SOLUCIÓN EXÁMENES

ESTADÍSTICA I

INFORMÁTICA SISTEMAS

CENTRO DE PALMA DE MALLORCA

ESTADÍSTICA I (SISTEMAS Y GESTIÓN)

Ejercicios de Estadística Descriptiva tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001

A) SISTEMAS

1. *(Febrero 1995, 1ª semana)*[1] El nº de operaciones realizadas en una hora en cada uno de los 10 terminales de un determinado dispositivo han sido:

3, 5, 4, 2, 3, 3, 6, 5, 6, 3.

Se pide:

a) Establecer la distribución de frecuencias calculando las frecuencias absolutas y las relativas.

b) Calcular los valores de posición (media y moda) y de dispersión (desviación típica, varianza y coeficiente de variación).

Cálculos auxiliares: $\sum x_i = 40$; $\sum x_i^2 = 178$

2. *(Febrero 1996)* Durante una semana se han contado las operaciones de teleproceso realizadas en una entidad de crédito, obteniéndose, en función de las variables *X = Día de la semana* e *Y = Hora de realización de la operación*, los resultados siguientes:

X / Y	8 a 10	10 a 12	12 a 14
Lunes	3	5	4
Martes	2	4	3
Miércoles	1	2	3
Jueves	3	6	6
Viernes	6	6	6

a) Calcular la distribución marginal de la variable X y la de Y condicionada a *X = Jueves*, utilizando las frecuencias relativas.

b) Si por cuestiones de densidad de tráfico en las líneas, las tarifas establecidas por la compañía telefónica, en función de la hora en la que se produce la operación son:

[1] En la segunda semana pusieron un ejercicio "calcado" a éste, por lo cual no lo reproduje aquí.

Entre las 8 y las 10: 2 ptas. (por operación),

Entre las 10 y las 12: 3 ptas.,

Entre las 12 y las 14: 4 ptas.,

calcular el coste medio por operación en esa semana y qué día de la semana ha resultado más barato el coste medio.

3. *(Febrero 1998)* De la observación de una variable X se han recogido los siguientes valores: 1, 2, 1, 3, 4, 2, 2, 1, 3, 2.

Se pide:

a) Construir la tabla de frecuencias *(absoluta, relativa y relativa acumulada).*

b) Calcular la media de dicha distribución.

c) Obtener una medida de la dispersión respecto a la media.

4. *(Febrero 1998)* La tabla siguiente muestra la distribución bidimensional de las variables X e Y *(en frecuencias relativas),* para una colección de individuos determinada.

X / Y	**1**	**2**	**3**	**4**	**5**
1	0,20	0,06	0,04	0,02	0,01
2	0,01	0,02	0,10	0,01	0,02
3	0,02	0,10	0,05	0,01	0,01
4	0,08	0,15	0,04	0,03	0,02

Se pide:

a) Obtener la distribución de frecuencias *marginal* correspondiente a la variable X.

b) Calcular la media y la desviación típica de la distribución de Y *condicionada* por X = 2.

5. *(Septiembre 1998)* La tabla siguiente contiene las frecuencias *relativas acumuladas* obtenidas a partir de una colección de observaciones de la variable discreta X.

X	**Frec. rel. acum.**
-2	0,1
-1	0,3
0	0,4
1	0,7
2	0,9
3	1,0

Se pide:

a) Calcular la media de dicha distribución.

b) Obtener el porcentaje de observaciones que cumplen la condición $\{0 \leq X < 3\}$.

6. *(Febrero 1999)* Con el fin de describir el comportamiento de una variable cuantitativa X bajo unas condiciones determinadas, se procedió a su observación. Los datos registrados se resumen en la siguiente tabla de frecuencias:

X	0	1	2	3	4
Frec. abs. acumulada	2	3	7	8	10

a) Calcule la media de esta distribución.

b) Obtenga una medida de la dispersión.

7. *(Febrero 1999)* La tabla siguiente resume los datos obtenidos en la observación de una variable cuantitativa X.

X	1	2	3	4	5
Frec. relativa	0,1	0,2	0,4	0,1	0,2

Se pide:

a) Calcular el coeficiente de variación.

b) Obtener los percentiles correspondientes al 25% y al 75%.

8. *(Septiembre 1999)* La tabla siguiente resume los datos obtenidos en la observación de una variable cuantitativa X.

X	1	2	3	4	5
Frec. relat. acumulada	0,1	0,4	0,8	0,9	1,0

a) Calcular la media y la mediana.

b) Obtener el coeficiente de variación.

c) ¿Con qué frecuencia se obtienen los resultados que cumplen la condición $\{3 \leq X \leq 5\}$?

B) GESTIÓN

9. *(Febrero 1995)*[2] El nº de operaciones erróneas realizadas por 10 operadores durante un día han sido: 4, 6, 5, 3, 4, 4, 7, 6, 7 y 4; calcular el nº medio de operaciones erróneas, la varianza y el coeficiente de variación.

Cálculos auxiliares: $\sum_{i=1}^{10} x_i = 50$; $\sum_{i=1}^{10} x_i^2 = 268$.

10. *(Febrero 1995)* [2] La distribución conjunta de dos variables estadísticas X e Y está dada por la tabla:

X / Y	1	2	3
1	4	2	1
2	3	6	2
3	2	5	8
4	3	5	9

Hallar la distribución marginal de X y su media; hallar la distribución de Y condicionada por X = 3.

Soluciones

1. Media = 4, moda = 3, varianza = 1'80, desviación típica = 1'34, coeficiente de variación = 0'3350.
2. Coste medio = 3'12, día de la semana = viernes.
3. Media = 2'1, varianza = 0'89.
4. Media = 3'06, desviación típica = 0'9741.
5. Media = 0'6, porcentaje de observaciones = 60%.
6. Media = 2, varianza = 1'8.
7. Coeficiente de variación = 0'3935, percentil del 25% = 2, percentil del 75% = 4.
8. Media = 2'8, mediana = 3, coeficiente de variación = 0'3857, frecuencia $\{3 \leq X \leq 5\}$ = 60%.
9. Media = 5, varianza = 1'8, coeficiente de variación = 0'2683.
10. Media = 2'84.

[2] De acuerdo con las instrucciones para el curso 2001/2002 de la página web de la asignatura, los exámenes de cursos anteriores al 1998/1999 NO SON ORIENTATIVOS para preparar la materia del curso actual.

<u>ESTADÍSTICA I (SISTEMAS Y GESTIÓN)</u>

<u>Ejercicios de Probabilidad tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001</u>

A) SISTEMAS

1. *(Febrero 1995)* Un proceso se ejecuta en uno cualquiera de los 10 periféricos de un dispositivo, elegido el periférico por el propio proceso al azar. Hay periféricos de tres tipos: 5 de tipo B_1, 3 de tipo B_2 y 2 de tipo B_3. Cada periférico puede producir un fallo en el proceso con diversa probabilidad. La probabilidad de fallo si el periférico es de tipo B_1, es 0'01; si es de tipo B_2, es 0'02; y si es de tipo B_3, es 0'04.
a) Calcular la probabilidad de que se produzca un fallo en el proceso.
b) Supuesto que se ha producido el fallo, ¿cuál es la probabilidad de que el periférico en el que se ha realizado el proceso sea del tipo B_3?

2. *(Febrero 1995)* Una planta de montaje utiliza microcircuitos suministrados por tres fabricantes F_1, F_2 y F_3. El 40% de dichos elementos se compra a F_1, el 35% a F_2 y el 25% restante a F_3. El porcentaje de circuitos defectuosos para F_1, F_2 y F_3 es 5, 7 y 8% respectivamente. Sabiendo que los circuitos se almacenan en la planta sin tener en cuenta su procedencia,
a) determinar la probabilidad de que una unidad montada en la planta contenga un circuito defectuoso;
b) calcular la probabilidad de que un circuito defectuoso proceda del fabricante F_2.

3. *(Febrero 1996)* Un sistema está constituido por tres componentes, A, B y C, que funcionan de forma independiente. Las probabilidades de fallo de cada componente son: P(A) = 0'1, P(B) = 0'2 y P(C) = 0'3. Se pide calcular:
a) La probabilidad de que fallen dos componentes del sistema, exactamente.
b) La probabilidad de que si fallan dos componentes, sean A y B.

4. *(Septiembre 1996)* Cuando un determinado proceso de fabricación opera bajo control produce un 5% de unidades defectuosas. Por el contrario, cuando se encuentra fuera de control el porcentaje de elementos defectuosos es 30%. La probabilidad de que el proceso se encuentre bajo control es 0'92. Si una unidad seleccionada al azar resulta defectuosa, ¿cuál es la probabilidad de que el proceso funcione correctamente?

5. *(Febrero 1997)* Un jugador tiene en su bolsillo dos dados de tipo A, tres de tipo B y cuatro de tipo C. Los dados de tipo A tienen tres caras con el número 3 y otras tres con el 6. Los de tipo B tienen dos caras con cada uno de los números 2, 4 y 6. Los de tipo C tienen una cara con cada uno de los números del 1 al 6. Si el jugador elige al azar un dado y lo lanza, se pide:
a) Calcular la probabilidad de que salga un 6.
b) Suponiendo que haya salido un 3, calcular la probabilidad de que el dado elegido haya sido del tipo A.

6. *(Febrero 1997)* Una urna contiene cuatro bolas de cada uno de los tres colores blanco, rojo y negro. Se extraen tres bolas al azar sin reemplazamiento. Calcular la probabilidad de los siguientes sucesos:

a) Que las tres bolas sean de colores diferentes.

b) Que las tres bolas sean del mismo color.

c) Que haya exactamente dos bolas del mismo color.

