

Grado en Ingeniería Informática

Estadística

PRIMER PARCIAL

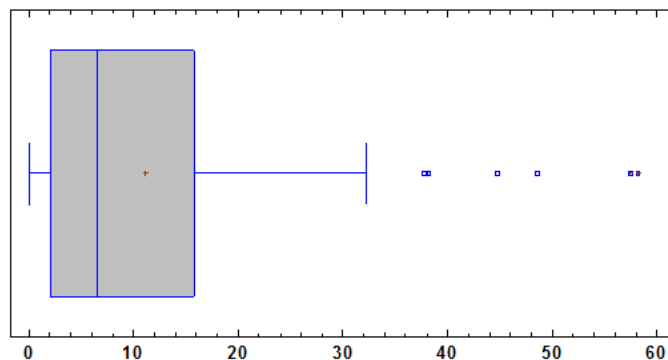
21 de marzo de 2016

Apellidos, nombre:	
Grupo:	Firma:

Instrucciones

1. Rellenar la información de cabecera del examen.
2. Responder a cada pregunta en la hoja correspondiente.
3. Justificar todas las respuestas.
4. No se permiten anotaciones personales en el formulario.
5. No se permite tener teléfonos móviles encima de la mesa. Sobre la mesa sólo se permite el DNI, calculadora, útiles de escritura, las tablas y el formulario.
6. No desgrapar las hojas.
7. Todas las preguntas puntúan lo mismo (sobre 10).
8. Se debe firmar en las hojas que hay en la mesa del profesor al entregar el examen. Esta firma es el justificante de la entrega del mismo.
9. Tiempo disponible: **2 horas**

1. Se han recogido los tiempos en meses que tardan en romperse 100 componentes electrónicos utilizados en el montaje de equipos informáticos, y los resultados los tenemos representados en este gráfico:



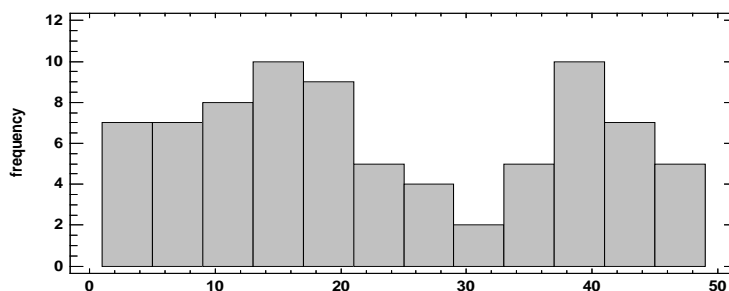
a) ¿Crees que los resultados pueden proceder de un modelo de distribución que conozcas? En caso afirmativo, ¿cuál sería? Justifica la respuesta. (2 puntos)

b) ¿Cuál sería el parámetro de posición más adecuado en este caso? ¿Por qué? Calcula su valor aproximado. (2 puntos)

c) ¿Cuál sería el parámetro de dispersión más adecuado en este caso? ¿Por qué? Calcula su valor aproximado. (2 puntos)

d) ¿Qué se puede decir en este caso acerca del valor del coeficiente de asimetría estandarizado? (2 puntos)

e) En un informe se indica que, a partir de los mismos datos, se ha obtenido otra representación gráfica, la cual se muestra a continuación. ¿Cuál es el nombre de dicho gráfico? ¿Crees que el gráfico es compatible con el de caja-bigotes mostrado anteriormente, o por el contrario se trata de un error en el informe? Justifica convenientemente tu respuesta. (2 puntos)



2. En una escuela de informática se ha llevado a cabo una encuesta sobre la marca del ordenador portátil que tienen los alumnos y se han obtenido los siguientes resultados:

	TOSHIBA	ACER	HP
CHICOS	120	80	20
CHICAS	60	40	10

Considerando que el conjunto de alumnos encuestados constituye la población objeto de estudio y asumiendo que cada alumno solamente dispone de un portátil, contesta a las siguientes preguntas definiendo claramente los sucesos considerados y justificando adecuadamente tus respuestas.

