Hoja de refuerzo 1

Probabilidad y Estadística

Ejercicio 1

Queremos analizar si existe relación lineal entre la temperatura (X) en la que se produce una reacción química y los porcentajes de producto obtenidos en la misma (Y). Se han analizado las temperaturas (ºC) y las cantidades de producto (%) de 5 reacciones químicas y los resultados son los siguientes:

x_i	150	170	200	250	300
y_i	77.4	78	84.5	89.2	95

a) ¿Puedes decir si que existe una relación lineal entre la temperatura y el porcentaje de producto obtenido? Interpretar el valor de r.

b) Predecir la temperatura en la que se produjo una reacción que dio como resultado un 85% de producto

c) Se desechan el 15% de las reacciones que consiguen el menor porcentaje de producto y el 5% de las que llegan al máximo porcentaje de producto, ¿entre qué dos valores una reacción se considera válida?

Sol: a) $r = 0.99$, b) $215.45°C$, c) $P_{15} = 77.4$ y $P_{95} = 95$.

Ejercicio 2

La probabilidad de comprar un periódico es $p(P) = 0,3$, la de una revista es $p(R) = 0,2$ y la de comprar ambos $p(P \cap R) = 0,08$. Se pide calcular las probabilidades de:

a) Comprar un periódico o una revista.

b) Comprar un periódico y no una revista.

c) Comprar una revista y no un periódico.

d) No comprar un periódico y comprar una revista.

e) No comprar un periódico o no comprar una revista.

f) No comprar un periódico y no comprar una revista.

Sol: a) 0.42, b) 0.22, c) 0.12, d) 0.12, e) 0.92, f) 0.58.

Ejercicio 3

Se trató de ajustar un modelo de regresión lineal simple para analizar la relación entre las variables

X: producción de trigo en Tm

Y: precio del kilogramo de harina (en euros)

Disponemos de los siguientes datos relativos a los últimos 5 años:

$\bar{x} = 28$; $\bar{y} = 0,414$

$\sum_{i=1}^{n} x_i^2 = 3958$; $\sum_{i=1}^{n} y_i^2 = 0,86$;

$\sum_{i=1}^{n} x_i y_i = 57,68$

a) ¿Qué puedes decir de la interdependencia entre las variables?

b) Predecir la producción de trigo un año en el que el precio del harina fue de 0,47 euros y dar una medida de la fiabilidad de dicha predicción

c) Calcular el coeficiente de determinación del modelo y dibujar la nube de puntos y la recta de regresión de x sobre y en un plano cartesiano lo más aproximadamente posible. A raiz de estos resultados, ¿considera que el modelo de predicción es bueno o malo? Razonar la respuesta.

d) Calcular la pendiente de la recta de regresión de y sobre x e interpretarla en el contexto del problema.

e) Si nos facilitan además las producciones en Tm de los 5 años de estudio, que son: 30, 28, 32, 25, 25. ¿Entre qué dos valores estará el 50% central de la distribución de producciones?

Sol: a) $r = -0.82$, b) 22.57, c) $R^2 = 0.6724$, d) m=-0.0074, e) $Q_1 = 25$, $Q_3 = 30$.

Ejercicio 4

El 1% de la población de un determinado lugar padece una enfermedad. Para detectar esta enfermedad se realiza una prueba diagnóstica. Esta prueba da positiva en el 97% de los pacientes que padecen la enfermedad; en el 98% de los individuos que no la padecen da negativa. Si elegimos al azar un individuo de esa población:

a) ¿Cuál es la probabilidad de que un individuo dé positivo y padezca la enfermedad?

b) Si sabemos que ha dado positiva, ¿cuál es la probabilidad de que padezca la enfermedad?

Sol: a) 0.0097, b) 0.32.

Ejercicio 5

Supongamos que el 90% de los vuelos despegan puntuales; el 80% de los vuelos llegan puntuales; y el 75% de los vuelos despegan y aterrizan con puntualidad.

a) ¿Cuál es la probabilidad de que un vuelo que despegó puntual, llegue también puntual?

b) ¿Y de que un vuelo que ha aterrizado puntual, también despegara puntual?

c) ¿Son los sucesos independientes?

Sol: a) 0.83, b) 0.93, c) No.

Ejercicio 6

El 42.5% de las empresas españolas tienen menos de 10 empleados. Un 31.86 tienen entre 10 y 49 empleados. Y un 16.64% tienen entre 50 y 249 empleados. Por otra parte, el porcentaje de empresas que realizaron el año pasado ventas por comercio electrónico fue del 14% para las de menos de 10 empleados, un 18% para las que tienen entre 10 y 49 empleados, un 29% para las que tienen entre 50 y 249 y, por último, un 40% para las que tienen más de 250 empleados. Se pide:

a) Calcular el porcentaje de empresas que realizaron ventas por comercio electrónico.

b) Sabiendo que una empresa realizó ventas por comercio electrónico, ¿cuál es la probabilidad de que tenga más de 250 empleados?

c) Y si no realizó ventas por correo electrónico, ¿cuál es la probabilidad de que tenga más de 250 empleados?

d) Calcular la probabilidad de que una empresa, elegida al azar, no realizara ventas por comercio electrónico y tenga 10 o más empleados.