7. *(Febrero 1998)* Se lanza un dado, y a continuación, tantas monedas equilibradas como puntos se obtuvieron en el lanzamiento del dado. Determinar:

a) La probabilidad de obtener cuatro caras.

b) La probabilidad de que el resultado del lanzamiento del dado fuera un 5 si se obtuvieron cuatro caras.

8. *(Febrero 1998)* Tenemos dos urnas: A con una bola blanca y dos negras, y B con dos blancas y dos negras. A continuación se extraen dos bolas (una de cada urna) y se cambian de urna. Determinar:

a) La probabilidad de que una nueva bola extraída de A sea blanca.

b) La probabilidad de que la bola que se obtuvo en B fuera negra si la obtenida en la segunda extracción de A es blanca.

9. *(Septiembre 1998)* De una urna que contiene tres bolas blancas y n bolas negras, $n \geq 4$, se extraen cuatro bolas sin reemplazamiento. Determinar la probabilidad de que:

a) No todas sean del mismo color.

b) El número de negras supere al de blancas.

c) No haya el mismo número de blancas que de negras.

10. *(Febrero 1999)* Se sabe que la cuarta parte de los habitantes de una gran ciudad tiene menos de 50 años y que la probabilidad de que una habitante muera de cáncer es 0’1 en el caso de que tenga menos de 50 años, y 0’2 en otro caso.

a) Determinar la probabilidad de que un habitante, elegido al azar, no muera de cáncer.

b) Si una persona murió de cáncer, ¿cuál es la probabilidad de que tuviera menos de 50 años?

11. *(Febrero 1999)* Una fábrica tiene tres plantas. El 20% de sus máquinas se montan en la primera planta, el 30% en la segunda y el 50% restante en la tercera. Se sabe además que la cadena de montaje falla, dando lugar a una máquina defectuosa, con probabilidades 0’1, 0’2 y 0’15 en cada una de las tres plantas.

a) Determinar la probabilidad de que una máquina cualquiera montada en esa fábrica sea defectuosa.

b) ¿Cuál es la probabilidad de que una máquina no defectuosa se haya montado en la primera planta?

12. *(Septiembre 1999)* De una urna que contiene *n* bolas blancas y tres negras ($n \geq 4$) se realizan cuatro extracciones sin reemplazamiento. Determinar la probabilidad de que:

a) Se hayan extraído todas las bolas negras de la urna.

b) No todas las bolas extraídas sean del mismo color.

13. *(Febrero 2000)* De los sucesos A, B y C se sabe que son independientes y que $P(B^c) = \frac{7}{8}$, $P(A \cap B) = \frac{1}{32}$, $P(A \cap C) = \frac{1}{8}$. Determínese:

a) P(A B)

b) P(B C /A)

14. *(Febrero 2001)* El conjunto de habitantes en situación de empleo de una región determinada se ha segmentado, en función de sus ingresos anuales, en tres grupos: el 20% tiene unos ingresos bajos, el 70% unos ingresos medios y el resto ingresos altos. Se sabe que en cada grupo la proporción de titulados universitarios es de 0'1, 0'9 y 0'3, respectivamente.
a) Si se elige al azar a un habitante de dicha región, ¿cuál es la probabilidad de que sea titulado universitario?
b) Si el individuo seleccionado es titulado universitario, ¿a cuál grupo pertenecerá más probablemente?

15. *(Septiembre 2001)* De una urna que contiene dos bolas blancas, tres bolas negras y una bola roja ser realizan cuatro extracciones sin reposición. Hallar la probabilidad de que:
a) entre las bolas extraídas aparezcan todos los colores;
b) se obtengan exactamente dos colores.

B) GESTIÓN

16. *(Febrero 1996)*[1] Una urna contiene a bolas blancas y b bolas negras. Se toman, sucesivamente, tres bolas al azar sin reemplazamiento.
a) Si la primera bola elegida era negra, ¿cuál es la probabilidad de que la tercera sea blanca?
b) Si la tercera bola elegida resultó ser negra, ¿cuál es la probabilidad de que la primera fuese blanca?

17. *(Febrero 1996)*[1] De una urna que contiene n bolas de cada uno de los colores: blanco, negro, rojo y azul, se toman cuatro bolas al azar. Hallar:
a) La probabilidad de que alguna bola sea azul.
b) La probabilidad de que haya k colores diferentes, $1 \leq k \leq 4$.

18. *(Febrero 1996)*[1] Un proceso se puede ejecutar en uno cualquiera de tres periféricos P, Q, R, con el siguiente protocolo: en el primer intento se elige un periférico al azar; si está listo, se ejecuta en él el proceso; si no lo está, se realiza el segundo intento, que consiste en elegir al azar uno de los dos periféricos restantes; si está listo, se ejecuta en él el proceso y, si no lo está, se termina el proceso sin ejecutar.
Cada periférico está listo o no con independencia del estado de los demás y de la elección del proceso. La probabilidad de que P esté listo es 0'80, la de Q es 0'90 y la de R es 0'95.
a) Hallar la probabilidad de que el proceso no se ejecute.
b) Hallar la probabilidad de que se ejecute en el dispositivo Q.
c) Si se ha ejecutado, hallar la probabilidad de que sea en el segundo intento.

19. *(Septiembre 1996)*[1] n bolas se colocan al azar en n urnas, U_1, U_2, ... , U_n. Calcular la probabilidad de los siguientes sucesos:
a) La urna U_1 está vacía.
b) Alguna urna está vacía.
c) La urna U_1 es la única vacía.
d) Sólo hay una urna vacía.

20. *(Septiembre 1996)*[1] Tres urnas U_1, U_2 y U_3 contienen: a bolas blancas y b bolas negras, a bolas blancas y a negras y b bolas blancas y b bolas negras, respectivamente. Tomamos una bola de cada urna. Si en total hay dos bolas blancas, hallar la probabilidad de que la bola elegida de la urna U_2 sea blanca.

[1] De acuerdo con las instrucciones para el curso 2001/2002 de la página web de la asignatura, los exámenes de cursos anteriores al 1998/1999 NO SON ORIENTATIVOS para preparar la materia del curso actual.

21. *(Febrero 1997)*[1] N tarjetas numeradas de 1 a N se barajan al azar y se examina el orden en que han quedado. Calcular la probabilidad de cada uno de los sucesos siguientes:
a) La tarjeta 2 es la primera.
b) La tarjeta 2 es posterior a la 1.
c) La tarjeta 2 es la siguiente a la 1.
d) Hay *r* tarjetas entre la 1 y la 2.

22. *(Febrero 1997)*[1] N tiradores disponen de *k* cartuchos cada uno y disparan independientemente. El tirador *i* tiene probabilidad p_i de acertar en cada disparo, $1 \leq i \leq N$, y cada tirador deja de disparar cuando hace diana. Hallar la probabilidad de cada uno de los sucesos siguientes:
a) Algún tirador hace diana.
b) Todos los tiradores hacen diana.
c) Entre todos, sólo consiguen una diana.
d) Si $p_1 = p_2 = ... = p_n = p$, hallar la probabilidad de que, entre todos, consigan *r* dianas.

23. *(Febrero 1997)*[1] Un dado tiene dos caras marcadas con 1, dos caras marcadas con 2 y dos caras marcadas con 3. Se lanza, repetidas veces, hasta que la suma de los resultados obtenidos es mayor que 5. ¿Cuál es la probabilidad de que sea mayor que 7?

24. *(Septiembre 1997)*[1] Una urna contiene *n* bolas numeradas de 1 a *n*. Se extraen *k* bolas sin reposición.
a) Hallar la probabilidad de que el mayor de los números que aparecen sea *r*.
b) ¿Cuál es el número mínimo de extracciones que hay que hacer para que pueda asegurarse que el número *n* aparecerá con una probabilidad mayor que 0'99?

25. *(Febrero 1998)*[1] Un sistema electrónico está formado por dos subsistemas A y B. La probabilidad de que A falle es 0'1; la probabilidad de que B falle y A no es 0'05 y la probabilidad de que ambos fallen es 0'08. Calcular:
a) P(A falle / B ha fallado).
b) P(A falle y B no).

26. *(Febrero 1998)*[1] Un conjunto C tiene *n* elementos. Se elige un subconjunto de C al azar. ¿Cuál es la probabilidad de que tenga intersección no vacía con otro subconjunto dado, B, $B \subset C$?

27. *(Febrero 1998)* [1] Un proceso se puede ejecutar en uno cualquiera de tres periféricos P_1, P_2, P_3, con el siguiente protocolo: primero, se intenta ejecutar en P_1; si está listo, se ejecuta en él el proceso; si no lo está, se realiza el segundo intento, que consiste en elegir al azar uno de los dos periféricos restantes; si está listo, se ejecuta en él el proceso y, si no lo está, se termina el proceso sin ejecutar.
Cada periférico está listo o no con independencia del estado de los demás y de la elección del proceso. La probabilidad de que P_i esté listo es p_i, i = 1, 2, 3.
a) Hallar la probabilidad de que el proceso no se ejecute.
b) Hallar la probabilidad de que se ejecute en el dispositivo P_2.
c) Si no se ha ejecutado, hallar la probabilidad de que sea en el 2º intento.

28. *(Septiembre 1998)*[1] *n* bolas numeradas del 1 al *n* se introducen al azar en *n* urnas. Llamaremos puntuación de una determinada urna, a la máxima numeración de las bolas que contiene; en el caso de que una urna quede vacía convendremos que su puntuación es de 0. Determinar la probabilidad de que:
a) La puntuación de la primera urna sea *k*.
b) La primera urna esté vacía y la última tenga puntuación *k*.

29. *(Febrero 1999)* De un conjunto de *n* monedas se escoge, al azar, un subconjunto cualquiera; luego, se lanzan las monedas escogidas. ¿Cuál es la probabilidad de obtener *k* caras?

