ESTADÍSTICA DESCRIPTIVA E INTRODUCCIÓN A LA PROBABILIDAD
DOBLE GRADO EN INFORMÁTICA Y MATEMÁTICAS ~ 20 DE JUNIO DE 2018

[3 puntos] Rodea la opción correcta en cada pregunta.
Puntuación por pregunta: respuesta correcta: +0.5; incorrecta -0.2; en blanco 0.

1. Sean *A, B* y *C* tres sucesos independientes dos a dos. Entonces siempre se cumple que:
a) $\mathrm{P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)}$
b) Si *C* es independiente de $A \cup B$, entonces: $\mathrm{P((\overline{A} \cup \overline{B})/ C) = 1 - P(A) \cdot P(B)}$
c) $\mathrm{P(A/ (B \cap C)) = P(A)}$
d) $\mathrm{P(A/ B) = P(B/A)}$

2. Sea X una variable aleatoria discreta con media m y desviación típica 0. Entonces se puede afirmar que:
a) X = m
b) $P(X \neq 0) = 0$
c) $P(X = 0) = 1$
d) $P(X \neq 0) = 1$

3. Indica la afirmación falsa:
a) Una variable aleatoria que es transformada de otra no tiene por qué ser siempre del mismo tipo que ésta.
b) La distribución geométrica modela el número de fracasos que han ocurrido antes de que ocurra el primer éxito.
c) Sea X una variable aleatoria continua en todo *R*, entonces $F_x(x) = 1 - P_{-x}(x)$, $\forall x \in R^+$
d) El momento no centrado de orden 2 de una variable aleatoria nunca puede ser inferior al cuadrado del momento no centrado de orden 1 de dicha variable.

4. Indica la respuesta correcta:
a) El coeficiente de variación de una variable tipificada es nulo.
b) El coeficiente de determinación, en el caso de la regresión lineal, coincide con el coeficiente de correlación lineal.
c) Si el valor de la vivienda se ha incrementado un 10%, 3%, 2% y 9% respectivamente durante los últimos 4 años, el incremento medio anual del valor de la vivienda durante dicho periodo ha sido de un 6%.
d) Todas las anteriores son falsas.

5. Sea X una variable aleatoria discreta y considérese la nueva v.a Y=g(X), donde g es una función continua y estrictamente decreciente. Si h=g^3, entonces la función de distribución de la variable Y es:
a) $F_y(y) = 1 - F_x\big(h(y)\big)$
b) $F_y(y) = 1 - F_x\big(h(y)\big) + P(Y = y)$
c) $F_y(y) = F_x\big(h(y)\big)$
d) $F_y(y) = \big[1 - F_x\big(h(y)\big)\big] \cdot |h'(y)|$

6. Indica la afirmación correcta:

a) La distribución de Poisson es un caso particular de la binomial, cuando n es grande y p pequeño.

b) En sucesivos experimentos independientes éxito/fracaso, con P(éxito)=p constante, el número de realizaciones hasta encontrar el r-ésimo éxito sigue una distribución binomial negativa.

c) Si p=0.4 en una distribución binomial, el cálculo de $\frac{7!}{4!\cdot 3!}(0.6)^4 \cdot (0.4)^3$, nos da la probabilidad de conseguir exactamente tres éxitos en 7 ensayos.

d) La distribución hipergeométrica toma valores entre 0 y n con probabilidades no nulas.

ESTADÍSTICA DESCRIPTIVA E INTRODUCCIÓN A LA PROBABILIDAD
DOBLE GRADO EN INFORMÁTICA Y MATEMÁTICAS ~ 20 DE JUNIO DE 2018

1. [2 puntos] En una plantación se estudian diferentes parcelas, observando el número de veces al año en que se utiliza fertilizante (variable X) junto con el rendimiento medido en kg/m^2 de cada parcela (variable Y). Los datos obtenidos durante un año han sido:

X \ Y	[6,8]	]8,10]	]10,14]	]14,16]
3	7	5	1	0
6	5	10	0	0
8	0	1	15	0
10	0	0	4	2
15	0	0	0	10

a) Para las parcelas que se fertilizaron en al menos 8 ocasiones, ¿qué rendimiento es el más habitual? ¿qué rendimiento superan el 75% de las parcelas?
b) Según un modelo lineal. ¿qué rendimiento tendrá una parcela que se fertiliza una vez al mes?, ¿qué cantidad de veces se habrá fertilizado una parcela que tiene un rendimiento de $11.5 kg/m^2$?
c) ¿Qué fiabilidad tienen las predicciones realizadas en el apartado anterior?