Sol: a) 20.11%, b) 0.18, c) 0.067, d) 0.43.

Ejercicio 7

Un transbordador puede llevar tanto automóviles como autobuses en un recorrido por una vía fluvial. Cada viaje cuesta al propietario 10 euros de base. La tarifa por automóvil es de 3 euros y la tarifa por autobús es de 8 euros. Cada día solo realiza un viaje.

Si X e Y representan, respectivamente, el número de automóviles y el número de autobuses que se transportan en un viaje, entonces dada la tabla de probabilidades conjuntas:

Y↓X→	0	1	2
0	0.01	0.01	0.03
1	0.03	0.08	0.07
2	0.03	0.06	0.06
3	0.07	0.07	0.13
4	0.12	0.04	0.03
5	0.08	0.06	0.02

Se pide:

a) Obtener la función de cuantía marginal de la variable Y. ¿Son las variables X e Y independientes?

b) Si un día el transbordador lleva más de 3 autobuses, ¿cuál es la probabilidad de que no lleve automóviles?

c) ¿Qué porcentaje de días el transbordador lleva más de 4 vehículos en total?

d) ¿Qué porcentaje de días el transbordador lleva más automóviles que autobuses?

e) Calcular la ganancia esperada de cada viaje.

Sol: a) no son independientes, b) 0.57 ,c) 36%, d) 11%, e) 15.8 euros.

Ejercicio 8

En una determinada compañía solo existen dos métodos de comunicación entre empleados, email y chat (E y C), y solo dos dispositivos diferentes para enviar las comunicaciones, ordenador y teléfono (O y T). Por norma, un empleado puede usar indistintamente ambos dispositivos durante las comunicaciones, sin embargo solo un método en cada comunicación. Sean los sucesos y sus probabilidades los siguientes:

$E =$ Un empleado usa el email como método de comunicación, $P(E) = 0.6$.

$C =$ Un empleado usa el chat como método de comunicación, $P(C) = 0.4$.

$O =$ Un empleado usa el ordenador como dispositivo de comunicación, $P(O) = 0.7$.

$T =$ Un empleado usa el teléfono como dispositivo de comunicación.

$X =$ Se produce un corte en durante la comunicación.

Se sabe que, la probabilidad de que un empleado cualquiera utilice ambos dispositivos simultaneamente es 0.3. Y, por otra parte, se sabe que la probabilidad de que haya un corte en la comunicación es de 0.1 para los usuarios que usan el email, de 0.15 para los usuarios de chat, y de 0.05 para los que usan el ordenador.

Se pide:

a) Calcular la probabilidad de que haya un corte en las comunicaciones.

b) Calcular la probabilidad de que un empleado use única y exclusivamente el teléfono como dispositivo de comunicación.

c) Se sabe que a un empleado se le ha cortado la comunicación, ¿cuál es la probabilidad de que se haya realizado a través del ordenador?

d) Se sabe que un empleado no tiene ordenador, ¿cuál es la probabilidad de que se le corte la comunicación?

Sol: a) 0.12, b) 0.3, c) 0.29, d) 0.28.

Ejercicio 9

Se ha observado, durante un mes determinado, el gasto en electricidad y el ingreso total en 6 familias. Los resultados, expresados en miles de euros (ingresos) y en cientos de euros (gasto electricidad), son:

.	Ingresos	Gasto eléctrico
1a familia	4	0,2
2a familia	6	0,3
3a familia	8	0,5
4a familia	10	0,9
5a familia	12	1,0
6a familia	20	1,9

a) Calcular la recta de regresión de Y sobre X y la recta de regresión de X sobre Y.

b) Predecir el gasto en electricidad de una familia que tiene unos ingresos de 15.000 euros.

c) ¿Qué ingresos puedes estimar en una familia que ha pagado 135 euros de factura eléctrica?

d) ¿Cuál es el grado de fiabilidad de cada una de esas predicciones?

Sol: a) $y = -0.30 + 0.11x$, $x = 2.82 + 8.98y$, b) 1.35, c) 14943€, d) $r = 0.99$, predicciones muy fiables.

Ejercicio 10

El tiempo de vida de una lámpara especial sigue una distribución exponencial con media 100 hs.

a) ¿Cuál es la probabilidad de que una lámpara dure por lo menos 30 horas?

b) Si una lámpara ya lleva 50 horas de uso, ¿cuál es la probabilidad de que dure más de 80 horas en total?

c) Se seleccionan cinco lámparas, ¿Cuál es el número esperado de lámparas que duran por lo menos 30 hs (considerando las 5)?

Sol: a) 0.7408 , b) 0.7408, c) 3.704.

Ejercicio 11

La probabilidad de que el alumno que está leyendo este examen no haya asistido a clase durante el curso es de 2/3. Sin embargo, el alumno tiene que hacer este examen si quiere aprobar el curso. Este examen es difícil y la probabilidad de aprobarlo habiendo ido a clase es la misma que la de suspenderlo habiendo ido a clase. Pero tiene solamente un 0.25 de probabilidad de aprobarlo si no ha asistido antes a clase durante el curso.