30. *(Febrero 1999)* Sobre un espacio de probabilidades fijo, consideramos dos sucesos disjuntos A_1 y A_2, con probabilidades respectivas $P(A_1) = 0'3$ y $P(A_2) = 0'2$. Sea B otro suceso tal que $P(B/A_1) = 0'5$ y $P(B/A_2) = 0'7$. Calcular $P(B/A_1 \cup A_2)$.

31. *(Septiembre 1999)* De una urna que contiene *n* bolas blancas y *m* bolas negras se extrae al azar un subconjunto cualquiera de bolas.
a) Determinar la probabilidad de que ese subconjunto contenga alguna bola blanca.
b) Determinar la probabilidad de que dicho subconjunto contenga exactamente una bola blanca.

32. *(Febrero 2000)*
a) Se lanza un dado tres veces. Hallar la probabilidad de que la suma de los resultados obtenidos sea par.
b) Se lanza un dado cuatro veces. Hallar la probabilidad de que la suma de los resultados obtenidos sea par.

33. *(Febrero 2000)* Se lanza tres veces una moneda. Si sabemos que ha aparecido al menos una cara, ¿cuál es la probabilidad de que hayan aparecido dos o más?

34. *(Febrero 2000)* Una urna contiene 6 bolas numeradas del 1 al 6. Se extraen las bolas, una a una sin reemplazamiento, hasta que todas las bolas han sido extraídas. Determínese:
a) La probabilidad de que la última bola lleve un número par.
b) La probabilidad de que todas las bolas con número par aparezcan antes que la primera con número impar.

35. *(Septiembre 2000)* Una urna contiene diez tarjetas numeradas del 1 al 10. Se extraen tres tarjetas sin reemplazamiento.
Determinar la probabilidad de que:
a) Las tres tarjetas tengan número par.
b) La suma de los números obtenidos sea par.

36. *(Septiembre 2000)* Se lanzan dos monedas, y a continuación, a una urna que contiene una bola blanca se le añaden tantas bolas negras como caras hayan obtenido. Luego, se extrae una bola, al azar, de la urna. ¿Cuál es la probabilidad de que sea blanca?

37. *(Febrero 2001)* De una urna que contiene *a* bolas blancas, *b* bolas negras y *c* rojas se extraen tres bolas sin reposición.
Determinar:
a) La probabilidad de que todas sean del mismo color.
b) La probabilidad de que aparezcan al menos dos colores.

38. *(Febrero 2001)* De una urna que contiene *n* bolas de cada uno de los colores: blanco, negro y rojo, se realizan tres extracciones sin reemplazamiento. Determinar:
a) La probabilidad de obtener una bola blanca antes que una negra.
b) La probabilidad de que aparezca una bola roja entre dos bolas de colores distintos.

39. *(Febrero 2001)* Dos jugadores A y B compiten en un juego consistente en lanzar por turno una moneda equilibrada, comenzando por el jugador A. El juego termina cuando alguno de los dos obtiene cara; en este caso el jugador que obtuvo la cara es el ganador.
a) Hallar las probabilidades de ganar de cada jugador.
b) Si cada uno de ellos pone en su turno una peseta y el ganador se lo lleva todo, ¿cuál es el beneficio esperado de cada jugador?

40. *(Septiembre 2001)* Una clave está formada por un número de cuatro dígitos:
$x_1\ x_2\ x_3\ x_3$
donde $1 \leq x_i \leq 9$. Un dispositivo genera aleatoriamente números de cuatro dígitos, hasta que acierta con la clave.
a) Si el dispositivo, en cada intento, escoge cada dígito independientemente de los demás, entre los nueve posibles con igual probabilidad, ¿cuál es la probabilidad de que encuentre la clave en el k-ésimo intento?
b) Si el dispositivo, en cada intento, escoge un número al azar entre todos los posibles menos los ya probados, ¿cuál es la probabilidad de que encuentre la clave en el k-ésimo intento?

Soluciones

1. (a) 0'0190. (b) 0'4211.
2. (a) 0'0645. (b) 0'0380.
3. (a) 0'11. (b) 0'22. (0,1818)
4. 0'6571.
5. (a) $\frac{8}{27}$. (b) $\frac{3}{5}$.
6. (a) $\frac{16}{55}$. (b) $\frac{3}{55}$. (c) $\frac{36}{55}$.
7. (a) $\frac{29}{384}$. (b) $\frac{10}{29}$.
8. (a) $\frac{17}{45}$. (7/18) (b) $\frac{14}{17}$. (6/7)
9. (a) $\frac{12n^2+12}{(n+3)(n+2)(n+1)}$; (b) $\frac{12(n-1)(n-2)}{(n+3)(n+2)(n+1)}$; (c) $\frac{12n^2-36n+48}{(n+3)(n+2)(n+1)}$.
10. (a) 0'825. (b) 0'1429.
11. (a) 0'1550. (b) 0'8710.
12. (a) $\frac{24}{n^3+6n^2+11n+6}$; (b) $\frac{12n^2+12}{n^3+6n^2+11n+6}$
13. (a) $\frac{11}{32}$. (b) $\frac{9}{16}$.
14. (a) 0'68. (b) Al grupo con ingresos medios.
15. (a) $\frac{3}{5}$. (b) $\frac{2}{5}$.

16. (a) y (b) $\frac{a}{a+b-1}$.

17. $p_{[1]} = \frac{\binom{4}{1}\binom{n}{4}}{\binom{4n}{4}}$; $p_{[2]} = \frac{\binom{4}{2}\binom{n}{2}^2 + \binom{4}{1}\binom{3}{1}\binom{n}{3}\binom{n}{1}}{\binom{4n}{4}}$; $p_{[3]} = \frac{\binom{4}{1}\binom{3}{2}\binom{n}{2}\binom{n}{1}^2}{\binom{4n}{4}}$;

$p_{[4]} = \frac{\binom{n}{1}^4}{\binom{4n}{4}}$.

18. (a) 0'0187. (b) 0'3375. (c) 0'0998.

19. (a) $\left(\frac{n-1}{n}\right)^n$; (b) $1-\frac{n!}{n^n}$; (c) $\frac{n!(n-1)}{2n^n}$; (d) $\binom{n}{2}\frac{n!}{n^n}$.

20. $\frac{a+b}{2a+b}$.

21. (a) $\frac{1}{N}$. (b) $\frac{1}{2}$. (c) $\frac{1}{N}$. (d) $\frac{2(N-r-1)}{N(N-1)}$.

22. (a) 1 - $\prod_{i=1}^{N}(1-p_i)^k$; (b) $\prod_{i=1}^{N}\left[1-(1-p_i)^k\right]$;

(c) $\sum_{i=1}^{N}\left[1-(1-p_i)^k\right]\prod_{j\#i}(1-p_i)^k$; (d) $\binom{N}{r}\left[1-(1-p)^k\right]^r(1-p)^{k(N-r)}$.

23. 0’4979.

24. (a) $\frac{\binom{r-1}{k-1}}{\binom{n}{k}}$. (b) 0’99n.

25. (a) 0’6154. (b) 0’02.

26. $1-2^{-m}$ (m = nº de elementos de B).

27. (a) $1-p_1-(1-p_1)\frac{1}{2}(p_2+p_3)$.

(b) $(1-p_1)\ \frac{1}{2}p_2$

(c) $\frac{(1-p_1)\frac{1}{2}(p_2+p_3)}{p_1+(1-p_1)\frac{1}{2}(p_2+p_3)}$.

28. (a) $\frac{1}{2}\left(1-\frac{1}{n}\right)^{n-k}$ si $1 \le k \le n$; (b) $\left(1-\frac{1}{n}\right)^n$ si $k = 0$.

(b) $\frac{1}{n}\left(1-\frac{1}{n}\right)^{k-1}\left(1-\frac{2}{n}\right)^{n-k}$ si $1 \le k \le n$; (b) $\left(1-\frac{2}{n}\right)^n$ si $k = 0$.

29. $\binom{n}{k}\left(\frac{3}{4}\right)^n\left(\frac{1}{3}\right)^k$.

30. 0’58.

31. (a) $1-\dfrac{1}{2^n}$. (b) $\dfrac{n}{2^n}$.

32. (a) y (b) $\dfrac{1}{2}$.

33. $\dfrac{4}{7}$.

34. (a) $\dfrac{1}{2}$. (b) $\dfrac{1}{20}$.

35. (a) $\dfrac{1}{12}$. (b) $\dfrac{1}{2}$.

36. $\dfrac{7}{12}$.

37. (a) $\dfrac{\binom{a}{3}+\binom{b}{3}+\binom{c}{3}}{\binom{a+b+c}{3}}$. (b) $1-\dfrac{\binom{a}{3}+\binom{b}{3}+\binom{c}{3}}{\binom{a+b+c}{3}}$.

38. (a) $\dfrac{n(2n-1)}{(3n-1)(3n-2)}$. (b) $\dfrac{2n}{3(3n-1)}$.

39. (a) $p_A = \dfrac{2}{3}$; $p_B = \dfrac{1}{3}$. (b) $\left(-\dfrac{2}{9}\right)$ y $\dfrac{2}{9}$.

40[2]. (a) $\dfrac{6560^{k-1}}{(6561)^k}$. (b) $(6562\text{-}k)^{-1}$.

[2] Solución no “oficial” (pendiente de confirmar).

ESTADÍSTICA I (SISTEMAS Y GESTIÓN)

Ejercicios de Modelos Probabilísticos tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001

A) SISTEMAS

1. *(Febrero 1995)* La duración de una componente de un sistema sigue una distribución *normal* de media 2.000 horas y desviación típica 100. Calcular:

a) La probabilidad de que la componente dure entre 1.900 y 2.050 horas.
b) ¿Qué duración mínima de la componente puede garantizar el fabricante con una probabilidad de error del 5%?