2. [3 puntos] Sea X una variable aleatoria con función de distribución:

$$F(x) = \begin{cases} 0 & si \quad x < 1 \\ \frac{x-1}{a} & si \quad 1 \leq x < 2 \\ \frac{x+1}{6} & si \quad 2 \leq x < k \\ 1 & si \quad x \geq k \end{cases}$$

a) Determinar los valores de *k* y *a* para que F(x) sea la función de distribución de alguna variable aleatoria continua X.
b) Calcular el coeficiente de variación de X.
c) Sea Y una variable aleatoria tal que:

$$Y = \begin{cases} 0 & si \quad x < 1.5 \\ 1 & si \quad 1.5 \leq x < 2.5 \\ 2 & si \quad x \geq 2.5 \end{cases}$$

Calcular la función generatriz de momentos de Y y obtener a partir de ella su esperanza matemática.
d) Supongamos un algoritmo que genera un dígito binario (0,1) cada vez que se ejecuta, y que la probabilidad de que genere un dígito 0 es la misma que la $P(0.5<X<1.5)$ siempre.

- Si se ejecuta el algoritmo 10 veces, ¿Cuál es la probabilidad de obtener un uno 8 veces o más?
- ¿Cuál es el número esperado de ceros antes de obtener el tercer uno?

3. [2 puntos] Tres urnas U1, U2 y U3 contienen respectivamente 3 bolas blancas y 3 negras; 4 bolas blancas y 6 negras; y 4 bolas blancas y 2 negras.

a) Tomamos una bola de la primera urna y otra de la segunda. ¿Cuál es la probabilidad de obtener exactamente una bola blanca?
b) Se elige una urna al azar y se seleccionan dos bolas simultáneamente. ¿Cuál es la probabilidad de que ambas sean negras? En el caso de que ambas sean negras, ¿cuál es la probabilidad de haberlas extraído de la tercera urna?
c) Tomamos una bola de la segunda urna y la introducimos en la tercera. Tras este proceso, ¿cuál es el número esperado de bolas blancas en la tercera urna?

Exámenes EDIP

1er examen (2018)

X = nº veces fertilizante al año
Y = rendimiento en kg/m²

1.

X\Y	7	9	12	15	$n_{i.}$	$n_{i.} x_i$	$n_{i.} x_i^2$	$x_i \sum n_{ij} y_j$
3	7	5	1	0	13	39	117	318
6	5	10	0	0	15	90	540	1150
8	0	1	15	0	16	128	1024	1512
10	0	0	4	2	6	60	600	780
15	0	0	0	10	10	150	2250	2250
$n_{.j}$	12	16	20	12	60	467	4531	5610
$n_{.j} c_j$	84	144	240	180	648			
$n_{.j} c_j^2$	588	1296	2880	2700	7464			

b) ¿Rendimiento de una parcela fertilizada una vez al mes? ¿Veces que se habrá fertilizado si su rendimiento es 11.5 kg/m²?

Una vez al mes ⇒ 12 veces al año

$\boxed{\text{Recta } Y/X}$ $\quad y = \dfrac{\sigma_{xy}}{\sigma_x^2} x + \bar{y} - \dfrac{\sigma_{xy}}{\sigma_x^2} \bar{x}$

$\bar{x} = \dfrac{467}{60} = 7.78\hat{3}$ veces $\quad \bar{y} = \dfrac{648}{60} = 10.8$ kg/m²

$\sigma_{xy} = \dfrac{5610}{60} - 7.78\hat{3} \cdot 10.8 = 9.44$

$\sigma_x^2 = \dfrac{4531}{60} - 7.78\hat{3}^2 = 14.936$ veces²

$y = 0.632x + 5.8807 \xRightarrow{\text{Predicción}} y = 0.632 \cdot 12 + 5.8807 = 13.4647 \dfrac{kg}{m^2}$

$\boxed{\text{Recta } X/Y}$ $\quad x = \dfrac{\sigma_{xy}}{\sigma_y^2} y - \dfrac{\sigma_{xy}}{\sigma_y^2} \bar{y} + \bar{x}$

$\sigma_y^2 = \dfrac{7464}{60} - 10.8^2 = 7.76$

$x = 1.2165y - 5.3548 \xRightarrow{\text{Predicción}} x = 1.2165 \cdot 11.5 - 5.3548 = 8.63495$ veces al año

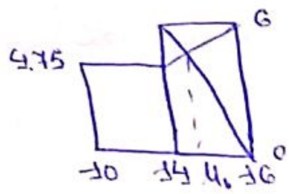
c) ¿Fiabilidad?

