ejercicios2IE.pdf

(Definitivo) Relación 2 resuelta

3º Inferencia Estadística

Grado en Matemáticas

Facultad de Ciencias
Universidad de Granada

Relación de ejercicios 2: Distribuciones en el muestreo de poblaciones normales. [1]

Ejercicio 1.

Se toma una muestra aleatoria simple de tamaño 5 de una variable aleatoria con distribución $\mathcal{N}(2'5, 36)$. Calcular:

a) Probabilidad de que la cuasivarianza muestral esté comprendida entre 1.863 y 2.674.
b) Probabilidad de que la media muestral esté comprendida entre 1.3 y 3.5, supuesto que la cuasivarianza muestral está entre 30 y 40.

SOLUCIÓN

$X \rightsquigarrow \mathcal{N}(2'5, 36), (X_1, \dots, X_n)$ m.a.s X, $n = 5$.

a) ¿$P(1{,}863 < S^2 < 2{,}674)$?

Sabemos que $\frac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2(n-1) \Rightarrow \frac{4S^2}{36} = \frac{S^2}{9} \rightsquigarrow \chi^2(4)$.

$$P(1{,}863 < S^2 < 2{,}674) = P\left(\frac{1{,}863}{9} < \frac{S^2}{9} < \frac{2{,}674}{9}\right) = P\left(0{,}207 < \frac{S^2}{9} < 0{,}2971\right)$$
$$= P\left(\frac{S^2}{9} > 0{,}207\right) - P\left(\frac{S^2}{9} > 0{,}2971\right) = 0{,}995 - 0{,}99 = 0{,}005$$

b) Ahora nos están preguntando $P\left(1{,}3 < \bar{X} < 3{,}5 \;/\; 30 < S^2 < 40\right)$. Por el lema de Fisher, sabemos que $\bar{X}$ y S^2 son independientes, luego la ocurrencia del suceso al que estamos condicionando no supondrá ninguna variación en la probabilidad del suceso $1{,}3 < \bar{X} < 3{,}5$. Así, se tiene

$$P\left(1{,}3 < \bar{X} < 3{,}5 \;/\; 30 < S^2 < 40\right) = P(1{,}3 < \bar{X} < 3{,}5) = P\left(\frac{1{,}3 - 2{,}5}{\sqrt{7{,}2}} < \frac{\bar{X} - 2{,}5}{\sqrt{7{,}2}} < \frac{3{,}5 - 2{,}5}{\sqrt{7{,}2}}\right)$$
$$= P(-0{,}44721 < Z < 0{,}37268) = P(Z < 0{,}37268) - P(Z < -0{,}44721)$$
$$= P(Z < 0{,}37628) - P(Z \geq 0{,}44721) = P(Z < 0{,}37628) - (1 - P(Z \leq 0{,}44721))$$
$$= P(Z < 0{,}37268) + P(Z \leq 0{,}44721) - 1 = 0{,}64531 + 0{,}67263 - 1 = 0{,}31794$$

donde hemos usado que $\bar{X} \rightsquigarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right) = \mathcal{N}(2'5, 7'2)$.

[1] Si encuentras algún error en el documento, escribe un correo electrónico a *melferizq@gmail.com*.

Ejercicio 2.

La longitud craneal en una determinada población humana es una variable aleatoria que sigue una distribución normal con media 185.6mm y desviación típica 12.78mm. ¿Cuál es la probabilidad de que una muestra aleatoria simple de tamaño 20 de esa población tenga media mayor que 190mm?

SOLUCIÓN

X = longitud craneal (en mm), $X \rightsquigarrow \mathcal{N}(185'6, (12'78)^2)$.

Tomamos una muestra aleatoria simple de tamaño $n = 20 \Rightarrow (X_1, \ldots, X_{20})$ m.a.s. de X.

$$P(\bar{X} > 190) = P\left(\frac{\bar{X} - 185{,}6}{12{,}78/\sqrt{20}} > \frac{190 - 185{,}6}{12{,}78/\sqrt{20}}\right) = P(Z > 1{,}5397) = 1 - P(Z \leq 1{,}5397) = 0{,}062$$

Ejercicio 3.

¿De qué tamaño mínimo habría que seleccionar una muestra de una variable con distribución normal $\mathcal{N}(\mu, 4)$ para poder afirmar, con probabilidad mayor que 0.9, que la media muestral diferirá de la poblacional menos de 0.1?

SOLUCIÓN

$X \rightsquigarrow (\mu, 4)$, ¿$n / P(|\bar{X} - \mu| < 0{,}1) > 0{,}9$? $\Rightarrow$ ¿$n / P(-0{,}1 < \bar{X} - \mu < 0{,}1) > 0{,}9$?

Sea $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \dfrac{\bar{X} - \mu}{2/\sqrt{n}} \rightsquigarrow \mathcal{N}(0, 1)$.

$$P\left(\frac{|\bar{X} - \mu|}{2/\sqrt{n}} < \frac{0{,}1}{2/\sqrt{n}}\right) = P(-0{,}05\sqrt{n} < Z < 0{,}05\sqrt{n}) = 2P(Z < 0{,}05\sqrt{n}) - 1 > 0{,}9$$

$$P(Z < 0{,}05\sqrt{n}) > 0{,}95 \Rightarrow 0{,}05\sqrt{n} = 1{,}64486 \Rightarrow n = 1082{,}22 \Rightarrow n \geq 1083$$

Ejercicio 4.

Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable con distribución normal. Calcular la probabilidad de que la cuasivarianza muestral sea menor que un 50 % de la varianza poblacional para $n = 16$ y para $n = 1000$.

SOLUCIÓN

$\boxed{16}$ Sabemos que $\dfrac{15S^2}{\sigma^2} = Y \rightsquigarrow \chi^2(15)$. Por tanto,

$$P(S^2 < 0{,}5\sigma^2) = P\left(Y < \frac{0{,}5\sigma^2 \cdot 15}{\sigma^2}\right) = P(Y < 7{,}5) = 1 - P(Y \geq 7{,}5) = \cdots = 0{,}05938$$

$\boxed{1000}$ Sabemos que $\dfrac{999S^2}{\sigma^2} = Y \rightsquigarrow \chi^2(999) \rightsquigarrow \mathcal{N}(999, 1998)$. Por tanto,

$$\begin{aligned} P(S^2 < 0{,}5\sigma^2) &= P(Y < 499{,}5) = P\left(Z < \frac{499{,}5 - 999}{\sqrt{1998}}\right) \\ &= P(Z < -11{,}1747) = 1 - P(Z < 11{,}7474) \approx 0 \end{aligned}$$

Ejercicio 5.

Sean S_1^2 y S_2^2 las cuasivarianzas muestrales de dos muestras independientes de tamaños $n_1 = 5$ y $n_2 = 4$ de dos poblaciones normales con la misma varianza. Calcular la probabilidad de que $S_1^2/S2^2$ sea menor que 5.34 o mayor que 9.12.

$\boxed{\text{SOLUCIÓN}}$

Sea $X_1 \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1^2), X_2 \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2^2)$, con $n_1 = 5, n_2 = 4$. Sabemos que, supuestas $\sigma_1^2 = \sigma_2^2$, se tiene que $S_1^2/S_2^2 \rightsquigarrow F(4, 3)$. Por tanto,

$$P\left(\frac{S_1^2}{S_2^2} < 5{,}34\right) + P\left(\frac{S_1^2}{S_2^2} > 9{,}12\right) = P\left(\frac{S_1^2}{S_2^2} < 5{,}34\right) + 1 - P\left(\frac{S_1^2}{S_2^2} < 9{,}12\right) = 0{,}9 + 1 - 0{,}95 = 0{,}95$$

Ejercicio 6.

Se consideran dos poblaciones de bombillas cuyas longitudes de vida siguen una ley normal con la misma media y desviaciones típicas 425 y 375 horas, respectivamente. Con objeto de realizar un estudio comparativo de ambas poblaciones, se considera una muestra aleatoria simple de 10 bombillas en la primera población y una de tamaño 6 en la segunda. ¿Cuál es la probabilidad de que la diferencia entre las medias muestrales del primer y segundo grupo sea menor que la observada en dos realizaciones muestrales que dieron 1325 horas y 1215 horas, respectivamente?

$\boxed{\text{SOLUCIÓN}}$

Si X e Y son variables que miden el tiempo de vida de dos poblaciones de bombillas (en horas), entonces:

$$X \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1^2), Y \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2^2) \Rightarrow \bar{X}_{10} \rightsquigarrow \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{10}\right), \bar{Y}_6 \rightsquigarrow \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{6}\right)$$

Nos dicen que $\mu_1 = \mu_2 = \mu$, por tanto, se tiene

$$\bar{X}_{10} - \bar{Y}_6 \rightsquigarrow \mathcal{N}\left(0, \frac{\sigma_1^2}{10} + \frac{\sigma_2^2}{6}\right) = \mathcal{N}(0, 203'71^2)$$

Así, entonces:

$$P(\bar{X}_{10} - \bar{Y}_6 < 110) = P\left(\frac{\bar{X}_{10} - \bar{Y}_6}{203{,}71} < \frac{110}{203{,}71}\right) = P(Z < 0{,}5399) \approx 0{,}7054$$

Ejercicio 7.

Sean $X_1, \dots, X_n, X_{n+1}$ variables aleatorias independientes e idénticamente distribuidas según una $\mathcal{N}(\mu, \sigma^2)$, y sean $\bar{X}, S^2$ la media y la cuasivarianza muestral de $X_1, \dots, X_n$. Calcular la distribución de

$$\frac{X_{n+1} - \bar{X}}{S}\sqrt{\frac{n}{n+1}}$$

SOLUCIÓN

Es evidente que $X_i \rightsquigarrow \mathcal{N}(\mu, \sigma^2), \forall i \in \{1, \dots, n+1\}$. Además, ya sabemos que $\bar{X} \rightsquigarrow \mathcal{N}(\mu, \sigma^2/n)$. Así, entonces se tendrá que

$$X_{n+1} - \bar{X} \rightsquigarrow \mathcal{N}\left(\mu - \mu, \sigma^2 + \frac{\sigma^2}{n}\right) = \mathcal{N}\left(0, \sigma^2\left(1 + \frac{1}{n}\right)\right)$$