Datos auxiliares: F(0'5) = 0'6915; F(1) = 0'8413; F(1'5) = 0'9332; F(1'645) = 0'95; F(2) = 0'9772, siendo F la función de distribución *normal* estándar.

2. *(Febrero 1995)* El tiempo que transcurre entre la llegada de dos trenes consecutivos a una estación de metro sigue una distribución de probabilidad exponencial de media igual a 5 minutos. Calcular:

a) La probabilidad de que tarde entre 4 y 5 minutos y medio.
b) ¿Qué tiempo máximo entre trenes puede garantizar la compañía metropolitana con probabilidad de error del 10%?.

Datos auxiliares: $e^{-0,8}$ = 0'4493; $e^{-1'1}$ = 0'3329; ln 0'1 = (–2'3026)

3. *(Septiembre 1995)* La probabilidad de que no haya errores en determinado proceso cuando se realiza es 0'9. Suponiendo que las diversas realizaciones del proceso son independientes, en una serie de 10 realizaciones,

a) Calcular la probabilidad de que no haya error en al menos 8 realizaciones.
b) Calcular el nº esperado de realizaciones sin errores y su desviación típica.

4. *(Septiembre 1995)* Suponiendo que la variable aleatoria X tiene distribución *normal* de media $\mu = 50$ y varianza $\sigma^2 = 100$, calcular:

a) El valor de C tal que $P(X \leq C)$ = 0'4207.
b) Dos números A y B, equidistantes de μ y tales que $P(A \leq X \leq B)$ = 0'966.

Datos auxiliares: Si Z tiene distribución N (0,1), $P_{0'20} = P(0 \leq Z \leq 0'20)$ = 0'0793; $P_{0'40}$ = 0'1554; $P_{1'22}$ = 0'3888; $P_{2'12}$ = 0'4830.

5. *(Febrero 1996)* El nº de errores que se producen diariamente en una transmisión de datos con un dispositivo sigue una distribución de Poisson de parámetro 5. Se pide:

a) Calcular la probabilidad de que en un día se produzcan al menos dos errores.
b) Calcular la probabilidad de que el nº de errores sea menor que la cifra resultante de restar de su esperanza su desviación típica.

Datos auxiliares: e^{-5} = 0'006738.

6. *(Febrero 1996)* Un fabricante de impresoras láser establece entre las especificaciones de su modelo HE-2500 que los tonner que utiliza tienen una vida media de 4.000 hojas con una desviación típica de 500. Suponiendo que la distribución del nº de hojas que se pueden imprimir con un tonner siga una distribución de probabilidad normal, se pide:

a) Calcular la probabilidad de que el nº de hojas que se imprimen con un tonner esté entre 3.600 y 4.200.
b) ¿Qué número H de hojas podría decir el fabricante que pueden imprimirse como mínimo con un tonner, asumiendo una probabilidad del 10% de equivocarse, es decir, de modo que la probabilidad de que el nº de hojas impresas sea menor que H, sea igual a 0'1?.

Datos auxiliares: F(0'4) = 0'6554; F(0'8) = 0'7881; F(1) = 0'8413; F(1'28) = 0'9; F(1'64) = 0'95, siendo F la función de distribución normal estándar.

7. *(Septiembre 1996)* La duración de cada operación que no se realiza en un terminal de teleproceso de una oficina bancaria es una variable aleatoria que sigue una distribución uniforme de media 10 segundos y varianza 3.

a) Calcular la probabilidad de que una operación dure menos de 11 segundos.
b) Calcular un valor C tal que la probabilidad de que el proceso dure más que C sea 0'80.

8. *(Febrero 1997)* Un sistema contiene tres componentes conectados en serie que funcionan independientemente. Suponiendo que el tiempo de vida de estos componentes (medido en horas) sigue una distribución *exponencial* de media 1.000, 3.000 y 6.000 horas, respectivamente, calcular la probabilidad de que el sistema no falle antes de 100 horas.

Datos auxiliares:

(i) La densidad *exponencial* es de la forma $f(x) = \begin{cases} (1 / \beta) \exp(- x / \beta), & x > 0 \\ 0, & \text{en otro caso} \end{cases}$

(ii) La conexión en serie implica que el sistema falla cuando lo hace alguno de sus componentes.

9. *(Septiembre 1997)* El tiempo (en horas) necesario para reparar una máquina es una variable aleatoria que se distribuye según una ley *exponencial* con parámetro $\beta = 2$. Se pide:

a) Calcular la probabilidad de que una reparación dure más de 2 horas.
b) Obtener la probabilidad de que se necesiten más de 10 horas, si se llevan 9 horas de reparación.

Datos auxiliares:

La densidad *exponencial* es de la forma $f(x) = \begin{cases} (1 / \beta) \exp(- x / \beta), & x > 0 \\ 0, & \text{en otro caso} \end{cases}$

10. *(Febrero 1998)* La vida útil de una componente electrónica puede considerarse como una variable aleatoria con distribución exponencial de media 5 años. Si se hacen 500 instalaciones de dicha componente, ¿cuál es la probabilidad de que en 2 años se produzcan más de 10 fallos?

Datos auxiliares: $e^{-0'4} = 0'67$.

La densidad exponencial es: $f_\beta(x) = (1 / \beta)\, e^{-x/\beta}$; $\beta \in \mathbb{R}^+$, $x > 0$.

11. *(Febrero 1998)* Si X es una variable aleatoria con distribución *normal* de media 2 y varianza 4, calcular la probabilidad de que:

a) X se encuentre en el intervalo (0, 4);

b) X tome un valor que diste de la media más de 2 unidades.

Datos auxiliares: $P(Z \leq -1) = 0'1587$, siendo Z una variable aleatoria con distribución N(0, 1).

12. *(Febrero 1999)* Una empresa compra grandes cantidades de una clase de dispositivos electrónicos. La decisión para aceptar o rechazar un lote de dichos elementos se toma en base a una muestra aleatoria de 100 unidades. Si el lote se rechaza cuando se encuentran 3 ó más unidades defectuosas, ¿cuál es la probabilidad de rechazar un lote con un 1% de dispositivos defectuosos?

13. *(Febrero 1999)* Se supone que el voltaje medido en cierto circuito eléctrico tiene una distribución *normal* con media 120 y desviación típica 2. Calcule la probabilidad de que una medición proporcione un voltaje:

a) Superior a 118.

b) Entre 116 y 118.

Datos auxiliares: F(1) = 0'8413 y F(2) = 0'9773, siendo F la función de distribución de la N(0, 1).

14. *(Septiembre 1999)* La probabilidad de error en la transmisión de un *bit* a través de un canal de comunicación es $(1 / 10^3)$.

a) Escriba una expresión para la probabilidad exacta de que se produzcan más de 3 errores cuando se transmite un bloque de 10^3 bits.

b) Calcule un valor aproximado para dicha probabilidad.

Observación: Suponga independencia.

Datos auxiliares: F(1) = 0'8413, F(2) = 0'9772, F(2'5) = 0'9938 y F(3) = 0'9987, siendo F la función de distribución de la N(0, 1).

15. *(Septiembre 2000)* El tiempo de vida – en miles de horas – de cierta clase de componentes electrónicas sigue una distribución *normal*. Se sabe que el 15'87% de los componentes dura al menos 104 mil horas y que el 54'68% tiene una duración que no dista más de 3 mil horas del tiempo medio de vida. Determínese la media y la varianza de dicha distribución.

Datos auxiliares: Si F es la función de distribución de la *normal* estándar, F(-1) = 0'1587, F(2) = 0'9772, F(0'1176) = 0'5468, F(0'75) = 0'7734.

16. *(Febrero 2001)* Un botánico ha observado que la longitud de las hojas de una determinada especie de árbol sigue una distribución N (μ, σ). Se sabe que el 15'87% de las hojas de uno de estos árboles tiene una longitud de al menos 6 cm. y que la longitud del 99% de las hojas no es superior a 7 cm. Determínese la media y la desviación típica de la distribución.

Datos auxiliares: Si F es la función de distribución de una N (0, 1), F(-1) = 0'1587, F(0'1587) = 0'564, F(-2'33) = 0'01.

B) GESTIÓN

17. *(Febrero 1995)*[1] La duración de un proceso sigue una distribución exponencial de media igual a 10 minutos. Hallar la probabilidad de que el proceso dure menos de 8 minutos y la probabilidad de que dure más de 15 minutos.

18. *(Febrero 1995)*[1] La duración de un sistema sigue una distribución normal de media 700 horas y varianza 10.000. Calcular la probabilidad de que la componente dure menos de 800 horas. Hallar la probabilidad de que dure más de 650 horas.

19. *(Septiembre 1995)*[1] Si X es una variable normal de media μ y varianza σ^2, calcular la probabilidad $P(\mu - \sigma \leq X \leq \mu + \sigma)$ y hallar $\lambda > 0$ para que se verifique: $P(\mu - \lambda\sigma \leq X \leq \mu + \lambda\sigma) = 0'95$.

20. *(Febrero 1997)*[1] El tiempo, en horas que tarda en averiarse una componente de un sistema es aleatorio, con función de densidad exponencial dada por: $f(t) = \lambda e^{-\lambda t}$, $\lambda > 0$, $t > 0$.
Las componentes se sustituyen por otra nueva cada vez que se averían o, por razones de seguridad, cuando llevan funcionando T horas.

a) Hallar el tiempo medio que funciona el sistema hasta la 1ª sustitución.
b) Si sólo se dispone de una componente de repuesto, hallar la probabilidad de que el sistema funcione más de 3T/2 horas.

