional, es decir,

$$P(|(\bar{x} - \bar{y}) - (\mu_x - \mu_y)| \leq 1).$$

5.2.1. Solución

En la asignatura, las cuasivarianzas muestrales son

$$S_{1x}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2, \quad S_{1y}^2 = \frac{1}{m-1} \sum_{j=1}^m (Y_j - \bar{y})^2.$$

La varianza agrupada (pooled) usando cuasivarianzas es

$$s_p^2 = \frac{(n-1)S_{1x}^2 + (m-1)S_{1y}^2}{n+m-2} = \frac{9 \cdot 1.44 + 11 \cdot 1.00}{20} = \frac{23.96}{20} = 1.20, \quad s_p = \sqrt{1.20} = 1.10.$$

Además,

$$df = n + m - 2 = 20, \quad \sqrt{\frac{1}{n} + \frac{1}{m}} = \sqrt{\frac{1}{10} + \frac{1}{12}} = \sqrt{0.1833} = 0.43,$$

$$s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = 1.10 \cdot 0.43 = 0.47.$$

El estadístico es

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{20}.$$

- Aquí $\bar{x} - \bar{y} = 8.20 - 8.80 = -0.60$. Bajo $\mu_x - \mu_y = 0$,

$$t^* = \frac{-0.60}{0.47} = -1.28.$$

Entonces, por simetría,

$$P(\bar{x} - \bar{y} < 0) = P(t < -1.28) = P(t > 1.28).$$

De la tabla oficial (cola derecha) para $df = 20$, $P(t_{20} > 1.28) = 0.1070$.

$$\boxed{0.1070}$$

b)

$$P(|(\bar{x} - \bar{y}) - (\mu_x - \mu_y)| \leq 1) = P\left(|t| \leq \frac{1}{0.47}\right) = P(|t| \leq 2.13).$$

Usamos simetría y cola derecha:

$$P(|t| \leq 2.13) = 1 - 2P(t_{20} > 2.13).$$

De la tabla oficial (cola derecha) para $df = 20$, $P(t_{20} > 2.13) = 0.0250$. Luego,

$$1 - 2 \cdot 0.0250 = 0.9500.$$

6. Proporciones muestrales en grandes muestras (z aprox.)

En muchos contextos económicos la variable de interés es dicotómica (contrata/no contrata, paga/no paga, entrega a tiempo/no a tiempo). En estos casos se trabaja con la proporción muestral \hat{p} , que es la media de una variable Bernoulli.

Una muestra

Si $X_i \in \{0, 1\}$ con $P(X_i = 1) = \pi$ y se toma una muestra aleatoria simple de tamaño n , entonces

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i \approx N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right).$$

La tipificación (aprox.) es

$$z = \frac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}} \approx N(0, 1).$$

Dos muestras

Con dos muestras independientes, tamaños n y m , y probabilidades π_x, π_y :

$$\hat{p}_x - \hat{p}_y \approx N\left(\pi_x - \pi_y, \sqrt{\frac{\pi_x(1-\pi_x)}{n} + \frac{\pi_y(1-\pi_y)}{m}}\right),$$

y la tipificación (aprox.) es

$$z = \frac{(\hat{p}_x - \hat{p}_y) - (\pi_x - \pi_y)}{\sqrt{\pi_x(1-\pi_x)/n + \pi_y(1-\pi_y)/m}} \approx N(0, 1).$$

Condición de uso (imprescindible)

Antes de usar la aproximación normal se debe verificar que hay suficientes éxitos y fracasos esperados. En la asignatura se exige el criterio más restrictivo:

- Una muestra: $n\pi \geq 10$ y $n(1-\pi) \geq 10$.
- Dos muestras: $n\pi_x \geq 10$, $n(1-\pi_x) \geq 10$, $m\pi_y \geq 10$ y $m(1-\pi_y) \geq 10$.

En los ejercicios de esta sección, el objetivo es: (i) identificar π o π_x, π_y a partir del enunciado, (ii) comprobar la condición, (iii) tipificar con z y (iv) resolver con la tabla oficial de la normal estándar de cola derecha $P(z > z_0)$, usando complementos y simetría cuando sea necesario.

6.1. Ejercicio 13

En una plataforma de comercio electrónico, se sabe por experiencia que la proporción de pedidos que llegan con retraso es $\pi = 0.40$. Se selecciona una muestra aleatoria simple de $n = 225$ pedidos y se define \hat{p} como la proporción muestral de pedidos con retraso.