- a) ¿Cuál es el porcentaje de alumnas encuestadas? *(2 puntos)*
- b) ¿Cuál es la probabilidad de que un alumno sea chico y tenga un ordenador de marca ACER? *(2 puntos)*
- c) Si el ordenador pertenece a una alumna, ¿cuál es la probabilidad de que sea HP? *(2 puntos)*
- d) Sean los sucesos $A = \text{“ser chico”}$ y $B = \text{“tener ordenador de la marca TOSHIBA”}$, ¿son independientes dichos sucesos? *(2 puntos)*
- e) Sabiendo que el ordenador no es de la marca ACER, ¿cuál es la probabilidad de que pertenezca a una alumna? *(2 puntos)*

3. Una empresa fabrica un tipo de tarjeta de red, las cuales se venden en lotes de gran tamaño. La empresa ha constatado que en promedio hay un 1% de tarjetas defectuosas en cada lote. Para evaluar la calidad del lote se toma una muestra aleatoria de 20 tarjetas, y se determina el número de unidades defectuosas en dicha muestra.

a) Define la variable aleatoria e indica su distribución. *(2 puntos)*

b) Calcula la probabilidad de encontrar como máximo 2 tarjetas defectuosas. *(4 puntos)*

c) Calcula el número de tarjetas de red x , de forma que la probabilidad de encontrar como máximo x tarjetas defectuosas sea, como mucho, del 95%. *(4 puntos)*

4. El tiempo de ocupación de un terminal informático por cada usuario sigue un modelo exponencial, siendo la mediana de 5 minutos.

a) Con el fin de optimizar adecuadamente el sistema, ¿cuál debe ser el tiempo de ocupación t de manera que el 90% de los usuarios haya finalizado sus consultas?
(3 puntos)

b) Si llega un nuevo usuario y encuentra que el terminal está siendo ocupado por otro que lleva ya 5 minutos trabajando, ¿cuál es la probabilidad de que el usuario tenga que esperar otros 5 minutos adicionales hasta que quede libre el terminal? Comenta el resultado obtenido e indica cuál es la propiedad que permite el cálculo de la probabilidad solicitada.
(4 puntos)

c) Repetir el cálculo de probabilidad del apartado anterior asumiendo que el tiempo de ocupación del terminal por cada usuario fuese una variable uniforme comprendida entre cero y 15 minutos.
(3 puntos)

SOLUCIÓN

1a) Los datos de tiempo pueden modelizarse por medio de una distribución exponencial ya que el gráfico box-whisker indica una distribución asimétrica positiva y el primer cuartil es próximo a cero.

1b) La mediana es un parámetro de posición más representativo que la media en distribuciones muy asimétricas, como es este caso, ya que no está condicionada por la presencia de valores extremos (que se muestran en el gráfico). Corresponde a la línea vertical del interior de la caja (valor aproximado: **6,5**).

1c) El intervalo intercuartílico (IIC), también llamado rango o recorrido intercuartílico, es un parámetro de dispersión más apropiado que la varianza o el rango en distribuciones muy asimétricas porque no está afectado por los valores extremos. Su valor aproximado es: $IIC = C_3 - C_1 = 16 - 2 = 14$ meses

1d) Cuando el coeficiente de asimetría estandarizado (CAE) es pequeño (entre -2 y 2), puede asumirse que los datos son una muestra aleatoria tomada de una distribución simétrica. En este caso, el gráfico box-whisker indica claramente una distribución asimétrica positiva porque el bigote derecho es bastante más largo que el izquierdo y la mediana está desplazada hacia la izquierda. Por tanto, el CAE será superior a 2.

1e) Este gráfico es un histograma (también llamado “histograma de frecuencias”). El diagrama box-whisker indica claramente una distribución asimétrica positiva, pero esto no se muestra en el histograma, así que ambos gráficos no corresponden a los mismos datos. Según el histograma hay 27 valores superiores a 32, pero el box-whisker solamente muestra 6 datos. Sumando la altura de las barras, el histograma se ha construido con 79 valores, mientras que el box-whisker se ha obtenido con 100. El histograma parece reflejar una posible mezcla de poblaciones en los datos representados, lo que corrobora el hecho de que los datos utilizados en los dos gráficos son diferentes y como sugiere el enunciado se trataría de un error en el informe.