21. *(Febrero 1998)*[1] Un sistema electrónico está formado por dos subsistemas A y B. La duración del subsistema A es una variable aleatoria X con función de densidad:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

con $\lambda > 0$. La duración del subsistema B es una variable aleatoria Y, independiente de X, que tiene función de densidad:

$$g(y) = \begin{cases} \mu e^{-\mu y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

con $\mu > 0$. Hallar $P(X + Y > t)$, $t > 0$. y el tiempo medio que transcurre hasta que falla algún subsistema.

22. *(Septiembre 1998)*[1] Se lanza un dado repetidas veces, hasta que aparezca el resultado del 1^er^ lanzamiento. Determinar la distribución del nº de lanzamientos realizados y su media.

23. *(Febrero 2001)* El tiempo de funcionamiento en horas de una componente de un sistema es una variable aleatoria exponencial con función de densidad:

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Por motivos de seguridad, esta componente se sustituye por una nueva cuando el sistema lleva funcionando T horas. Determinar:
a) la probabilidad de que la componente falle antes de ser sustituida.
b) El tiempo medio que funciona hasta que es sustituida.

[1] De acuerdo con las instrucciones para el curso 2001/2002 de la página web de la asignatura, los exámenes de cursos anteriores al 1998/1999 NO SON ORIENTATIVOS para preparar la materia del curso actual.

24. *(Febrero 2001)* El funcionamiento de una cadena de montaje depende de dos dispositivos que trabajan independientemente. La cadena falla cuando alguno de los dos dispositivos deja de funcionar. Se sabe que la distribución del tiempo de vida de cada uno de ellos es exponencial con densidades respectivas dadas por:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 1 & \text{en otro caso} \end{cases} \qquad f(y) = \begin{cases} \mu e^{-\mu y} & \text{si } y > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Determinar la distribución del tiempo de funcionamiento de la cadena y su media.

25. *(Septiembre 2001)* Una lista de *n* registros está ordenada al azar. Para buscar un registro se examina secuencialmente la lista empezando por el primero, hasta dar con el buscado. ¿Cuál es el número esperado de registros que habrá de examinar?

Soluciones

1. (a) 0’5328. (b) 1.835 horas.
2. (a) 0’1164. (b) 12 minutos (aprox.).
3. (a) 0’4305. (b) 7’2 y 0’8485.
4. (a) C = 48. (b) A = 28’8 y B = 71’2.
5. (a) 0’9596. (b) 0’1247.
6. (a) 0’4435. (b) 3.360 hojas.
7. (a) 2 / 3. (b) 8’2.
8. 0,8607.
9. (a) 0’3679. (b) 0’6065.
10. $1-\sum_{i=1}^{10}\binom{500}{i}0'33^i 0'67^{500-i} \cong 1$
11. (a) 0’6826. (b) 0’3174.
12. 0’2642.
13. (a) 0,8413. (b) 0’8186.
14. (a) $1-\sum_{x=0}^{3}\binom{1000}{x}(0'001)^x(0'999)^{1000-x}$. (b) 0’0908.
15. μ = 100; σ^2 = 16.
16. μ = 5’2481; σ = 0’7519.
17. 0’5507 y 0’2231, respectivamente.
18. 0’8413 y 0’6915, respectivamente.
19. 0’6826 y 1’96, respectivamente.
20. (a) $\frac{1}{\lambda}$ $(1 - e^{-\lambda T})$; (b) $(1 + \frac{\lambda T}{2})e^{-3\lambda T/2}$.
21. P(X + Y > t) = $\left(\frac{-\mu}{\lambda-\mu}\right)e^{-\lambda\cdot t} + \frac{\lambda}{\lambda-\mu}e^{-\mu\cdot t}$, t > 0; E (T) = $\frac{1}{\lambda+\mu}$.
22. P(N = k) = $\left(\frac{5}{6}\right)^{k-2}\frac{1}{6}$, k = 2, 3, ...; E(N) = 7.

23. (a) $1 - e^{-\alpha T}$; (b) $\dfrac{1}{\alpha}(1 - e^{-\alpha T})$.

24. $$g(t) = \left\{ \begin{array}{ll} (\lambda+\mu)\, e^{-(\lambda+\mu)t} & \text{si } t > 0 \\ 0 & \text{en otro caso} \end{array} \right\};$$ media = $1 / (\lambda+\mu)$.

25[2]. $\dfrac{1+n}{2}$.

[2] Solución no "oficial" (pendiente de confirmar).

ESTADÍSTICA I (SISTEMAS Y GESTIÓN)

Ejercicios de Regresión Lineal tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001

A) SISTEMAS

1. *(Septiembre 1995)* El tiempo que emplea un dispositivo en realizar determinado cálculo, es función del tamaño del archivo tratado. En 10 observaciones se han obtenido los siguientes datos:

Tamaño (x)	362	387	254	317	428	231	276	324	441	510
Tiempo (y)	22	25	20	22	28	17	19	23	25	29

a) Calcular la recta de regresión de y sobre x.
b) Calcular la proporción de la varianza de y explicada por el modelo.
Datos auxiliares: $\bar{x}$ = 353; $\bar{y}$ = 23; s_y^2 = 13’2; s_x^2 = 7200’6; s_{xy} = 292’5.

2. *(Febrero 1996)* El tiempo que tarda un dispositivo en imprimir un listado, está relacionado con el nº de observaciones a imprimir. Si en una muestra de 10 observaciones se han obtenido los siguientes datos:

Nº de docum. (X)	24	25	19	25	27	18	20	22	28	32
Tiempo (Y)	44	51	45	44	53	34	39	46	52	62

a) Calcular la recta de regresión que relaciona el tiempo de impresión con el nº de documentos.
b) Calcular el coeficiente de correlación entre las dos variables.
Datos auxiliares: $\overline{X}$ = 24; $\overline{Y}$ = 47; s_x^2 = 17’2; s_y^2 = 55’8; s_{xy} = 28’1.

3. *(Septiembre 1996)* El tiempo de ocupación (Y) de un ordenador parece depender del nº de operaciones (X) que realiza. Observando estas variables durante 10 sesiones se han obtenido los siguientes resultados:

X	9	13	15	20	12	18	16	17	12	18
Y	23	39	44	62	34	52	44	46	23	53

a) Determinar la recta de regresión de Y sobre X.
b) Calcular el grado de ajuste de la recta de regresión a los datos.
Datos auxiliares: $\overline{X}$ = 15; $\overline{Y}$ = 42; s_x^2 = 10’6; s_y^2 = 144; s_{xy} = 37’4.

4. *(Febrero 1997)* En un estudio sobre la influencia de la memoria RAM (x_1) y la velocidad de la CPU (x_2) en los rendimientos (y) de los ordenadores de una empresa, se han obtenido los siguientes parámetros: $\bar{y}$ = 10; $\bar{x}_1$ = 12; $\bar{x}_2$ = 20, s_y = 4; s_1 = 5; s_2 = 10; ρ_{y1} = 0’8; ρ_{y2} = 0’9; ρ_{12} = 0’6.
a) Hallar el plano de regresión de y sobre (x_1, x_2).
b) Calcular el coeficiente de correlación parcial $\rho_{y2.1}$.

5. *(Septiembre 1997)* El tiempo que se tarda en hacer una conexión con una determinada página Web a través de Internet parece estar relacionado con el nº de personas que están accediendo a ella. Se han estudiado los tiempos de acceso en 10 días diferentes, resultando los siguientes datos:

Personas accediendo (*X*)	20	16	18	36	25	44	31	27	62	21
Tiempo (*Y*)	51	33	44	70	50	107	55	46	100	44

a) Determinar la recta de regresión de *Y* sobre *X*.
b) Calcular el grado de ajuste de la recta de regresión a los datos.
Datos auxiliares: $\overline{X} = 30$; $\overline{Y} = 60$; $s_x^2 = 181'2$; $s_y^2 = 555'2$; $s_{xy} = 288'9$.

6. *(Febrero 2000)* Suponga que 25 observaciones del par de variables *(X, Y)* proporcionan los siguientes resultados: $\sum x_i = 75$, $\sum y_i = 50$, $\sum x_i^2 = 625$, $\sum y_i^2 = 228$, $\sum x_i y_i = 30$.
a) Calcule la pendiente de la recta de regresión de *Y* sobre *X*.
b) ¿Qué valor toma la varianza residual?

7. *(Febrero 2000)* Suponga que el coeficiente de correlación calculado a partir de 10 observaciones del par *(X, Y)* es r = 0'60.
a) Si $\bar{y} = 4$ y $\sum_{i=1}^{10} y_i^2 = 250$, ¿cuánto vale la varianza residual para la regresión de *Y* sobre *X*?
b) Si $\sqrt{V_X} = 6$ (desviación típica de la variable *X*), ¿cuál es la expresión de la recta de regresión a la que hace referencia el apartado anterior?

8. *(Septiembre 2000)* Suponga que 13 observaciones del par de variables *(X, Y)* proporcionan los siguientes resultados: $\sum x_i = 169$, $\sum y_i = 26$, $\sum x_i^2 = 2797$, $\sum y_i^2 = 71$, b = - 0'15 (pendiente de la recta de regresión).
a) Calcule el coeficiente de correlación entre *X* e *Y*.
b) Obtenga el error cuadrático medio cometido con la recta de regresión de *Y* sobre *X*.

9. *(Febrero 2001)* Se dispone de 25 observaciones de dos características, *X* e *Y*, de una población. El cálculo de la recta de regresión de *Y* sobre *X* produjo los siguientes resultados: $y = 2'5 + 4'0\ x$.
Sabiendo que $V_x / V_y = 0'04$ (cociente de varianzas), se pide:
a) Calcular el coeficiente de correlación muestral entre ambas variables.
b) ¿Qué porcentaje de la varianza de *Y* representa la varianza residual?