Calcular la probabilidad de que en la muestra la proporción de pedidos con retraso sea mayor que 0.41 (equivalentemente, que el número de pedidos con retraso sea mayor que 92).

6.1.1. Solución

Condición de uso (criterio de la asignatura):

$$n\pi = 225 \cdot 0.40 = 90 \geq 10, \quad n(1-\pi) = 225 \cdot 0.60 = 135 \geq 10.$$

Como $X \sim \text{Bin}(n, \pi)$ y $\hat{p} = X/n$,

$$X > 92 \iff \hat{p} > \frac{92}{225} = 0.41 \text{ (redondeado a dos decimales).}$$

Por aproximación normal,

$$\hat{p} \approx N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right) = N\left(0.40, \sqrt{\frac{0.40 \cdot 0.60}{225}}\right).$$

Calculamos la desviación típica:

$$\sqrt{\frac{0.24}{225}} = \sqrt{0.0010667} = 0.03.$$

Tipificamos:

$$z = \frac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}} \approx N(0, 1).$$

Entonces

$$P(\hat{p} > 0.41) = P\left(z > \frac{0.41 - 0.40}{0.03}\right) = P(z > 0.33).$$

De la tabla oficial (cola derecha), $P(z > 0.33) = 0.3707$.

0.3707

6.2. Ejercicio 14

En una entidad financiera, la proporción de clientes que contratan una tarjeta premium es $\pi = 0.30$. Se selecciona una muestra aleatoria simple de $n = 400$ clientes y se define \hat{p} como la proporción muestral de clientes que contratan la tarjeta premium.

Calcular la probabilidad de que la proporción muestral esté comprendida entre 0.26 y 0.34 (ambos inclusive).

6.2.1. Solución

Condición de uso (criterio de la asignatura):

$$n\pi = 400 \cdot 0.30 = 120 \geq 10, \quad n(1-\pi) = 400 \cdot 0.70 = 280 \geq 10.$$

Por aproximación normal,

$$\hat{p} \approx N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right) = N\left(0.30, \sqrt{\frac{0.30 \cdot 0.70}{400}}\right).$$

Desviación típica:

$$\sqrt{\frac{0.21}{400}} = \sqrt{0.000525} = 0.02.$$

Tipificamos:

$$z = \frac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}} \approx N(0, 1).$$

Entonces,

$$P(0.26 \leq \hat{p} \leq 0.34) = P\left(\frac{0.26 - 0.30}{0.02} \leq z \leq \frac{0.34 - 0.30}{0.02}\right) = P(-2.00 \leq z \leq 2.00).$$

Usando tabla oficial de cola derecha:

$$P(-2.00 \leq z \leq 2.00) = 1 - P(z > 2.00) - P(z < -2.00) = 1 - 2P(z > 2.00).$$

De la tabla oficial, $P(z > 2.00) = 0.0228$. Luego,

$$1 - 2 \cdot 0.0228 = 1 - 0.0456 = 0.9544.$$

$$\boxed{0.9544}$$

6.3. Ejercicio 15

Una empresa de suscripción digital compara dos versiones de su página de registro. Se sabe que la proporción de usuarios que completan el registro es $\pi_A = 0.40$ con la versión A y $\pi_B = 0.30$ con la versión B. Se toman dos muestras aleatorias simples e independientes:

- Versión A: $n = 400$ usuarios.
- Versión B: $m = 400$ usuarios.

Sea \hat{p}_A y \hat{p}_B la proporción muestral de registros completados en cada versión. Calcular la probabilidad de que la diferencia muestral sea mayor que 0.05, es decir,

$$P(\hat{p}_A - \hat{p}_B > 0.05).$$

6.3.1. Solución

Condición de uso (criterio de la asignatura):

$$\begin{aligned} n\pi_A &= 400 \cdot 0.40 = 160 \geq 10, & n(1 - \pi_A) &= 400 \cdot 0.60 = 240 \geq 10, \\ m\pi_B &= 400 \cdot 0.30 = 120 \geq 10, & m(1 - \pi_B) &= 400 \cdot 0.70 = 280 \geq 10. \end{aligned}$$

