

SOLUCIÓN DEL EXAMEN

16 de Diciembre de 2014

Ejercicio 1

1. Dado que el índice de Presión Sistólica en Sangre PSS tiene distribución $N(125, 144)$, entonces la variable aleatoria $Z := \frac{PSS-125}{12}$ tiene distribución $N(0, 1)$. Usando esta estandarización podemos calcular las probabilidades de cada categoría:

$$\begin{aligned} P(\text{Normal}) &= P(PSS \leq 120) = P\left(Z \leq \frac{120 - 125}{12}\right) = P(Z \leq -0.416) = \\ &= 1 - P(Z \leq 0.416) = 1 - \Phi(0.42) = 1 - 0.6628 = 0.337. \end{aligned}$$

$$\begin{aligned} P(\text{Pre-hipertenso}) &= P(120 \leq PSS \leq 140) = P\left(\frac{120 - 125}{12} \leq Z \leq \frac{140 - 125}{12}\right) = \\ &= P(-0.416 \leq Z \leq 1.25) = \Phi(1.25) - \Phi(-0.42) = 0.894 - 0.337 = 0.557 \end{aligned}$$

$$\begin{aligned} P(\text{Presión Alta}) &= P(PSS \geq 140) = P\left(Z \geq \frac{140 - 125}{12}\right) = \\ &= P(Z \geq 1.25) = 1 - \Phi(1.25) = 1 - 0.894 = 0.106. \end{aligned}$$

2. Una forma es recordar que el intervalo que contiene el 68% de la probabilidad es simplemente el intervalo $I = [\mu - \sigma, \mu + \sigma]$, donde $\mu = E(X) = 125$ y $\sigma = \sqrt{\text{Var}(X)} = 12$. Por lo tanto $I = [113, 137]$.

Otra forma es buscar un intervalo de la forma $I = [\mu - a, \mu + a]$ tal que

$$P(PSS \in I) = P\left(Z \in \left[\frac{a}{\sigma}, \frac{a}{\sigma}\right]\right) = 0.68$$

Usando las propiedades de función de distribución de la normal estandar y observando la tabla correspondiente obtenemos que $P(Z \in [-1, 1]) \approx 0.68$. Se obtiene entonces que $\frac{a}{12} = 1$, es decir $a = 12 = \sigma$.

3. (a) Utilizando la Fórmula de Probabilidad Total obtenemos:

$$\begin{aligned} P(I) &= P(N)P(I|N) + P(PH)P(I|PH) + P(PA)P(I|PA) \\ &= 0.337 \times 0.15 + 0.557 \times 0.25 + 0.55 \times 0.106 = 0.248. \end{aligned}$$

- (b) Utilizando la Fórmula de Bayes obtenemos:

$$P(N|I) = \frac{P(N) \times P(I|N)}{P(I)} = \frac{0.337 \times 0.15}{0.248} = 0.2038$$

4. Realizamos un test de hipótesis paramétrico a una cola para datos Normales para decidir si la presión media se redujo o no. Planteamos:

$$\begin{cases} H_0 : \mu \geq \mu_0, \\ H_1 : \mu < \mu_0 \end{cases}$$

donde $\mu_0 = 125$, $\alpha = 0.05$, $\sigma = 12$ y $\hat{X} = 122.77$.

El estadístico de prueba es:

$$Z = \sqrt{n} \left(\frac{\hat{X} - \mu_0}{\sigma} \right) = \sqrt{10} \left(\frac{122.77 - 125}{12} \right) \approx -0.5876$$

Por otro lado $z_\alpha = 1,645$ y la región de rechazo esta dada por la condición $Z \leq -z_\alpha$, luego no rechazamos H_0 .

Se concluye que podemos suponer con un 95% de confianza que la presión media no se redujo.

Ejercicio 2

Sea $f(x) = 2 - 2x$ la función definida en $[0, a]$.

1. Dado que f es positiva en $[0, a]$, para que sea función de densidad es necesario que $\int_0^a f(x)dx = 1$.

$$\int_0^a f(x)dx = 2 \int_0^a (1 - x)dx = 1 \rightarrow 2a - a^2 = 1 \Rightarrow a = 1.$$

Es posible observar también que la función f es positiva hasta $x = 1$ y en ese caso forma un triángulo de área 1, indicando que es una densidad.

2. $F_X(x) = P(X \leq x) = \int_0^x f(y)dy$. Por lo tanto tenemos tres casos:

- si $x < 0$, $f(x) = 0$ y por lo tanto $F_X(x) = 0$.
- si $0 \leq x \leq 1$ entonces $F_X(x) = 2 \int_0^x 1 - ydy = 2x - x^2$.
- si $x > 1$ entonces $F_X(x) = 2 \int_0^1 1 - ydy + 2 \int_1^x 1 - ydy = 2 \int_0^1 1 - ydy = 1$.

En resumen:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 2x - x^2 & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

3. $P(X \geq \frac{1}{2}) = 1 - P(X < \frac{1}{2}) = 1 - F_X(\frac{1}{2}) = 1 - \left(2 \cdot \frac{1}{2} - \left(\frac{1}{2} \right)^2 \right) = \frac{1}{4}$.

