Estadística. Grado en Informática. Curso 2010-11
Departamento de Estadística. Escuela Politécnica Superior
Examen Final, 1ª parte. Febrero 2011

Respuestas

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a															
b															
c															

<u>Marcar la respuesta correcta en cada una de las preguntas que siguen.</u> NO OLVIDAR trasladar las respuestas marcadas a la tabla que está situada en la parte superior (si no es así, no se corregirá la prueba). RECUERDE que las preguntas incorrectas RESTAN un tercio.

1.- La mediana como medida de centralización es aplicable a variables
- **a)** En escala nominal o superior
- **b)** En escala ordinal o superior
- **c)** En escala por intervalos o superior

2.- La varianza como medida de dispersión,
- **a)** Siempre será aplicable
- **b)** No siempre será aplicable, dependerá del tipo de variables
- **c)** Se aplicará únicamente a variables numéricas, agrupadas en intervalos

3.- El método de mínimos cuadrados para obtener los coeficientes de regresión en un modelo:
- **a)** No es aplicable si el modelo es de tipo parabólico
- **b)** Es aplicable únicamente si el modelo es de tipo lineal simple
- **c)** Es aplicable a cualquier tipo de modelo

4.- Si Cov (X,Y) ≥ 0
- **a)** $R_{xy} > 0$
- **b)** $R_{yx} \geq 0$
- **c)** $R_{xy} = 0$

5.- Tipificar unos datos, consiste en
- **a)** Determinar cuál es el máximo y cuál es el mínimo
- **b)** Centrar los datos y dividirlos por su desviación típica
- **c)** Centrar los datos y dividirlos por su varianza

6.- Sea A y B dos sucesos tales que P(A) = P(B) = 0.2,
- **a)** $P(A \cap B) = 0.04$
- **b)** A y B son incompatibles
- **c)** $P(A \cup B) \leq 0.4$

7.- El Teorema de Bayes para el cálculo de probabilidades a posteriori
- **a)** Es aplicable a cualquier tipo de sucesos
- **b)** Los E_i sucesos que intervienen deben de ser independientes entre si.
- **c)** Requiere entre otros requisitos, que los E_i sucesos que intervienen sean excluyentes.

8.- La ley de Laplace de asignación de probabilidad puede aplicarse
- **a)** Es aplicable a cualquier tipo de sucesos
- **b)** Requiere entre otros requisitos, que los E_i sucesos que intervienen sean igualmente verosímiles.
- **c)** Los E_i sucesos que intervienen deben de ser independientes entre si

9.- Si $P(B) = P\left(B/A\right)$, entonces los sucesos son:
- **a)** Nunca puede ocurrir.
- **b)** Independientes.
- **c)** Incompatibles.

10.- Una fábrica de componentes industriales tiene tres proveedores: A, B y C. Las probabilidades de que un pedido provenga de alguno de ellos es 0.3, 0.6 y 0.1 respectivamente. El proveedor A proporciona productos con calidad Inferior a la tolerada con probabilidad 0.2, el proveedor B, con probabilidad 0.15 y para el proveedor C se desconoce. En el supuesto de que la probabilidad de que un componente elegido al azar sea de calidad Aceptable es de 0.8:

- **a)** $P\left(I/C\right) = 0.6$
- **b)** $P\left(I/C\right) = 0.5$
- **c)** $P\left(I/C\right) = 0.3$

11.- El lanzamiento y viaje de la nave espacial MRO hacia Marte, depende de varios ordenadores conectados en paralelo y que funcionan independientemente, de tal manera que la MRO llega a Marte funcionando al menos un ordenador. La probabilidad de funcionamiento de cada ordenador es 0.9, y deseamos conocer cuántos ordenadores se han conectado en paralelo para que la nave MRO pueda realizar su misión en Marte con probabilidad 0.999.

- **a)** 5
- **b)** 4
- **c)** 3

12.- Cierto curso está dividido en tres grupos, una asignatura es impartida en cada grupo por un profesor diferente, el examen final es el mismo para los tres grupos, los resultados obtenidos se muestran a continuación:

$$n_1 = n_2 = n_3 = 15 \quad ; \quad \vec{\mu}_{123} = \begin{pmatrix} 69.0\widehat{6} \\ 74 \\ 71.9\widehat{3} \end{pmatrix} \quad ; \quad \sum_{123} = \begin{pmatrix} 162.06\widehat{2} & 145.2049 & 207.1428 \\ 145.2049 & 132 & 186.7358 \\ 207.1428 & 186.7358 & 270.99\widehat{5} \end{pmatrix}$$

- **a)** El grupo 1 es el más homogéneo
- **b)** El grupo 2 es el más homogéneo
- **c)** El grupo 3 es el más homogéneo

13.- Los dos grupos que presentan mayor relación lineal, según los datos anteriores, son:

- **a)** El 1 y el 3
- **b)** El 1 y el 2
- **c)** El 2 y el 3

14.- **Con los datos del ejercicio 12, se ha calculado la recta de regresión lineal $\hat{G}_1 = a + b\,G_2$, que ha resultado ser:**

- **a)** $\hat{G}_1 = -12.33 + 1.1\,G_2$
- **b)** $\hat{G}_1 = 12.116 + 0.896\,G_2$
- **c)** $\hat{G}_1 = 12.33 - 1.1\,G_2$

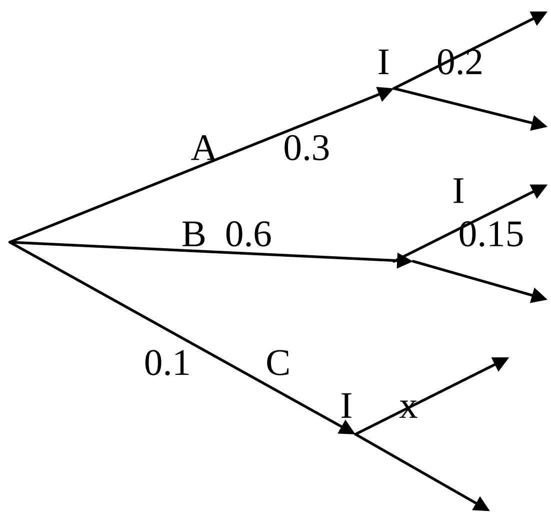
15.- La bondad del modelo anterior es:

- **a)** $R^2 = 0.9854$
- **b)** $R^2 = 0.9769$
- **c)** $R^2 = 0.9747$

Ejercicio 10.- Una fábrica de componentes industriales tiene tres proveedores: A, B y C. Las probabilidades de que un pedido provenga de alguno de ellos es 0.3, 0.6 y 0.1 respectivamente. El proveedor A proporciona productos con calidad inferior a la tolerada con probabilidad 0.2, el proveedor B, con probabilidad 0.15 y para el proveedor C se desconoce. En el supuesto de que la probabilidad de que un componente elegido al azar sea de calidad aceptable es de 0.8, verificar cuál es la hipótesis correcta.

Sean los sucesos: A ≡ Componente procede del proveedor A
B ≡ Componente procede del proveedor B
C ≡ Componente procede del proveedor C
I ≡ Calidad del componente Inferior a la tolerada

Según el enunciado: $\begin{cases} P(A)=0.3 \;;\; P(B)=0.6 \;;\; P(C)=0.1 \\ P\left(I/A\right)=0.2 \;;\; P\left(I/B\right)=0.15 \;;\; P\left(I/C\right)=X \\ P\left(\bar{I}\right)=0.8 \;\Rightarrow\; P(I)=0.2 \end{cases}$



Aplicando el teorema de la probabilidad total o de la Partición, para los tres posibles valores de P(I/C) = X, tenemos:

$$P(I)=P(A)\times P\left(I/A\right)+P(B)\times P\left(I/B\right)+P(C)\times P\left(I/C\right)=0.2$$

$$0.2=0.3\times 0.2+0.6\times 0.15+0.1\times X \Rightarrow X = P\left(I/C\right)=0.5$$

Ejercicio 11.- El lanzamiento y viaje de la nave espacial Mars Reconnaissance Orbiter (MRO) hacia Marte, depende de varios ordenadores conectados en paralelo y que funcionan independientemente, de tal manera que la MRO llega a Marte funcionando al menos un ordenador.

La probabilidad de funcionamiento de cada ordenador es 0.9, y deseamos conocer cuantos ordenadores se han conectado en paralelo para que la nave MRO pueda realizar su misión en Marte con probabilidad 0.999.

Sea el suceso $\mathbf{F_i}$ ≡ "Funciona el ordenador nº **i**",

siendo la probabilidad de dicho suceso: $P\left(F_i\right)=0.9$

y conectamos **N** ordenadores en paralelo,

Para que el sistema funcione (suceso **F**), ha de funcionar 1 o 2 o 3 o los N ordenadores, es decir:

$$\begin{aligned}P\left(F\right)=P\Big[&\left(F_1\cap\overline{F}_2\cap\overline{F}_3\cap...\cap\overline{F}_N\right)\cup\left(\overline{F}_1\cap F_2\cap\overline{F}_3\cap...\cap\overline{F}_N\right)\cup...\cup\left(\overline{F}_1\cap\overline{F}_2\cap\overline{F}_3\cap...\cap F_N\right)\cup\\ &\cup\left(F_1\cap F_2\cap\overline{F}_3\cap...\cap\overline{F}_N\right)\cup......\cup..............................\cup\left(F_1\cap F_2\cap F_3\cap...\cap F_N\right)\Big]=0.999\end{aligned}$$