Además sé que $\dfrac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$ y, por el lema de Fisher, es independiente de X_{n+1} y de $\bar{X}$. Demostremos que es independiente de $X_{n+1} - \bar{X}$:

$$\begin{aligned} M_{(X_{n+1}, \bar{X}, S^2)}(t, u, v) &= M_{X_{n+1}}(t) M_{(\bar{X}, S^2)}(u, v) \\ &= M_{X_{n+1}}(t) M_{\bar{X}}(u) M_{S^2}(v) = M_{(X_{n+1}, \bar{X})}(t, u) M_{S^2}(v) \end{aligned}$$

Tipifico la variable $X_{n+1} - \bar{X}$ para obtener una $\mathcal{N}(0,1)$ y así poder obtener una t de Student:

$$\frac{X_{n+1} - \bar{X}}{\sqrt{\sigma^2(1 + 1/n)}} \rightsquigarrow \mathcal{N}(0,1) \Rightarrow \frac{\dfrac{X_{n+1} - \bar{X}}{\sqrt{\sigma^2(1+1/n)}}}{\sqrt{\dfrac{S^2(n-1)/\sigma^2}{n-1}}} = \frac{X_{n+1} - \bar{X}}{S}\sqrt{\frac{n}{n+1}} \rightsquigarrow t(n-1)$$

Ejercicio 8.

Sean $(X_1, \dots, X_n), (Y_1, \dots, Y_m)$ muestras aleatorias simples independientes de poblaciones $\mathcal{N}(\mu_1, \sigma^2)$ y $\mathcal{N}(\mu_2, \sigma^2)$, respectivamente. Sean $\alpha, \beta \in \mathbb{R}$ y $\bar{X}, \bar{Y}, S_1^2, S_2^2$ las medias y las cuasivarianzas de las dos muestras. Calcular la distribución de

$$\frac{\alpha(\bar{X} - \mu_1) + \beta(\bar{Y} - \mu_2)}{\sqrt{\dfrac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}\sqrt{\dfrac{\alpha^2}{n} + \dfrac{\beta^2}{m}}}$$

SOLUCIÓN

$$X \rightsquigarrow \mathcal{N}(\mu_1, \sigma^2) \text{ e } Y \rightsquigarrow \mathcal{N}(\mu_2, \sigma^2) \to \bar{X} \rightsquigarrow \mathcal{N}\left(\mu_1, \frac{\sigma^2}{n}\right) \text{ y } \bar{Y} \rightsquigarrow \mathcal{N}\left(\mu_2, \frac{\sigma^2}{m}\right)$$

Por tanto, se tiene entonces que

$$\alpha\bar{X} \rightsquigarrow \mathcal{N}\left(\alpha\mu_1, \alpha^2 \cdot \frac{\sigma^2}{n}\right) \text{ y } \beta\bar{Y} \rightsquigarrow \mathcal{N}\left(\beta\mu_2, \beta^2 \cdot \frac{\sigma^2}{m}\right)$$

Así que

$$\alpha\bar{X} + \beta\bar{Y} \rightsquigarrow \mathcal{N}\left(\alpha\mu_1 + \beta\mu_2, \alpha^2 \cdot \frac{\sigma^2}{n} + \beta^2 \cdot \frac{\sigma^2}{m}\right) \Rightarrow \frac{\alpha\bar{X} + \beta\bar{Y} - (\alpha\mu_1 + \beta\mu_2)}{\sqrt{\sigma^2\left(\frac{\alpha^2}{n} + \frac{\beta^2}{m}\right)}} \rightsquigarrow \mathcal{N}(0,1) \quad (1)$$

Sabemos, por otro lado, que $\frac{(n-1)S_1^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$, $\frac{(m-1)S_2^2}{\sigma^2} \rightsquigarrow \chi^2(m-1)$, y además son variables aleatorias independientes. Así que, por la reproductividad de la χ^2 se tiene entonces que

$$\frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma^2} \rightsquigarrow \chi^2(n+m-2) \quad (2)$$

El lema extendido de Fisher nos asegura la independencia entre S_1^2, S_2^2 y $\bar{X}, \bar{Y}$. Tenemos por tanto que (1) y (2) son independientes, por lo que

$$\frac{\dfrac{\alpha\bar{X} - \beta\bar{Y} - (\alpha\mu_1 + \beta\mu_2)}{\sqrt{\sigma^2\left(\frac{\alpha^2}{n} + \frac{\beta^2}{m}\right)}}}{\sqrt{\dfrac{1}{\sigma^2}\dfrac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}} = \frac{\alpha(\bar{X} - \mu_1) + \beta(\bar{Y} - \mu_2)}{\sqrt{\dfrac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}\sqrt{\dfrac{\alpha^2}{n} + \dfrac{\beta^2}{m}}} \rightsquigarrow t(n+m-2)$$