$r^2 = \dfrac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} = 0.76886$ ⇒ El ajuste lineal explica el 76.886% de los casos. Tiene una fiabilidad aceptable a)

$Y/x\geq 8$	c_j	a_j	$n_{3j}+n_{4j}+n_{5j}$	N_j	h_j
[6,8]	7	2	0	0	0
]8,10]	9	2	1	1	0.5
]10,14]	12	4	19	20	4.75
]14,16]	15	2	12	32	6

El rendimiento más habitual es la moda, que se hallará en el intervalo modal:]14,16]

Semejanza de triángulos
⇓

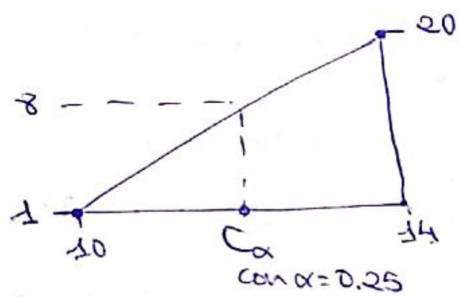


$$\frac{6-4.75}{M_0-14} = \frac{6}{16-M_0}$$

$$\Rightarrow -1.25M_0+20 = 6M_0-84 \Rightarrow M_0 = 14.3448\ \frac{kg}{m^2}$$

El rendimiento que superan el 75% de las parcelas es $C_{0.25}$. Se calcula usando semejanza de triángulos y la curva de distribución:

$0.25\cdot 32 = 8 \Rightarrow$ Se hallará en]10,14]



$$\frac{14-10}{20-1} = \frac{C_\alpha-10}{8-1}$$
⇓
$$C_{0.25} = 11.4737\ \frac{kg}{m^2}$$

2. $$F(x) = \begin{cases} 0 & \text{si } x<1 \\ \frac{x-1}{a} & \text{si } 1\leq x<2 \\ \frac{x+1}{6} & \text{si } 2\leq x<k \\ 1 & \text{si } x\geq k \end{cases}$$

a) Determinar k y a.
Para ello es necesario que sea continua:

$$\frac{k+1}{6} = 1 \Rightarrow k+1=6 \Rightarrow \boxed{k=5}$$

$$F(2) = 1/2 \Rightarrow \lim_{x\to 2^-} F(x) = \frac{1}{2} = \frac{1}{a} \Rightarrow \boxed{a=2}$$

b) Coef. variación X.

Necesitamos $E[X]$ y $\sqrt{Var(X)} = \sqrt{E(X^2) - (E(X))^2}$, y para ello necesitamos la función de densidad:

$$f(x) = F'(x) = \begin{cases} 0 & \text{si } x < 1 \\ 1/2 & \text{si } 1 \leq x < 2 \\ 1/6 & \text{si } 2 \leq x < 5 \\ 0 & \text{si } x \geq 5 \end{cases}$$

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx = \int_1^2 \frac{x}{2}\,dx + \int_2^5 \frac{x}{6}\,dx = \left[\frac{x^2}{4}\right]_1^2 + \left[\frac{x^2}{12}\right]_2^5 =$$

$$= \frac{4}{4} - \frac{1}{4} + \frac{25}{12} - \frac{1}{6} = \frac{8}{3} = 2.\hat{6}$$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x)\,dx = \int_1^2 \frac{x^2}{2}\,dx + \int_2^5 \frac{x^2}{6}\,dx = \left[\frac{x^3}{6}\right]_1^2 + \left[\frac{x^3}{18}\right]_2^5 =$$

$$= \frac{8}{6} - \frac{1}{6} + \frac{125}{18} - \frac{8}{18} = \frac{23}{3} = 7.\hat{6}$$

$$\sigma = \sqrt{7.\hat{6} - 2.\hat{6}^2} = 0.7454$$

$$C.V.(X) = \frac{\sigma}{\bar{x}} = 0.2795085$$

c) $$Y = \begin{cases} 0 & \text{si } x < 1.5 \\ 1 & \text{si } 1.5 \leq x < 2.5 \\ 2 & \text{si } x \geq 2.5 \end{cases}$$

Función generatriz de momentos de Y y obtener con ella su esperanza.

$P(Y=0) = P(X < 1.5) = 1/4$

$P(Y=1) = P(1.5 \leq X < 2.5) = P(X < 2.5) - P(X \leq 1.5) =$
$= 1/3$

$P(Y=2) = P(X \geq 2.5) = 1 - P(X \leq 2.5) = 5/12$

$$M_x(t) = \sum_{i=0}^{2} e^{ty_i} \cdot P[Y = y_i] = e^t \cdot \frac{1}{3} + e^{2t} \cdot \frac{5}{12} + \frac{1}{4}$$

$E[X] = M_x'(0)$

$$M_x'(t) = \frac{e^t}{3} + \frac{5e^{2t}}{6} \Rightarrow E[X] = \frac{7}{6} = 1.1\hat{6}$$

d) $Z = \begin{cases} 0 & \text{si } 0.5 < X < 1.5 \\ 1 & \text{en cualquier otro caso} \end{cases}$

$P(Z=0) = P(0.5 < X < 1.5) = P(X < 1.5) - P(X < 0.5) =$

$= 1/4 = 0.25$

$P(Z=1) = 1 - P(Z=0) = 3/4 = 0.75$

• Algoritmo ejecutado 10 veces. ¿Prob. de obtener un 1 ocho o más veces?