También es posible calcular esta probabilidad como el área de un triángulo de base $1/2$ y altura 1.

$$4. E(X) = 2 \int_0^1 x(1-x)dx = 2 \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = \frac{1}{3}.$$

$$5. E(Y) = E(2X + 1) = 2E(X) + 1 = 2\frac{1}{3} + 1 = \frac{5}{3}.$$

Para calcular la varianza de Y es necesario calcular la varianza de X y para ello necesitamos $E(X^2)$.

$$E(X^2) = 2 \int_0^1 x^2(1-x)dx = 2 \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 = \frac{1}{6} \text{ y la varianza de } X \text{ resulta entonces:}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

$$\text{Finalmente } \text{Var}(Y) = \text{Var}(2X + 1) = 4\text{Var}(X) = 4\frac{1}{18} = \frac{2}{9}.$$

Ejercicio 3

1. Test de Rachas: El estadístico que se obtiene es $R = 3 < \frac{2n-1}{3} = \frac{11}{3}$.

Si se plantea el test a una cola:

$$\begin{cases} H_0 : \text{La muestra es aleatoria} \\ H_1 : \text{La muestra presenta pocas rachas} \end{cases}$$

Entonces la tabla da directamente el p-valor $\alpha^* = 0.4139 > 0.1$. Por lo tanto, no es posible rechazar la hipótesis nula.

La conclusión es la misma si se plantea un test a dos colas, es decir:

$$\begin{cases} H_0 : \text{La muestra es aleatoria} \\ H_1 : \text{La muestra no es aleatoria} \end{cases}$$

Test de Spearman: el vector posición (o vector de rangos) es $R_i = (6, 4, 1, 2, 5, 3)$. Por lo tanto el estadístico es:

$$R_S = 1 - 6 \frac{\sum_{i=1}^6 (R_i - i)^2}{N \times (N^2 - 1)} = 1 - 6 \frac{\sum_{i=1}^6 (R_i - i)^2}{6 \times 35} = 1 - \frac{1}{35} (25 + 4 + 4 + 4 + 0 + 9) = 1 - \frac{46}{35} = -0.3143$$

Si se plantea el test a una cola:

$$\begin{cases} H_0 : \text{La muestra es aleatoria} \\ H_1 : \text{Hay dependencia negativa} \end{cases}$$

Entonces la tabla da directamente el p-valor $\alpha^* = 0.282 > 0.1$. Por lo tanto, no es posible rechazar la hipótesis nula.

La conclusión es la misma si se plantea un test a dos colas, es decir:

$$\begin{cases} H_0 : \text{La muestra es aleatoria} \\ H_1 : \text{La muestra no es aleatoria} \end{cases}$$

Visto los resultados del test de Rachas y el test de Correlación de Rangos de Spearman, podemos suponer que la muestra es aleatoria (i.i.d).

2. Vamos a realizar el test de ajuste de Kolmogorov-Smirnov a la distribución propuesta: $\mathcal{U}[60.5, 64.5]$.

$$\begin{cases} H_0 : \text{La muestra se ajusta a } F_0 = \mathcal{U}[60.5, 64.5] \\ H_1 : \text{No } H_0 \end{cases}$$

X_i^*	$F_o(X_i^*) = \frac{X_i^* - 60.5}{4}$	$\frac{i}{n}$	$\frac{i-1}{n}$	$ F_o(X_i^*) - \frac{i}{n} $	$ F_o(X_i^*) - \frac{i-1}{n} $
61.26	0.19	0.167	0.000	0.024	0.19
61.40	0.225	0.333	0.167	0.108	0.059
63.05	0.637	0.500	0.334	0.137	0.3045
63.91	0.852	0.666	0.500	0.186	0.3525
64.11	0.902	0.833	0.667	0.069	0.2365
64.24	0.935	1.000	0.834	0.065	0.102

El estadístico Kolomogorov Smirnov resulta entonces $KS = 0.3525$. La tabla indica que el p-valor es mayor que 0.2. Por lo tanto no es posible rechazar la hipótesis nula. Es decir podemos asumir que la muestra se ajusta a la distribución propuesta.

3. Para testear si ambas muestras tienen la misma distribución, vamos a realizar un test de Kolmogorov-Smirnov de comparación de dos muestras.

$$\begin{cases} H_0 : Z \sim X \\ H_1 : \text{No } H_0 \end{cases}$$

Z_i^*	Muestra	$F_m^*(Z^*)$	$F_n^*(Z^*)$	$ F_m^*(Z^*) - F_n^*(Z^*) $
60.52	Y	0.000	0.167	0.167
61.15	Y	0.000	0.334	0.334
61.26	X	0.167	0.334	0.167
61.27	Y	0.167	0.500	0.334
61.40	X	0.334	0.500	0.167
62.40	Y	0.334	0.667	0.334
63.05	X	0.500	0.667	0.167
63.34	Y	0.500	0.834	0.334
63.56	Y	0.500	1.000	0.500
63.91	X	0.667	1.000	0.334
64.11	X	0.834	1.000	0.167
64.24	X	1.000	1.000	0.000

El estadístico de Kolmogorov Smirnov es $mnD_{mn} = 6 \times 6 \times \frac{1}{2} = 18$. La tabla indica que el p-valor es mayor que 0.1. Por lo tanto no es posible rechazar la hipótesis nula. Es decir, podemos asumir que la distribución de la muestra tomada en el año 2012 coincide con la de la muestra tomada en el año 2011.

4. Asumiendo que el promedio mensual sigue una distribución Uniforme en el intervalo $[60.5, 64.5]$ se obtiene que

$$P(X > 63.5) = 1 - P(X \leq 63.5) = 1 - \frac{63.5 - 60.5}{4} = 1 - \frac{3}{4} = \frac{1}{4}.$$