Métodos Estadísticos Aplicados a la Ingeniería **Examen Temas 1-4**
Ingeniería Industrial (E.I.I.) **23/4/09**

Apellidos y Nombre: **Calificación:**

Cuestión 1.1.- Se ha calculado el percentil 85 sobre las estadísticas de siniestralidad en el sector de la construcción durante el último año y se ha obtenido el valor de 2,5. El significado de este dato es…
a)… el 2,5 por ciento de los trabajadores en el sector de la construcción sufren menos de 85 accidentes al año.
b)… el 15 por ciento de los trabajadores en el sector de la construcción sufren menos de 2,5 accidentes al año.
c)… el 85 por ciento de los trabajadores en el sector de la construcción sufren menos de 2,5 accidentes al año.
Cuestión 1.2.- El número de intervalos de clase para una variable estadística agrupada plantea que:
a) si hay muchas clases se produce pérdida de información.
b) si hay pocas clases hay mucha dispersión y excesivo detalle.
c) si hay clases semi-vacías en los extremos del rango debemos agrupar con menos clases.
Problema 1.- Sea la distribución referida a beneficios anuales de 38 empresas madrileñas más modestas:

Beneficio (Miles de €)	Nº empresas
230-280	5
280-330	7
330-580	14
580-630	9
630-780	3

Se pide
a) Calcular el beneficio medio de estas 38 empresas madrileñas
b) ¿Cuál es el beneficio mayor de la mitad de las empresas más modestas?
c) Determinar el beneficio más frecuente.
d) Estudiar la dispersión de esta distribución a partir del recorrido intercuartílico, desviación típica y coeficiente de variación de Pearson. Interpretar los resultados obtenidos.
e) Estudiar la forma de esta distribución. Comentar el resultado.
Cuestión 2.1.- Después de estudiar la relación existente entre la flexión y la extensión de cuello de los alumnos de la UEx, obtenemos que el valor de la covarianza es -0,57. ¿El valor de r saldrá positivo o negativo?
a) Saldrá positivo porque la relación es inversa.
b) Saldrá negativo también porque el signo de la covarianza y del coeficiente de correlación lineal de Pearson siempre coinciden.
c) No podemos saber el signo de r sabiendo el de la covarianza porque no están relacionados.
d) Todas son falsas.
e) Necesitamos conocer previamente R2
Cuestión 2.2.- La distribución conjunta de dos variables es siempre:
a) el producto de las dos distribuciones marginales.
b) el producto de la distribución marginal de una variable y la distribución condicionada de otra variable.
c) suma la unidad si las dos variables son independientes.
Problema 2.- Dada la siguiente tabla de correlaciones entre la edad (X) y las horas semanales frente al televisor (Y) [X e Y son las marcas de clase correspondientes a distribuciones dadas en intervalos]

X/Y	5	15	40	70
10	3	7	6	4
20	6	14	12	8
40	5	10	9	6
60	1	4	3	2

a) Consideradas aisladamente, ¿qué variable presenta mayor dispersión relativa?, ¿y mayor concentración?
b) Calcule la covarianza entre la edad del individuo y las horas semanales empleadas en ver TV. ¿Qué puede comentar con este resultado?
c) Calcule la distribución de las horas de TV condicionado a los individuos de 20 años de edad.
d) Calcule la distribución de la edad condicionada a aquellos que ven 5 horas de TV a la semana.
e) Calcule la media aritmética de horas de TV que ve el intervalo más joven de la población (X=10)
Cuestión 3.1.- Si dos sucesos A y B son incompatibles (excluyentes):
a) La intersección es el conjunto vacío.
b) La probabilidad de la intersección es cero.
c) La probabilidad de la unión es la suma de las probabilidades.
d) Todo lo anterior es cierto.
e) Sólo la a) y la c) son correctas.

Cuestión 3.2.- El 2% de la población padece diabetes. Si de ellos, el 30% no está diagnosticado, esta cantidad puede entenderse como una probabilidad…
a) De un suceso intersección.
b) De un suceso de unión.
c) Condicionada.
d) A posteriori.
e) De un suceso complementario.

Problema 3.- Dos exámenes tienen lugar al mismo tiempo en dos aulas contiguas. Las dos aulas, denominadas A y B, están ocupadas por un 90% y un 50% de alumnas, respectivamente. En el aula B hay cuatro veces más estudiantes, alumnos y alumnas, que en el aula A. Las dos aulas se vacían simultáneamente en el mismo pasillo, donde antes no había nadie. ¿Cuál es la probabilidad que una alumna, elegida al azar en el pasillo, haya sido del aula A?

Cuestión 4.1.- La distribución de Poisson puede considerarse como una buena aproximación de una distribución Binomial (n,p) si:
a) p es muy grande y n muy pequeño.
b) sólo son posibles dos resultados: éxito y fracaso.
c) n es grande y p es pequeño.
d) su distribución es similar a la distribución normal tipificada.
e) todas son correctas.