10. *(Febrero 2001)* Con el fin de describir una nube de puntos bidimensional, formada por 20 pares (x_i, y_i), se calcularon algunos resúmenes estadísticos, resultando:
$\bar{x} = 2'80$, V_x (varianza) = 2'56, $\sum_i x_i y_i = 1014$, $\bar{y} = 15'45$, V_y (varianza) = 40'45.

Se pide:
a) Calcular el coeficiente de correlación muestral entre *X* e *Y*.
b) Obtener el error cuadrático medio cometido con la recta de regresión de *Y* sobre *X*. ¿Cuál será el signo de la pendiente de dicha recta? Razone la respuesta.

11. *(Septiembre 2001)* Suponga que los valores de las variables X e Y se asocian como se muestra en la tabla siguiente:

X	2	4	1	5
Y	3	5	1	3

a) Calcule la media y la desviación típica de X.
b) Partiendo de que la pendiente de la recta de regresión de Y sobre X es 0'6, obtenga el coeficiente de correlación muestral.

B) GESTIÓN

12. *(Septiembre 1995)*[1] El tiempo que tarda una impresora en terminar un listado está relacionado con el nº de registros a imprimir. En una muestra de 10 listados se han obtenido los datos siguientes:

Nº de registros (X)	36	38	25	31	42	23	27	32	44	51
Tiempo de impr. (Y)	221	254	203	224	285	173	194	231	262	309

Calcular la recta de regresión que relaciona el tiempo de impresión con el nº de registros y el coeficiente de correlación entre X e Y.
Datos auxiliares: $\overline{X} = 35$, $\overline{Y} = 235$, $\sigma_X^2 = 71$, $\sigma_Y^2 = 1614'4$, $\sigma_{XY} = 325$.

13. *(Febrero 1996)*[1] Tres variables y, x_1, x_2 tienen medias: $\overline{y} = 1'349$, $\overline{x}_1 = 0'567$ y $\overline{x}_2 = 0'542$; desviaciones típicas: $\sigma_y = 0'4753$, $\sigma_1 = 0'2729$, y $\sigma_2 = 0'2112$, y matriz de correlaciones:

$$\begin{pmatrix} 1 & \rho_{y1} & \rho_{y2} \\ \rho_{1y} & 1 & \rho_{12} \\ \rho_{2y} & \rho_{21} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0'9344 & 0'6292 \\ 0'9344 & 1 & 0'5023 \\ 0'6292 & 0'5023 & 1 \end{pmatrix}$$

Hallar:
a) La recta de regresión de x_2 y x_1.
b) El coeficiente de correlación parcial $\rho_{y2.1}$. Comprobar que se verifica $\rho_{y2.1}^2 \le \rho_{y2}^2$.
c) ¿Se cumple siempre $\rho_{y2.1}^2 \le \rho_{y2}^2$? Razonar la respuesta.

14. *(Febrero 1997)*[1] Tres variables y, x_1, x_2 tienen matriz de correlaciones igual a:

$$R = \begin{pmatrix} 1 & \rho_{y1} & \rho_{y2} \\ \rho_{y1} & 1 & 0 \\ \rho_{y2} & 0 & 1 \end{pmatrix}$$

Hallar el coeficiente de correlación múltiple $\rho_{y(12)}^2$. ¿Se cumple $\rho_{y1.2} = \rho_{y1}$? Razonar la respuesta.

15. *(Febrero 1998)*[1] Consideremos tres variables Y, X_1, X_2 con matriz de correlaciones:

$$\begin{pmatrix} 1 & \rho_{y1} & \rho_{y2} \\ \rho_{1y} & 1 & \rho_{12} \\ \rho_{2y} & \rho_{21} & 1 \end{pmatrix}$$

Probar que si X_1 y X_2 están incorreladas, se cumple $\rho_{y2.1}^2 \ge \rho_{y2}^2$.

---

[1] De acuerdo con las instrucciones para el curso 2001/2002 de la página web de la asignatura, <u>los exámenes de cursos anteriores al 1998/1999 NO SON ORIENTATIVOS</u> para preparar la materia del curso actual.

16. *(Febrero 1999)* Tres variables y, x_1, x_2 tienen matriz de correlaciones:

$$R = \begin{pmatrix} 1 & \rho_{y1} & \rho_{y2} \\ \rho_{y1} & 1 & -1 \\ \rho_{y2} & -1 & 1 \end{pmatrix}$$

Hallar el coeficiente de correlación múltiple $\rho^2_{y(12)}$.

Soluciones

1. (a) y = 8’6682 + 0’0406 x. (b) 0’9001.
2. (a) Y = 7’791 + 1’634 X. (b) 0’907.
3. (a) Y = -10’93 + 3’53 X. (b) 91’64%.
4. (a) y = 0’8500 + 0’3250 x_1 + 0’2625 x_2. (b) 0’8750.
5. (a) Y = 12’1680 + 1’5944 X. (b) 0’8296.
6. (a) $b = \frac{-3}{10}$. (b) $\frac{92}{25}$.
7. (a) 5’76. (b) y = 0’3 x (en desviaciones respecto a las medias).
8. (a) (-0’8429). (b) 0’4231.
9. (a) 0’8. (b) 36%.
10. (a) 0’7311. (b) 18’8292. Positiva.
11. (a) 1’5811. (b) 0’6708.
12. Y = 74’49 + 4’58 X. ρ_{XY} = 0’9599.
13. (a) x_2 – 0’542 = 0’3887 (x_1 – 0’567) (b) $\rho_{y2.1} = 0'519$. Se cumple la desigualdad indicada en el enunciado.

    (c) En general, no se cumple dicha desigualdad.
14. $\rho^2_{y(12)} = \rho^2_{y1} + \rho^2_{y2}$. $\rho_{y1.2} = \rho_{y1}$ si $\rho_{y2} = 0$.
15. $\rho^2_{y2.1} = \dfrac{\rho^2_{y2}}{1-\rho^2_{y1}} \geq \rho^2_{y2}$.
16. $\rho^2_{y(12)} = \rho^2_{y1} = \rho^2_{y2}$.

ESTADÍSTICA I (SISTEMAS)

Ejercicios de Regresión y Análisis de la Varianza tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001

1. *(Febrero 2000)* Construir la tabla de análisis de la varianza para los datos siguientes (4 grupos con 3 observaciones por grupo):

Observaciones↓ / Grupos→	1	2	3	4
	110	111	113	118
	109	116	108	123
	105	109	109	125
Suma	324	336	330	366
Media	108	112	110	122

¿Son significativas las diferencias observadas en la variable respuesta?

Datos auxiliares: $F_{3,\,8;\,0'05} = 4'0662$, $\sum_{i=1}^{4}\sum_{j=1}^{3}\left(x_{ij} - \bar{x}_{..}\right)^2 = 428$.

2. *(Febrero 2000)* Completar la siguiente tabla de análisis de la varianza que corresponde a un experimento unifactorial con *tres* niveles – grupos – y unos tamaños muestrales de *seis*, *cuatro* y *cinco* unidades, respectivamente:

Fuente de variación→	Entre grupos	Dentro de los grupos
Suma de cuadrados	66’93	□
G. l.	□	□
Cuadrado medio	□	8’83
Estadístico		□

¿Qué se puede concluir a partir de estos resultados?

Datos auxiliares: $F_{2,\,12;\,0'05} = 3'8853$, $F_{3,\,15;\,0'01} = 5'4170$, $F_{3,\,10;\,0'05} = 3'7083$.

3. *(Septiembre 2000)* Un experimento orientado a detectar diferencias en el funcionamiento (medido en términos de una variable X) de 4 equipos informáticos proporcionó los siguientes datos:

Equipos	1	2	3	4
	76’42	80’41	74’20	86’20
	78’62	82’26	72’68	86’04
	80’40	81’15	78’84	84’36
	78’20	79’20	80’32	80’68
Media ($\overline{x_{i.}}$)	78’41	80’75	76’51	84’32
Desv. ($\overline{x_{i.}} - \overline{x_{..}}$)	(-1’59)	0’75	(-3’49)	4’32
$(\overline{x_{i.}} - \overline{x_{..}})^2$	2’53	0’56	12’18	18’66

Se pide:
a) Construir la tabla de análisis de la varianza.
b) ¿Qué se puede concluir respecto al objetivo planteado?

Datos auxiliares: $F_{3,\ 12;\ 0'05} = 3'4903$, $\sum_{i=1}^{4}\sum_{j=1}^{4}\left(x_{ij} - \overline{x_{..}}\right)^2 = 208'42$.

4. *(Febrero 2001)* Para estudiar la asociación entre dos variables X e Y, se hace uso de un modelo de regresión lineal. La estimación de los parámetros de dicho modelo, a partir de una muestra de 15 pares (x_i, y_i), proporcionó los siguientes resultados:

Modelo $y = \alpha + \beta \cdot x$; $\hat{\alpha} = 7'699$; $\hat{\beta} = (-0'060)$

Además: $\bar{x} = 43'533$, v_x (varianza) = 434'94, $cov_{x,\ y} = (-26'32)$,
$\bar{y} = 5'067$, v_y (varianza) = 3'39

¿Se puede confirmar estadísticamente, al nivel $\alpha = 0'05$, que el parámetro β es negativo?