2a) Suceso F: la estudiante es mujer. Hay 110 alumnas y 220 alumnos (330 en total). Aplicando la regla de Laplace, hay 110 individuos que satisfacen el suceso: $P(F) = 110/330 = 1/3 = 33,3\%$

2b) Suceso H: el estudiante es hombre. Suceso A: el estudiante tiene un portátil ACER. Aplicando la regla de Laplace, hay 80 individuos que satisfacen ambos sucesos a la vez de modo que: $P(M \cap A) = 80/330 = 0,242$

2c) Suceso M: la estudiante es mujer. Suceso HP: tiene un portátil de marca HP. Hay 10 mujeres con portátil HP en el subconjunto de 110 mujeres. Por tanto, aplicando la regla de Laplace: $P(HP/M) = 10/110 = 0,091$

2d) Dos sucesos son independientes si satisfacen:

$$P(A/B) = P(A) ; P(A/B) = 120/180 = 2/3; \quad P(A) = 220/330 = 2/3$$

$$P(B/A) = P(B); P(B/A) = 120/220 = 0,545; P(B) = 180/330 = 0,545$$

$$P(A \cap B) = P(A) \cdot P(B); \quad 120/330 = (220/330) \cdot (180/330) = 0,364$$

Dado que se cumplen estas condiciones, A y B son sucesos independientes.

2e) Suceso M: la estudiante es mujer. Suceso A: tiene un portátil de marca ACER. Entre los 210 portátiles que no son ACER (180 Toshiba y 30 HP), 70 de ellos pertenecen a chicas. Aplicando la regla de Laplace:

$$P(M/\bar{A}) = \frac{60+10}{180+30} = \frac{70}{210} = \frac{1}{3} = \mathbf{0,33}$$

3a) Variable aleatoria X: número de unidades defectuosas encontradas en la muestra de 20 tarjetas de red.

Esta variable X sigue un modelo de distribución binomial B (n=20, p=0,01) ya que el valor mínimo es 0 y el máximo es 20.

$$\begin{aligned} \mathbf{3b)} \quad P(X \leq 2) &= P(X=0) + P(X=1) + P(X=2) = \binom{20}{0} \cdot 0,01^0 \cdot 0,99^{20} + \\ &+ \binom{20}{1} \cdot 0,01^1 \cdot 0,99^{19} + \binom{20}{2} \cdot 0,01^2 \cdot 0,99^{18} = 0,81791 + 0,16523 + 0,01586 = 0,9990 \end{aligned}$$

Se obtiene el mismo resultado aproximando la Binomial a una distribución Poisson con $\lambda = 0,2$ leyendo en la curva 2 del ábaco.

3c) $P(X \leq x) < 0,95$. Según el resultado obtenido en 3b):

$P(X=0) = 0,82$; $P(X \leq 1) = 0,82 + 0,16 = 0,98$ que es $> 0,95$. Por tanto, la solución tiene que ser cero; es decir, no es posible encontrar ningún valor x distinto de cero que satisfice la condición requerida.

$$\mathbf{4a)} \quad P(T > 5) = 0,5 = e^{-\alpha \cdot 5} ; \quad \alpha = -(\ln 0,5)/5 = 0,1386$$

$$P(T < t) = 0,9; \quad P(T > t) = 0,1 = e^{-0,1386 \cdot t} ; \quad t = -(\ln 0,1)/0,1386 = \mathbf{16,61 \text{ minutos}}$$

4b) T: tiempo de ocupación del terminal. La probabilidad de esperar más de 10 minutos es equivalente a la probabilidad de que el terminal esté ocupado más de 10 minutos. Según la propiedad de falta de memoria de la exponencial, y teniendo en cuenta que 5 es el valor de la mediana, la probabilidad que nos piden es: $P[(T > 10)/(T > 5)] = P(T > 5) = \mathbf{0,5}$

4c) En este caso, $X \approx U(0; 15)$; $P(T < t) = (t-0)/(15-0)$

$$P[(T > 10)/(T > 5)] = \frac{P[(T > 10) \cap (T > 5)]}{P(T > 5)} = \frac{P(T > 10)}{P(T > 5)} = \frac{1 - P(T < 10)}{1 - P(T < 5)} = \frac{1 - 10/15}{1 - 5/15} = 0,5$$