Cuestión 4.2.- Qué propiedad o propiedades caracterizan a una distribución normal tipificada frente a una distribución normal cualquiera:
a) El área encerrada bajo su función de densidad es igual a 1.
b) Su media es 1 y su desviación típica es 0.
c) Su rango de valores oscila entre 0 y 3.
d) Su media es 0 y su desviación típica es 1.
e) son ciertas **a)** y **c)**.

Problema 4.- Un partido político decide, con vistas a las próximas elecciones, hacer una pequeña encuesta entre sus votantes, con el fin de averiguar si la gestión que está realizando es aprobada por los electores. Una vez realizado el estudio se obtiene que el 49% de los electores está en desacuerdo con la gestión.
a) Calcular la probabilidad de que el tercer elector que se encuentre sea el primero que esté en desacuerdo con la gestión.
b) ¿Cuál es la probabilidad de que el número de electores que están de acuerdo con la gestión sean tres o más de seis encuestados?
c) ¿Cuál es la probabilidad de que el séptimo elector que se encuentre sea el tercero que está de acuerdo con la gestión?
d) Calcular el valor esperado y la varianza del número de electores que están en desacuerdo con la gestión de veinte encuestados.

NOTAS SOBRE EL EXAMEN

La duración del examen es de 2 horas aproximadamente. Se valorará positivamente la claridad y la concisión en la respuesta de las preguntas teóricas, así como, en la resolución de los problemas.
El alumno podrá elegir bien entre las cuestiones y problema asociados al tema 1 o al tema 2, siendo obligatorio contestar y resolver las cuestiones y problemas asociados a los temas 3 y 4.
El formulario (extensión máxima la cara de un folio) se ha de entregar con el examen, pudiéndose recuperar una vez calificado el examen. La presencia de ejercicios o teoría en el mismo supone un suspenso directo.
La nota mínima exigida para eliminar este examen estará entre 6,5 y 7, decisión final del profesorado.
Los problemas 1, 2 y 4 se valorarán sobre 2,5 puntos
Las cuestiones se valorarán sobre 0,5
El problema 3 vale 2 puntos

SOLUCIONES AL EXAMEN PARCIAL DE MÉTODOS ESTADÍSTICOS APLICADOS A LA INGENIERÍA

Cuestión 1.1.- La respuesta correcta es la **c),** véase teoría.
Cuestión 1.2.- La respuesta correcta es la **c)**, véase teoría.
Problema 1

Beneficio (Miles de €)	x_i	Nº empresas	N_i	F_r
230-280	255	5	5	0,1316
280-330	305	7	12	0,1842
330-580	455	14	26	0,3684
580-630	605	9	35	0,2368
630-780	705	3	38	0,0789
		N=38		**F_r=1**

a) El beneficio medio de estas 38 empresas madrileñas es de 456,32 (miles de euros)

$$\overline{B} = \frac{\sum_{i=1}^{5} x_i n_i}{N} = \frac{255*5+305*7+455*14+605*9+705*3}{38} = 456{,}3158$$

b) El beneficio mayor de la mitad de las empresas más modestas es:

$\frac{N}{2} = 19$ → Este valor se encuentra, mirando la columna de N_i, entre 12 y 26. Por tanto la mediana vale 455 (miles de euros).

c) El beneficio más frecuente, mirando la columna de F_r, es de 455 (miles de euros).

d)

- El rango intercuartílico: **IRQ**=Q3-Q1=300

$$r\frac{N}{q} = 1\frac{38}{4} = 9{,}5 \rightarrow Ni < r\frac{N}{q} < N_{i+1} \Rightarrow 5 < 9{,}5 < 12$$

Q_1=305

$$g\,r\frac{N}{q} = 3\frac{38}{4} = 28{,}5 \rightarrow Ni < r\frac{N}{q} < N_{i+1} \Rightarrow 26 < 28{,}5 < 35$$

Q_3=605

- La desviación típica es:

$$S_B = \sqrt{\frac{\sum_{i=1}^{5}\left(x_i - \overline{X}\right)^2 n_i}{N}} = \sqrt{\frac{(255-456{,}32)^2*5+(305-456{,}32)^2*7+(455-456{,}32)^2*14+(605-456{,}32)^2*9+(705-456{,}32)^2*3}{38}} = 140{,}25$$

- El coeficiente de variación de Pearson:

$$C_V = \frac{S_B}{\overline{B}} = 0{,}3073$$

Podemos deducir a partir de las diferentes magnitudes de dispersión calculadas que los datos mostrados en la tabla presentan una gran dispersión.

e)

- Coeficiente de asimetría de Fisher

$$g_1 = \frac{\frac{\sum_{i=1}^{5}\left(x_i - \overline{B}\right)^3 n_i}{N}}{S_B^3} = 0{,}1017>0 \rightarrow \text{Asimetría positiva o hacia la derecha}$$