5. *(Septiembre 2001)* Se pretende hacer un *análisis de la varianza* a partir de los siguientes datos:

				Media
Grupo 1	3	4	3	3'333
Grupo 2	1	3	2	2'000
Grupo 3	8	6	5	6'333

a) Calcule la suma de cuadrados *entre grupos*.
b) Obtenga la suma de cuadrados *dentro de los grupos*.
c) Calcule e interprete el valor del estadístico F en este caso.
Datos auxiliares: La variabilidad total es 36'889. $F_{2,\ 6;\ 0'05} = 5'1433$, $F_{3,\ 3;\ 0'05} = 9'2766$, $F_{3,\ 9;\ 0'05} = 3'8625$.

Soluciones

1.

Fuente de la variación	*Suma de cuadrados*	*Gr. Lib.*	*Media cuadrados*	*F*
Entre grupos	348	3	116	
Dentro grupos	80	8	10	11’6

Las diferencias observadas en la variable respuesta son significativas, para un $\alpha = 0'05$.

2.

Fuente de variación→	Entre grupos	Dentro de los grupos
Suma de cuadrados	66’93	105’96
G. L.	2	12
Cuadrado medio	33’465	8’83
Estadístico		3’7899

Las diferencias observadas en la variable respuesta no son significativas, para un $\alpha = 0'05$.

3.

Fuente de la variación	*Suma de cuadrados*	*Gr. Lib.*	*Media cuadrados*	*F*
Entre grupos	135’73	3	45’24	
Dentro grupos	72’69	12	6’06	7’47

Las diferencias observadas en el funcionamiento de los equipos son significativas, para un $\alpha = 0'05$.

4. Se puede confirmar que β es negativo.

5. (a) 29’552. (b) 7’337. (c) $F = 12'083 \Rightarrow$ Las diferencias son significativas.

ESTADÍSTICA I (SISTEMAS)

Ejercicios de Estimación (Puntual y por Intervalos de Confianza) tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001

1. *(Febrero 1996)* El valor de la medida X es fundamental para el buen funcionamiento de una determinada componente de tipo electrónico. Por esta razón el proceso de fabricación de dicha componente debe ser controlado al máximo. Se pide:

a) Estimar la media poblacional mediante un intervalo con un nivel de confianza del 99%, sabiendo que en una muestra aleatoria de 25 artículos la media de X fue 3 y la desviación típica 0’25.

b) ¿Cuántas observaciones serían necesarias si se pretende que la amplitud de dicho intervalo sea inferior a 0’1?

Datos auxiliares: $t_{26;\,0'005} = 2'779$; $t_{25;\,0'005} = 2'787$; $t_{24;\,0'005} = 2'797$.

2. *(Febrero 1998)* El tiempo de vida de una determinada componente electrónica sigue una distribución exponencial dada por $f_\lambda(x) = \frac{1}{\lambda} e^{-x/\lambda}$; $\lambda \in \mathbb{R}^+$, $x > 0$.

En una muestra de 10 de estas componentes se obtuvieron las siguientes duraciones, en miles de horas: 52, 50, 10, 25, 40, 45, 56, 20, 35, 27.

Determinar la estimación máximo verosímil del parámetro λ.

3. *(Septiembre 1998)* De una determinada especie de ratas africanas sólo se sabe que su peso se distribuye según una ley normal de varianza 9 grs^2. Capturadas 8 de estas ratas se obtuvieron los siguientes pesos (en gramos): 125, 124’5, 127, 126’5, 123, 120, 124, 127’5.

a) Determinar un intervalo de confianza, al 95%, para el peso medio de dicha especie de ratas.

b) ¿Cuál sería el tamaño muestral necesario para que la estimación del peso medio difiera del peso medio real en menos de una unidad, con una probabilidad de al menos 0’95?

Datos auxiliares: Si F es la función de distribución de una N(0, 1), F(1’96) = 0’975, F(-2’575) = 0’005.

4. *(Febrero 1999)* El tiempo de vida de ciertos aparatos sigue una distribución *normal*. Extraída una muestra de 10 de estos aparatos, se observaron los siguientes tiempos de vida (medidos en miles de horas): 10, 15, 8, 7, 11, 6, 4, 9, 12, 8.

Obtenga:

a) Un estimador insesgado del tiempo de vida medio.

b) Un intervalo de confianza, al 95%, para el tiempo de vida medio. ¿De qué modo interpreta el nivel de confianza del intervalo anterior?

Datos auxiliares: $t_{10;\,0'025} = 2'228$, $z_{0'025} = 1'96$, $t_{9;\,0'05} = 1'833$, $t_{9;\,0'025} = 2'262$.

Soluciones

1. (a) (2’86, 3’14). (b) Al menos 196.
2. 36.
3. (a) (122’61, 126’77). (b) n = 35.
4. (a) 9. (b) (1’8469, 16’1531).

ESTADÍSTICA I (SISTEMAS)

Ejercicios de Contraste de Hipótesis tomados de exámenes reales, desde febrero de 1995 hasta septiembre de 2001[1]

1. *(Febrero 1995)* Para contrastar dos métodos de enseñanza se consideraron dos grupos de 10 estudiantes cada uno. Finalizado el período de prueba se realizó una evaluación, obteniéndose los siguientes resultados:

Método	n	$\overline{X}_i$	$\sum\left(X_i - \overline{X}_i\right)^2$
1	10	82	6.060
2	10	79	6.900

¿Es significativa la diferencia entre medias? (*Suponga varianzas iguales*)
Datos auxiliares: $t_{18,\,0'1} = 1'330$; $t_{18,\,0'05} = 1'734$; $t_{18,\,0'025} = 2'101$.

2. *(Febrero 1995)* Se ha establecido recientemente un nuevo programa de seguridad para las empresas de un determinado sector industrial. Con el fin de valorar dicho plan se ha registrado, para 10 plantas industriales de características similares, el nº de horas perdidas a causa de accidentes laborales, resultando:

Planta	Antes	Después	Dif. (D_i)
1	30	23	7
2	18	21	(-3)
3	24	22	2
4	32	28	4
5	16	14	2
6	15	15	0
7	23	24	(-1)
8	25	21	4
9	28	23	5
10	18	25	(-7)

Determinar, con un nivel de significación $\alpha = 0'05$, si el programa es efectivo.
Datos auxiliares: $\Sigma D_i = 13$, $\Sigma D_i^2 = 173$; $t_{9,\,0'1} = 1'383$; $t_{9,\,0'05} = 1'833$; $t_{9,0'025} = 2'262$.

3. *(Septiembre 1995)* La tabla adjunta contiene una muestra de los valores de una variable, relacionada con la velocidad de proceso de una clase de equipos informáticos, antes (*X*) y después (*Y*) de la incorporación de una nueva componente electrónica. ¿Es significativa la diferencia observada?

X	6'98	7'08	8'34	5'30	6'26	6'77	7'03	5'56	5'97	6'64	7'03	7'69
Y	6'95	6'94	7'17	5'15	6'28	6'81	6'59	5'34	5'98	6'51	6'84	6'99
D = X - Y	0'03	0'14	1'17	0'15	(-0'02)	(-0'04)	0'44	0'22	(-0'01)	0'13	0'19	0'70

Datos auxiliares: $t_{11,\,0'05} = 1'796$; $t_{12,\,0'05} = 1'782$; $t_{11,\,0'025} = 2'201$; $t_{12,\,0'025} = 2'179$.
$\sum_i D_i = 3'10$; $\sum_i D_i^2 = 2'199$; $\left(\sum_i D_i\right)^2/12 = 0'801$.

[1] En la resolución de todos estos ejercicios, se presupone que las poblaciones son *normales*.

4. *(Febrero 1996)* Con el fin de comparar la velocidad de dos equipos informáticos cuando se realiza cierta clase de trabajos, se consideraron 9 tareas y se midió el tiempo empleado por cada uno de dichos equipos para su ejecución. Los resultados obtenidos se muestran en la siguiente tabla:

Tarea	1	2	3	4	5	6	7	8	9
Eq. 1	3	2	1	4	3	3	1	1	2
Eq. 2	1	2	2	2	3	2	1	2	3

¿Es significativa la diferencia observada?
Datos auxiliares: $t_{10,\ 0'025} = 2'228$; $t_{9,\ 0'025} = 2'262$; $t_{8,\ 0'025} = 2'306$.

5. *(Septiembre 1996)* Un fabricante de componentes electrónicas para ordenadores hace publicidad de sus productos asegurando que sólo una cantidad inferior al 2% presenta defectos de origen. Una empresa dedicada al montaje de equipos informáticos, impresionada por este mensaje publicitario, ha efectuado una compra masiva de dichas componentes. Para contrastar la afirmación del fabricante, se decide analizar una muestra aleatoria de tamaño 300 y resulta que 10 elementos son defectuosos. ¿Qué se puede concluir a partir de este dato?
Datos auxiliares: F(2'33) = 0'9901; F(1'65) = 0'9505; F(1'96) = 0'9750, siendo F la función de distribución *normal estándar*.

6. *(Febrero 1997)* En una muestra aleatoria de 400 individuos se encontraron 12 cumpliendo una condición determinada. ¿Hay una evidencia suficiente para concluir que la proporción de individuos de esta clase es superior al 2%? (Utilice $\alpha = 0'05$).
Nota: no dan datos auxiliares.

7. *(Febrero 1997)* La tabla siguiente resume los datos obtenidos en dos muestras aleatorias independientes, procedentes de dos poblaciones *normales*.

Muestra	Tamaño	$\overline{x}_i$	$\sum\left(x_{ij} - \overline{x}_i\right)^2$
1	15	35'2	35
2	12	34'0	25

¿Se puede concluir, al nivel de significación 5%, que la media de la 1ª población es mayor que la media de la 2ª?
Datos auxiliares: $t_{25,\ 0'05} = 1'708$; $t_{27,\ 0'025} = 2'052$; $t_{29,\ 0'025} = 2'045$; $t_{27,\ 0'05} = 1'703$.