Por aproximación normal y muestras independientes,

$$\hat{p}_A - \hat{p}_B \approx N\left(\pi_A - \pi_B, \sqrt{\frac{\pi_A(1 - \pi_A)}{n} + \frac{\pi_B(1 - \pi_B)}{m}}\right).$$

Aquí $\pi_A - \pi_B = 0.10$ y

$$\sqrt{\frac{0.40 \cdot 0.60}{400} + \frac{0.30 \cdot 0.70}{400}} = \sqrt{\frac{0.24}{400} + \frac{0.21}{400}} = \sqrt{\frac{0.45}{400}} = \sqrt{0.001125} = 0.03.$$

Tipificamos:

$$z = \frac{(\hat{p}_A - \hat{p}_B) - (\pi_A - \pi_B)}{\sqrt{\pi_A(1 - \pi_A)/n + \pi_B(1 - \pi_B)/m}} \approx N(0, 1).$$

Entonces,

$$P(\hat{p}_A - \hat{p}_B > 0.05) = P\left(z > \frac{0.05 - 0.10}{0.03}\right) = P(z > -1.67).$$

Por simetría y usando tabla de cola derecha,

$$P(z > -1.67) = 1 - P(z \leq -1.67) = 1 - P(z > 1.67).$$

De la tabla oficial, $P(z > 1.67) = 0.0475$. Luego,

$$1 - 0.0475 = 0.9525.$$

$$\boxed{0.9525}$$

6.4. Ejercicio 16

En una compañía aseguradora, la proporción de pólizas que presentan al menos un parte en el primer año es $\pi = 0.20$. Se selecciona una muestra aleatoria simple de $n = 400$ pólizas y se define \hat{p} como la proporción muestral de pólizas con parte.

Calcular la probabilidad de que en la muestra la proporción de pólizas con parte sea mayor que 0.23 (equivalentemente, que el número de pólizas con parte sea mayor que 92).

6.4.1. Solución

Condición de uso (criterio de la asignatura):

$$n\pi = 400 \cdot 0.20 = 80 \geq 10, \quad n(1 - \pi) = 400 \cdot 0.80 = 320 \geq 10.$$

Como $X \sim B(n, \pi)$ y $\hat{p} = X/n$,

$$X > 92 \iff \hat{p} > \frac{92}{400} = 0.23.$$

Por aproximación normal,

$$\hat{p} \approx N\left(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}\right) = N\left(0.20, \sqrt{\frac{0.20 \cdot 0.80}{400}}\right).$$

Desviación típica:

$$\sqrt{\frac{0.16}{400}} = \sqrt{0.0004} = 0.02.$$

Tipificamos:

$$z = \frac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}} \approx N(0, 1).$$

Entonces

$$P(\hat{p} > 0.23) = P\left(z > \frac{0.23 - 0.20}{0.02}\right) = P(z > 1.50).$$

De la tabla oficial (cola derecha), $P(z > 1.50) = 0.0668$.

$$\boxed{0.0668}$$

7. Ejercicios integradores

En esta sección se proponen **4 ejercicios largos** (diseñados para resolverse “a boli” en ~15 minutos cada uno). El objetivo no es solo calcular, sino **identificar correctamente la casuística** a partir del enunciado económico: variable de interés, parámetro conocido/desconocido, tamaño muestral, tipo de estimador (\bar{x} , S^2 , S_1^2 , \hat{p} , $\hat{p}_A - \hat{p}_B$) y distribución/tipificación apropiada (z , χ^2 , t , t pooled).

Criterios de presentación y corrección (convención del curso): - Redondeo: **estadístico** a 2 decimales; **probabilidad** a 4 decimales. - Tablas oficiales: se usa siempre la **cola derecha**; cuando sea necesario, se aplican **simetría** y **complemento**. - En proporciones, es obligatorio comprobar la condición más restrictiva:

$$n\pi \geq 10, \quad n(1 - \pi) \geq 10 \quad (\text{y análogo en dos muestras}).$$

Cada ejercicio incluye varios apartados, de modo que el alumno practique tanto el **planteamiento** como el **cálculo numérico** sin depender de la narrativa.

7.1. Ejercicio L1 (integrador): control de calidad en envasado

Una planta de bebidas controla el **contenido neto** (en ml) de sus botellas. Por experiencia, el contenido puede modelizarse como aproximadamente normal: $X \sim N(\mu, \sigma)$.