- Coeficiente de curtosis de Fisher

$$g_2 = \frac{\frac{\sum_{i=1}^{5}\left(x_i - \overline{B}\right)^4 n_i}{N}}{S_B^4} = 1{,}8877<3 \rightarrow \text{Platicúrtica}$$

Cuestión 2.1.- La respuesta correcta es la **b)**, véase teoría.
Cuestión 2.2.- La respuesta correcta es la **b)**, véase teoría.
Problema 2.-

X/Y	5	15	40	70	$n_{i.}$
10	3	7	6	4	20
20	6	14	12	8	40
40	5	10	9	6	30
60	1	4	3	2	10
$n_{.j}$	15	35	30	20	100

a) Para la variable X.

x_i	$n_{i.}$
10	20
20	40
40	30
60	10

La media de la variable X es: **28**
La desviación de la variable X es: **15,36**
El coeficiente de variación es: **0,5486**

Para la variable Y

y_j	$n_{.j}$
5	15
15	35
40	30
70	20

La media de la variable Y es: **32**
La desviación de la variable Y es: **22,77**
El coeficiente de variación es: **0,7116**

Como el coeficiente de variación es más pequeño para la variable X que la de la variable Y, podemos decir que en ella es menor la dispersión, y por tanto, los valores de X están más concentrados entorno a la media.

b) La **covarianza** es igual a **1**, podemos deducir que existe una relación directa entre la edad de la persona y el número de horas que ve la televisión semanalmente.

c) La distribución de las horas de TV condicionado a los individuos de 20 años de edad es:

X/Y	5	15	40	70
20	6	14	12	8

d) La distribución de la edad condicionada a aquellos que ven 5 horas de TV a la semana es:

X/Y	5
10	3
20	6
40	5
60	1

e) La media aritmética de horas de TV que ve el intervalo más joven de la población (X=10) es de **32 horas semanales**.

X/Y	5	15	40	70	
10	3	7	6	4	**20**

Cuestión 3.1.- La respuesta correcta es la **d)**, véase teoría
Cuestión 3.2.- La respuesta correcta es la **c)**, véase teoría
Problema 3.- Vamos a denotar por:
M: que el alumno elegido al azar sea mujer
A: que el alumno elegido al azar sea del aula A
B: que el alumno elegido al azar sea del aula B
De acuerdo con los datos que nos ofrece el problema, así como la pregunta misma que se realiza nos lleva a pensar en la aplicación del teorema de Bayes. En nuestro caso

$$P(A/M)=\frac{P(M/A)P(A)}{P(M/A)P(A)+P(M/B)P(B)}$$

Con ayuda del enunciado identificamos los diferentes términos que aparecen en la expresión anterior.

P(M/A)=0,9
P(M/B)=0,5
P(A)=1/5
P(B)=4/5

Sustituyendo en la expresión tenemos que:

P(A/M)=0,31

Cuestión 4.1.- La respuesta correcta es la **c)**, véase teoría
Cuestión 4.2.- La respuesta correcta es la **d)**, véase teoría

Problema 4.-

Denotemos el suceso por **A**=la gestión sea aprobada por un votante, de tal forma que tenemos que:x

$$P(A) = 0{,}51$$
$$P(\overline{A}) = 0{,}49$$

a) Nos hablan que el tercer elector sea el primero en estar en desacuerdo con la gestión. Por tanto, la distribución de la variable discreta es la distribución geométrica.

Sea **X** el número del primer elector en desacuerdo con la gestión.

X~G(p=0,49)

entonces

$$P(X = x) = pq^{x-1} \Rightarrow P(X = 3) = (0{,}49)(0{,}51)^2 = 0{,}12745$$

b) Sea **Y** el número de los seis encuestados que está de acuerdo con la gestión. La distribución que sigue es una binomial

Y~B(n=6,p=0,51)

entonces

$$P(X \geq 3) = 1 - P(X < 3) = 1 - \sum_{x=0}^{2} \binom{n=6}{x} p^x q^{n-x} = 1 - \left[\binom{6}{0}(0{,}49)^6 + \binom{6}{1}(0{,}51)(0{,}49)^5 + \binom{6}{2}(0{,}51)^2(0{,}49)^4 \right] = 0{,}6748$$

c) Sea **Z** el número del tercer elector que está de acuerdo con la gestión. La distribución que sigue es una binomial negativa.

Z~BN(r=3,p=0,51)

entonces

$$P(Z = x) = \binom{x-1}{r-1} p^r q^{x-r} \Rightarrow P(Z = 7) = \binom{6}{2}(0{,}51)^3(0{,}49)^4 = 0{,}1147$$

d) Sea **W** el número de los 20 encuestados en desacuerdo con la gestión. La distribución seguida es la binomial y por tanto, la esperanza y la varianza es:

W~B(n=20,p=0,49)

entonces

$$E(W) = np = 20(0{,}49) = 9{,}8$$
$$V(W) = npq = 20(0{,}49)(0{,}51) = 4{,}998$$