$X \rightsquigarrow B(10, 0.75)$ $X \equiv$ nº unos al usar el algoritmo 10 veces

$P(X \geq 8) = \binom{10}{8} 0.75^8 \cdot 0.25^2 + \binom{10}{9} 0.75^9 \cdot 0.25 + \binom{10}{10} 0.75^{10} \cdot 0.25^0 =$

$= 0.28157 + 0.187712 + 0.0563135 = 0.525596$

• ¿Nº esperado de ceros antes de obtener el 3er uno?

nº fracasos (obtener ceros) hasta el 3er éxito

$X \equiv$ nº fracasos (obtener ceros) hasta el 3er éxito (obtener unos)

$X \rightsquigarrow BN(3, 0.75)$

$E[X] = \frac{r(1-p)}{p} = \frac{3(1-0.75)}{0.75} = 1$ cero

→ Hagámoslo a la inversa. Si salen 1 ocho o más veces, 0 salen 2 o menos, por lo que:

$X \equiv$ nº ceros al usar el algoritmo 10 veces

$X \rightsquigarrow B(10, 0.25)$

$P(X \leq 2) = \binom{10}{0} 0.25^0 \cdot 0.75^{10} + \binom{10}{1} 0.25 \cdot 0.75^9 + \binom{10}{2} 0.25^2 \cdot 0.75^8 =$

$= 0.056314 + 0.187712 + 0.28157 = 0.52559$

3. $U_1 \equiv$ 3 blancas y 3 negras

$U_2 \equiv$ 4 blancas y 6 negras

$U_3 \equiv$ 4 blancas y 2 negras

a) Dos bolas (una de U_1 y otra de U_2). ¿Prob. de sacar exactamente una bola blanca?

$P(\text{Blanca en } U_1) + P(\text{Blanca en } U_1) =$

$= \frac{\binom{3}{1}\binom{6}{1}}{\binom{16}{2}} + \frac{\binom{3}{1}\binom{4}{1}}{\binom{16}{2}} = \frac{3 \cdot 6}{120} + \frac{3 \cdot 4}{120} = \frac{1}{4} = 0.25$

$P(1 \text{ blanca}) = P(\text{Blanca } 1 \cap \text{Negra } 2) + P(\text{Blanca } 2 \cap \text{Negra } 1) =$

$= \frac{3}{6} \cdot \frac{6}{10} + \frac{4}{10} \cdot \frac{3}{6} = \frac{1}{2} = 0.5$

b) Urna al azar y se le sacan 2 bolas. ¿Prob. ambas negras? Si son negras, ¿prob. de haberlas extraído de la 3ª urna?

Por el T^ma^ de la prob. total:

$$P(NN) = P(U_1) \cdot P(NN/U_1) + P(U_2) \cdot P(NN/U_2) + P(U_3) \cdot P(NN/U_3) =$$

$$= \frac{1}{3}\left(\frac{3}{6} \cdot \frac{2}{5} + \frac{6}{10} \cdot \frac{5}{9} + \frac{2}{6} \cdot \frac{1}{5}\right) = \frac{1}{5} = 0.2$$

$$P(U_3/NN) = \frac{P(U_3 \cap NN)}{P(NN)} = \frac{P(U_3) \cdot P(NN/U_3)}{P(NN)} = \frac{\frac{1}{3} \cdot \frac{2}{6} \cdot \frac{1}{5}}{\frac{1}{5}} =$$

$$= \frac{1}{9} = 0.\overline{1}$$

c) Sacamos una bola de la 2ª y la introducimos en la 3ª. ¿Nº esperado de bolas blancas en la 3ª urna?

$X \sim B(1, 0.4)$ (Distribución de Bernoulli)

$X \equiv$ nº bolas blancas extraído

$P(X=0) = 1 - 0.4 = 0.6 \quad P(X=1) = 0.4$

$E[X+4] = E[X] + E[4] = 0 \cdot 0.6 + 1 \cdot 0.4 + 4 = 4.4$ bolas blancas en U_3 se esperan

Porque la 3ª urna ya tiene 4 bolas blancas