En una inspección se toma una muestra aleatoria simple de tamaño $n = 16$. La especificación técnica del proceso indica $\sigma = 4$ ml (valor histórico de referencia). En el informe interno, el técnico calcula además la **cuasivarianza muestral** y reporta $S_1^2 = 16.00$.

- 1) **Riesgo de variabilidad:** calcular la probabilidad de que la cuasivarianza sea mayor que 26.67, es decir, $P(S_1^2 > 26.67)$.
- 2) **Riesgo de sobrellenado medio:** suponiendo $\mu = 500$ ml, calcular la probabilidad de que la media muestral supere 501.75, es decir, $P(\bar{x} > 501.75 \mid \mu = 500)$, usando el S_1^2 reportado.
- 3) **Interpretación:** ¿cambia el resultado del apartado 1) si además se informa que $\bar{x} = 501.00$ ml? (Responder con un cálculo/probabilidad.)

7.1.1. Por qué se modeliza así (explicación breve)

- El contenido neto es una variable continua y, en procesos industriales estables, la normalidad es una aproximación razonable para variaciones alrededor de un objetivo (ruido agregado de muchas fuentes).
- Para una muestra normal:
 - La variabilidad muestral se canaliza mediante S_1^2 y su distribución se obtiene con χ^2 cuando σ^2 es conocida.
 - Si σ es desconocida (o se sustituye por una dispersión muestral), la tipificación de \bar{x} conduce a una t de Student.

7.1.2. Solución

7.1.2.1. 1) Riesgo de variabilidad (distribución χ^2) Como $X_1, \dots, X_{16} \stackrel{iid}{\sim} N(\mu, \sigma)$ y $\sigma = 4$ es conocida, para la cuasivarianza S_1^2 se cumple:

$$\frac{(n-1)S_1^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Aquí $n-1 = 15$ y $\sigma^2 = 16$. Entonces

$$P(S_1^2 > 26.67) = P\left(\chi_{15}^2 > \frac{15 \cdot 26.67}{16}\right).$$

Calculamos el umbral (redondeo del estadístico a 2 decimales):

$$\frac{15 \cdot 26.67}{16} = \frac{400.05}{16} = 25.00.$$

Por tanto,

$$P(S_1^2 > 26.67) = P(\chi_{15}^2 > 25.00).$$

De la tabla oficial (cola derecha), para $gl = 15$, $P(\chi_{15}^2 > 25.00) = 0.0500$.

0.0500

7.1.2.2. 2) Riesgo de sobrellenado medio (distribución t) Como σ es desconocida en el pivote de la media (se sustituye por la dispersión muestral), se usa:

$$t = \frac{\bar{x} - \mu}{S_1/\sqrt{n}} \sim t_{n-1}.$$

Del informe $S_1^2 = 16.00 \Rightarrow S_1 = 4.00$ y $\sqrt{n} = \sqrt{16} = 4$. Por tanto,

$$\frac{S_1}{\sqrt{n}} = \frac{4.00}{4} = 1.00.$$

Con $\mu = 500$ y el umbral 501.75:

$$t^* = \frac{501.75 - 500}{1.00} = 1.75.$$

Luego,

$$P(\bar{x} > 501.75 \mid \mu = 500) = P(t_{15} > 1.75).$$

De la tabla oficial (cola derecha), para $gl = 15$, $P(t_{15} > 1.75) \approx 0.0500$.

0.0500

7.1.2.3. 3) Interpretación (independencia Fisher–Cochran) En una muestra normal, \bar{x} y S_1^2 son independientes. Por tanto,

$$P(S_1^2 > 26.67 \mid \bar{x} = 501.00) = P(S_1^2 > 26.67) = 0.0500.$$

0.0500

7.2. Ejercicio L2 (integrador): comparación de costes entre dos equipos

Dos equipos de una empresa de mantenimiento realizan el mismo tipo de intervención siguiendo el mismo protocolo y con las mismas herramientas. Se registra el **coste de material por intervención** (en euros), que puede modelizarse como aproximadamente normal.

Se toman dos muestras aleatorias simples e independientes y se obtiene el siguiente resumen:

- Equipo A: $n = 10$, $\bar{x} = 23.13$, $S_x^2 = 1.44$.
- Equipo B: $m = 12$, $\bar{y} = 22.20$, $S_y^2 = 1.00$.