8. *(Septiembre 1997)* Se supone que una componente electrónica, que ha sido utilizada con mucha frecuencia durante los últimos años, presenta defectos de fabricación en el 5% de los casos. Un equipo de investigadores de una planta que utiliza dichos componentes en sus montajes ha revisado una muestra aleatoria de 100 elementos y ha encontrado fallos en 8 de los 100 casos considerados. ¿Se puede concluir que la componente falla más de lo supuesto?
Datos auxiliares: $z_{0'05} = 1'645$.

9. *(Febrero 1998)* El departamento de control de calidad de una empresa dedicada al montaje de PC's afirma que el tiempo medio de vida de sus electroventiladores no es inferior a 3 años. Los archivos que guarda dicho departamento ofrecen los resultados que se observaron sobre una muestra de 45 electroventiladores para la que se obtuvo: $\overline{X}$ = 2'83 y cuasidesviación típica s = 0'596. ¿Avala la información contenida en los archivos la afirmación del departamento? Tómese un nivel de significación de α = 0'05.

Datos auxiliares: $z_{0'95}$ = (-1'645), $z_{0'005}$ = 2'575, (-0'17) $\sqrt{45}$ / 0'596 = (-1'91).

10. *(Febrero 1999)* Una compañía aérea afirma que la duración de sus vuelos entre las ciudades A y B se distribuye según una ley *Normal* de media no superior a 1'2 horas. Sin embargo, la competencia afirma que esto no es cierto ya que, después de tomar 9 vuelos al azar, observó una duración media de 1'78 horas y una cuasivarianza de 0'36. ¿Avalan estos datos la afirmación de la competencia? (Utilice α = 0'05).

Datos auxiliares: $t_{8,\,0'05}$ = 1'86; $z_{0'05}$ = 1'645, $z_{0'025}$ = 1'96.

11. *(Febrero 1999)* Un vendedor afirma que el tiempo medio de vida de las bombillas que oferta a sus clientes no es inferior a 10 meses. Extraída una muestra aleatoria de 64 bombillas, se obtuvo un tiempo de vida medio de 9'8 meses y una cuasivarianza muestral de 1. ¿Avala la muestra, al nivel de significación α = 0'025, la afirmación hecha por el vendedor?

Datos auxiliares: Si F es la función de distribución de la N (0, 1), F(1'96) = 0'975 y F(1'645) = 0'95.

12. *(Septiembre 1999)* Los beneficios diarios, en millones de pts., obtenidos por dos entidades bancarias siguen distribuciones *normales* con la misma varianza. Del último ejercicio económico se tomaron al azar dos muestras: de 11 días para la 1ª entidad, obteniéndose un beneficio medio de 100 y una cuasivarianza de 8'4 y de 17 días para la 2ª, que dio un beneficio medio de 96 y una cuasivarianza de 15. ¿Existen diferencias significativas entre los beneficios medios diarios de ambas entidades en el último ejercicio económico? Tómese α = 0'02.

Datos auxiliares: $t_{26,\,0'01}$ = 2'479; $t_{28,\,0'01}$ = 2'467; $z_{0'01}$ = 2'33.

13. *(Febrero 2000)* El nº de pulsaciones por minuto de dos secretarias A y B sigue una distribución *normal* de varianza 10 para la secretaria A y 12'5 para la B. En un examen de rapidez mecanográfica se sometió a ambas a 10 pruebas independientes, obteniéndose un nº medio de 198 pulsaciones por minuto para A y 195 para B.

a) A la vista de los resultados de las pruebas, ¿existen diferencias entre la rapidez mecanográfica de ambas secretarias? Tómese α = 0'05.
b) Si se toma α = 0'01, ¿cuáles son las conclusiones obtenidas? Relacione estas conclusiones con las del apartado anterior.

Datos auxiliares: Si F es la función de distribución de la *normal* estándar, F(-1'96) = 0'025, F(-1'645) = 0'05, F(-2'58) = 0'005.

14. *(Septiembre 2000)* A fin de comparar las cantidades de tabaco consumidas por los habitantes de una comarca procedentes de zona rural y los habitantes de la misma comarca procedentes de zona urbana, se tomaron 10 y 12 fumadores de cada zona obteniéndose – en número de cigarros por día – una media de 16 y cuasivarianza de 4 para el primer grupo, y una media de 11 y cuasivarianza de 5 para el segundo.

a) Plantee un contraste de hipótesis que permita comparar el consumo medio en ambos grupos; tómese α = 0'01.

b) ¿Qué se puede decir del p-valor de la muestra en relación al nivel de significación considerado?

Datos auxiliares: $t_{22,\ 0'01}$ = 2'508; $t_{20,\ 0'05}$ = 2'845; $t_{20,\ 0'01}$ = 2'528; $t_{22,\ 0'005}$ = 2'819. Suponga normalidad e igualdad de varianzas en ambos grupos.

15. *(Febrero 2000)* Nos dicen que un programa de ordenador genera observaciones de una distribución *normal estándar*, es decir, N (0, 1). En una muestra aleatoria de 100 observaciones producidas mediante dicho programa se obtienen los siguientes resultados:

6 observaciones menores que (-2).
20 entre (-2) y (-1).
30 entre (-1) y 0.
25 entre 0 y 1.
15 entre 1 y 2.
4 mayores que 2.

Sabiendo que una distribución N (0, 1) asigna las probabilidades:

Pr {menores que (-2)} = 0'02.
Pr {entre (-2) y (-1)} = 0'14.
Pr {entre (-1) y 0} = 0'34,

¿qué regla utilizaría para decidir, al nivel α = 0'01, si el programa funciona correctamente o no? *(No es necesario completar los cálculos, pero debe exponer claramente las operaciones que habría que hacer con los datos proporcionados en el enunciado).*

Datos auxiliares: $\chi^2_{3,0'01}$ = 11'345; $\chi^2_{5,\ 0'01}$ = 15'086; $\chi^2_{6,0'01}$ = 16'812.

16. *(Febrero 2001)* El departamento de control de calidad de una empresa afirma que el porcentaje de artículos defectuosos que salen de su cadena de montaje no supera el 5%. En un experimento posterior se obtuvo que de un total de 1.000 artículos había 70 defectuosos. ¿Avala dicho experimento la afirmación hecha por el departamento de control de calidad? Tómese un nivel de significación α = 0'05.

Datos auxiliares: Si F es la *función de distribución* de una N (0, 1), F(-1'645) = 0'05; F(1'96) = 0'975 y F(0'05) = 0'52.

17. *(Febrero 2001)* El tiempo de funcionamiento de ciertos equipos informáticos antes de que comiencen a presentar fallos sistemáticos es una variable aleatoria con distribución *normal*. Un comprador decide adquirir estos equipos si su tiempo medio de funcionamiento sin fallos es de al menos 5 años. A fin de tomar una decisión, analiza la experiencia de compradores anteriores, observando para una muestra de 9 equipos un tiempo medio de funcionamiento sin fallos de 4 años y 6 meses y una cuasivarianza de 1 año. ¿Qué decisión adoptaría en base a la experiencia de estos compradores?

Datos auxiliares: $z_{0'01}$ = 2'33; $t_{9,\ 0'01}$ = 2'821; $t_{8,\ 0'005}$ = 3'355; $t_{8,\ 0'01}$ = 2'896.

18. *(Septiembre 2001)* Una empresa dedicada al montaje de equipos informáticos distribuye entre sus clientes dos modelos: A y B. Cierto estudio de mercado concluyó que al menos tres de cada cuatro individuos del grupo de clientes potenciales preferían el modelo A. Un estudio paralelo del departamento de análisis de datos de la empresa mostró que, de 1.000 individuos tomados al azar entre los clientes del último año, 700 compraron el modelo A. Si el comportamiento de los antiguos clientes se puede extrapolar al del grupo de clientes potenciales, ¿avala este último estudio las conclusiones del estudio de mercado realizado? Tómese un nivel de significación α = 0'01.

Datos auxiliares: Si F es la *función de distribución* de una N (0, 1), F(-2'58) = 0'005; F(-2'33) = 0'01 y F(1'645) = 0'95.

Soluciones

1. La diferencia no es significativa, para un nivel $\alpha = 0'05$.
2. El programa no es efectivo.
3. La diferencia observada es significativa, para un nivel $\alpha = 0'05$.
4. La diferencia observada es significativa, para un nivel $\alpha = 0'05$.
5. Realmente parece difícil que $p = 0'02$ como asegura la empresa, pero a un nivel de significación del 5% podría ser aceptada.
6. No hay una evidencia suficiente para concluir que la proporción de individuos de esta clase es superior al 2%.
7. No se puede concluir que la media de la 1ª población sea mayor.
8. No se puede concluir, a un nivel de significación del 5%, que la componente falle más de lo supuesto.
9. La información contenida en los archivos avala la afirmación del departamento.
10. Los datos avalan la afirmación de la competencia.
11. La muestra no avala la afirmación hecha por el vendedor.
12. Existen diferencias significativas entre los beneficios medios diarios de ambas entidades.
13. a) Existe una diferencia significativa entre la rapidez mecanográfica de ambas secretarias.
    b) No hay una diferencia significativa entre la rapidez de las dos secretarias.
    Se llega a conclusiones opuestas en ambos apartados, en función de cuál sea el nivel de significación elegido.
14. a) Puede afirmarse que la proporción de fumadores es superior en la zona rural.
    b) El p-valor es prácticamente igual al nivel de significación (0'01), y al ser tan pequeño (próximo a cero) nos permite rechazar con gran seguridad H_0.
15. $\lambda = 15'50 > \chi^2_{5,\, 0'01}$. El programa no funciona correctamente.
16. No la avala.
17. Se decidiría no comprar los equipos informáticos a ese proveedor.
18. No lo avala.