Por el suceso complementario: $P\left(F\right)=1-P\left(\overline{F}\right)=1-P\left(\overline{F}_1\cap\overline{F}_2\cap\overline{F}_3\cap...\cap\overline{F}_N\right)=0.999$

$$\begin{aligned}&P\left(F\right)=1-P\left(\overline{F}\right)=1-P\left(\overline{F}_1\cap\overline{F}_2\cap\overline{F}_3\cap...\cap\overline{F}_N\right)=\\ &=1-\left(0.1\times0.1\times...\times0.1\right)=1-0.1^N=0.999\end{aligned}$$

haciendo operaciones:

$$0.1^N=1\times10^{-3}\Rightarrow N=\frac{Ln\left(1\times10^{-3}\right)}{Ln\,0.1}=\underline{3}\rfloor$$

Ejercicios 12-15.- Cierto curso está dividido en tres grupos, una asignatura es impartida en cada grupo por un profesor diferente, el examen final es el mismo para los tres grupos, siendo las notas obtenidas las siguientes (se puntúa de 0 a 100):

Grupo 1	Grupo 2	Grupo 3
47	56	43
52	59	48
52	59	50
57	61	55
63	67	61
64	69	67
69	73	72
71	76	78
72	76	80
72	80	80
78	83	83
81	83	85
81	84	89
86	90	91
91	94	97

Deseamos calcular:

1. **¿En que grupo las notas están más dispersas?**
2. **Obtener la matriz de correlación.**
3. **Modelo de regresión lineal simple $\hat{G}_1 = a + b\,G_2$. Bondad del modelo calculado. ¿Qué significa el valor obtenido?.**

DATOS AUXILIARES:

$$\mathbf{n_1 = n_2 = n_3 = 15}$$

$$\vec{\mu}_{G_1G_2G_3} = \begin{pmatrix} 69.0\widehat{6} \\ 74 \\ 71.9\widehat{3} \end{pmatrix} \quad ; \quad \Sigma_{G_1G_2G_3} = \begin{pmatrix} 162.06\widehat{2} & 145.2049 & 207.1428 \\ & 132 & 186.7358 \\ & & 270.99\widehat{5} \end{pmatrix}$$

1. **¿En que grupo las notas están más dispersas?**

Para determinar en que Grupo las notas están más dispersas, calcularemos el Coeficiente de Variación para cada Grupo $CV = \frac{S}{\overline{X}}$

Grupo 1: $CV_{G_1} = \frac{\sqrt{162.06\widehat{2}}}{69.0\widehat{6}} = 0.1843 \Rightarrow 18.43\%$

Grupo2: $CV_{G_2} = \frac{\sqrt{132}}{74} = 0.1552 \Rightarrow 15.52\%$

Grupo 3: $CV_{G_3} = \frac{\sqrt{270.99\widehat{5}}}{71.9\widehat{3}} = 0.2288 \Rightarrow 22.88\%$

Por tanto en el Grupo 3 las notas están más dispersas, y **las del Grupo 2 las más homogéneas.**

2. **Obtener la matriz de correlación**

$$\Gamma_{G_1G_2G_3} = \begin{pmatrix} 1 & R_{12} & R_{13} \\ & 1 & R_{23} \\ & & 1 \end{pmatrix} \; ; \text{ siendo } R_{xy} = \frac{S_{xy}}{S_x S_y}$$

de acuerdo con los datos auxiliares: Matriz de covarianza:

$$\sum\nolimits_{G_1G_2G_3} = \begin{pmatrix} 162.06\overset{\frown}{2} & 145.2049 & 207.1428 \\ & 132 & 186.7358 \\ & & 270.99\overset{\frown}{5} \end{pmatrix}$$

$$R_{12} = \frac{145.2049}{\sqrt{162.06\overset{\frown}{2}} \times \sqrt{132}} = 0.9927 \; ; \; R_{13} = \frac{207.1428}{\sqrt{162.06\overset{\frown}{2}} \times \sqrt{270.99\overset{\frown}{5}}} = 0.9884$$

$$R_{23} = \frac{186.7358}{\sqrt{270.99\overset{\frown}{5}} \times \sqrt{132}} = 0.9873$$

luego la matriz de correlación es:

$$\Gamma_{G_1G_2G_3} = \begin{pmatrix} 1 & 0.9927 & 0.9884 \\ 0.9927 & 1 & 0.9873 \\ 0.9884 & 0.9873 & 1 \end{pmatrix}$$

3. **Modelo de regresión lineal simple $\hat{G}_1 = a + b\,G_2$. Bondad del modelo calculado. ¿Qué significa el valor obtenido?.**

$$b = \frac{S_{12}}{S_2^2} = \frac{145.2049}{132} = 1.1 \; ; \; a = \overline{G_1} - b\,\overline{G_2} = 69.0\overset{\frown}{6} - 1.1 \times 74 = -12.3334$$

es decir: $\hat{G}_1 = -12.3334 + 1.1\,G_2$

<u>Bondad del modelo calculado:</u>

$$R^2 = (R_{12})^2 = (R_{21})^2 = 0.9927^2 = 0.9854$$

El modelo calculado explica el 98.54% de la variable Dependiente