En un informe interno se trabaja con una **única medida de dispersión** para ambos equipos y, en auditoría, se pide comprobar que el cálculo es consistente si se usa la definición alternativa de dispersión S_1^2 .

Calcular:

- a) La medida de dispersión común a ambos equipos y verificar el mismo resultado al recalcularla usando S_{1x}^2 y S_{1y}^2 .
- b) Suponiendo que la diferencia media real de costes fuese $\mu_x - \mu_y = 0.30$, calcular la probabilidad de observar una diferencia muestral **al menos tan grande como la observada**.
- c) Con el mismo supuesto de diferencia media real, calcular la probabilidad de que el error de estimación de la diferencia media muestral supere 1.04 euros:

$$P(|(\bar{x} - \bar{y}) - (\mu_x - \mu_y)| > 1.04).$$

7.2.1. Solución

7.2.1.1. a) **Dispersión común (dos formas equivalentes)** En la asignatura:

$$S_x^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{x})^2, \quad S_{1x}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2,$$

de modo que

$$nS_x^2 = (n-1)S_{1x}^2.$$

Análogamente, $mS_y^2 = (m-1)S_{1y}^2$.

(a1) Usando S_x^2, S_y^2

$$s_p^2 = \frac{nS_x^2 + mS_y^2}{n+m-2} = \frac{10 \cdot 1.44 + 12 \cdot 1.00}{20} = \frac{26.40}{20} = 1.32.$$

$$\boxed{s_p^2 = 1.32} \quad \Rightarrow \quad s_p = \sqrt{1.32} = 1.15.$$

(a2) Recalculando con S_{1x}^2, S_{1y}^2

Primero:

$$S_{1x}^2 = \frac{n}{n-1} S_x^2 = \frac{10}{9} \cdot 1.44 = 1.60, \quad S_{1y}^2 = \frac{m}{m-1} S_y^2 = \frac{12}{11} \cdot 1.00 = 1.09.$$

Después:

$$s_p^2 = \frac{(n-1)S_{1x}^2 + (m-1)S_{1y}^2}{n+m-2} = \frac{9 \cdot 1.60 + 11 \cdot 1.09}{20} = \frac{26.40}{20} = 1.32.$$

Coincide con (a1).

7.2.1.2. b) **Probabilidad unilateral (tabla t, cola derecha)** La diferencia observada es:

$$\bar{x} - \bar{y} = 23.13 - 22.20 = 0.93.$$

Se usa

$$t = \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n+m-2}.$$

Aquí $df = 10 + 12 - 2 = 20$ y $s_p = 1.15$. Calculamos:

$$\sqrt{\frac{1}{10} + \frac{1}{12}} = \sqrt{0.1833} = 0.43, \quad s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = 1.15 \cdot 0.43 = 0.49.$$

Con $\mu_x - \mu_y = 0.30$,

$$t^* = \frac{0.93 - 0.30}{0.49} = \frac{0.63}{0.49} = 1.28.$$

Entonces

$$P(\bar{x} - \bar{y} \geq 0.93 \mid \mu_x - \mu_y = 0.30) = P(t_{20} > 1.28).$$

De la tabla oficial (cola derecha), $P(t_{20} > 1.28) = 0.1070$.

$$\boxed{0.1070}$$

7.2.1.3. c) Probabilidad bilateral

$$P(|(\bar{x} - \bar{y}) - (\mu_x - \mu_y)| > 1.04) = P\left(|t| > \frac{1.04}{0.49}\right).$$

Como $1.04/0.49 = 2.13$,

$$P(|t| > 2.13) = 2P(t_{20} > 2.13).$$

De la tabla oficial (cola derecha), $P(t_{20} > 2.13) = 0.0250$. Luego,

$$2 \cdot 0.0250 = 0.0500.$$

0.0500

7.3. Ejercicio L3 (integrador): dos ciudades, gasto medio y morosidad

Una entidad financiera analiza dos mercados locales (Ciudad A y Ciudad B) para decidir dónde intensificar una campaña comercial. Se estudian dos indicadores:

- 1) **Gasto medio mensual** con tarjeta (en euros) de los clientes activos. Del histórico de la entidad se dispone de referencias estables de dispersión:

$\sigma_A = 30$ euros en Ciudad A y $\sigma_B = 40$ euros en Ciudad B.

En una auditoría se toman dos muestras aleatorias simples e independientes: - Ciudad A: $n = 100$, $\bar{x} = 305$. - Ciudad B: $m = 100$, $\bar{y} = 295$.

- 2) **Morosidad**: en cada ciudad se observa si el cliente ha incurrido en impago al menos una vez en los últimos 12 meses (sí/no). Del histórico se considera que las tasas son $\pi_A = 0.20$ y $\pi_B = 0.10$. En las mismas muestras anteriores ($n = m = 100$) se calcula \hat{p}_A y \hat{p}_B .

Calcular:

- a) La probabilidad de que la diferencia de gastos medios muestrales sea al menos 5 euros, es decir,

$$P(\bar{x} - \bar{y} > 5 \mid \mu_A - \mu_B = 0).$$

- b) Comprobar si se cumplen las condiciones de uso para la aproximación normal en morosidad (criterio de la asignatura) y, si se cumplen, calcular:

$$P(\hat{p}_A - \hat{p}_B > 0.20).$$

7.3.1. Solución

Por qué se modeliza así (justificación breve):

El gasto mensual es una magnitud continua y, en muestras moderadas/grandes, la media muestral se comporta aproximadamente normal. Además, el enunciado proporciona referencias históricas de dispersión por ciudad, lo que fija el error típico de $\bar{x} - \bar{y}$.

La morosidad es dicotómica (impaga/no impaga), por lo que se trabaja con proporciones muestrales y su aproximación normal en grandes muestras, verificando antes los éxitos y fracasos esperados.

7.3.1.1. a) Diferencia de medias muestrales (dos ciudades) Con muestras independientes,

$$\bar{x} - \bar{y} \sim N\left(\mu_A - \mu_B, \sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_B^2}{m}}\right).$$

Bajo $\mu_A - \mu_B = 0$:

$$\sigma_{\bar{x}-\bar{y}} = \sqrt{\frac{30^2}{100} + \frac{40^2}{100}} = \sqrt{\frac{900}{100} + \frac{1600}{100}} = \sqrt{9 + 16} = \sqrt{25} = 5.00.$$

Entonces:

$$P(\bar{x} - \bar{y} > 5) = P\left(z > \frac{5 - 0}{5.00}\right) = P(z > 1.00).$$

De la tabla oficial (cola derecha), $P(z > 1.00) = 0.1587$.

0.1587

7.3.1.2. b) Diferencia de proporciones muestrales (morosidad) Condición de uso (criterio de la asignatura: 10 éxitos y 10 fracasos esperados):

Para Ciudad A ($\pi_A = 0.20$, $n = 100$):

$$n\pi_A = 100 \cdot 0.20 = 20 \geq 10, \quad n(1 - \pi_A) = 100 \cdot 0.80 = 80 \geq 10.$$

Para Ciudad B ($\pi_B = 0.10$, $m = 100$):

$$m\pi_B = 100 \cdot 0.10 = 10 \geq 10, \quad m(1 - \pi_B) = 100 \cdot 0.90 = 90 \geq 10.$$

Se cumple la condición.

Por aproximación normal e independencia:

$$\hat{p}_A - \hat{p}_B \approx N\left(\pi_A - \pi_B, \sqrt{\frac{\pi_A(1 - \pi_A)}{n} + \frac{\pi_B(1 - \pi_B)}{m}}\right).$$

Aquí $\pi_A - \pi_B = 0.20 - 0.10 = 0.10$ y

$$\sqrt{\frac{0.20 \cdot 0.80}{100} + \frac{0.10 \cdot 0.90}{100}} = \sqrt{\frac{0.16}{100} + \frac{0.09}{100}} = \sqrt{0.0016 + 0.0009} = \sqrt{0.0025} = 0.05.$$

Tipificamos:

$$z = \frac{(\hat{p}_A - \hat{p}_B) - (\pi_A - \pi_B)}{\sqrt{\pi_A(1 - \pi_A)/n + \pi_B(1 - \pi_B)/m}} \approx N(0, 1).$$

Entonces:

$$P(\hat{p}_A - \hat{p}_B > 0.20) = P\left(z > \frac{0.20 - 0.10}{0.05}\right) = P(z > 2.00).$$

De la tabla oficial (cola derecha), $P(z > 2.00) = 0.0228$.

0.0228