Cuando el profesor corrige el examen, se encuentra con que el alumno ha suspendido, ¿cuál es la probabilidad de que el alumno haya asistido a clase?

Sol: 0.25.

Ejercicio 12

La relación que existe entre las variables X e Y se muestra en la siguiente tabla:

X, Y	1-16	16-31	31-46	46-60
10-19	1	1	1	0

X, Y	1-16	16-31	31-46	46-60
19-28	0	8	3	0
28-37	0	3	7	1
37-45	0	2	3	0

Se pide calcular la correlación y las rectas de regresión.

Sol: $r = 0.375$, $y = 0.44x - 18.16$, $x = 0.31y + 10.12$.

Ejercicio 13

En una urna se encuentran 5 pelotas numeradas el uno al cinco. Se toman dos pelotas al azar, sin reemplazo, y se anotan sus números. Determinar la distribución de probabilidad de:

a) El mayor de los números seleccionados.

b) La suma de los números seleccionados.

Sol: a) $P(X = k) = \frac{1}{10}, \frac{1}{5}, \frac{3}{10}, \frac{2}{5}$, para $k = 2, 3, 4, 5$, respectivamente, b) $P(X = k) = \frac{1}{10}, \frac{1}{10}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{10}, \frac{1}{10}$, para $k = 3, 4, 5, 6, 7, 8, 9$, respectivamente.

Ejercicio 14

Sea la siguiente distribución conjunta de probabilidad:

X, Y	-1	0	1
-1	0.2	0	0.2
0	0	0.2	0
1	0.2	0	0.2

Se pide determinar si son independientes.

Sol: Son dependientes.

Ejercicio 15

Una compañía informática clasifica el riesgo de no cumplir con los plazos de desarrollo en un proyecto en tres tipos: A, B y C, con probabilidades 0.35, 0.40 y 0.25, respectivamente. Se sabe que la probabilidad de que los desarrolladores la empresa contratadora se queje es de 0.4 en los proyectos de tipo A, 0.2 en los de tipo B y 0.3 en los de tipo C. Se pide:

a) ¿Cuál es la probabilidad de que la empresa contratadora se queje?

b) Si la compañía informática no tiene problemas con la empresa contratadora, ¿cuál es la probabilidad de que exista un riesgo de tipo B?

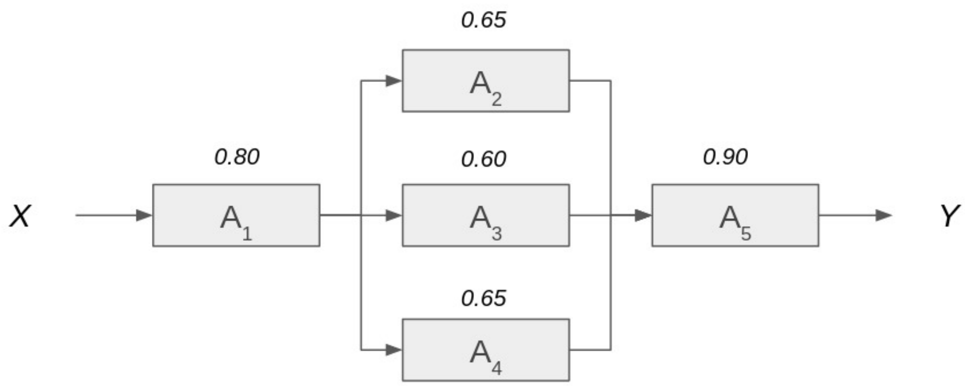
Sol: a) 0.295, b) 0.4539.

Ejercicio 16

Se considera un sistema eléctrico integrado como el que se muestra en el diagrama. Las probabilidades de que los componentes A_1, A_2, A_3, A_4 y A_5 funcionen correctamente se muestran también en el diagrama. Para que el sistema función completamente, debe pasar del nodo X al nodo Y. Se pide:

a) Calcular la probabilidad de que el sistema eléctrico funciona.

b) Calcular la probabilidad de que el sistema eléctrico funciona con, al menos, cuatro componentes.



Sol: a) 0.68, b) 0.5.

Ejercicio 17

En una agencia de viajes que cuenta con dos operadores se consideran las variables aleatorias

X=“número de paquetes vendidos al día por el operador A”

Y=“número de paquetes vendidos al día por el operador B”

En la tabla siguiente se muestran las correspondientes probabilidades conjuntas

X/Y	0	1	2
0	0,15	0,15	0,10
1	0,05	0,20	0,05
2	0,10	0,05	0,15

a) Obtener la función de cuantía marginal de la variable Y

b) De los días en que el operador B vende algún paquete, ¿cuál es la probabilidad de que el operador A no haya vendido ninguno?

c) ¿Qué porcentaje de días venden entre los dos más de 3 paquetes de viajes?

d) ¿Qué porcentaje de días vende más el operador A que el operador B?

e) ¿Son las variables X e Y independientes?

Sol: b) 0.357, c) 0.25, d) 0.20, e) No son independientes.

Ejercicio 18

En una agencia de viajes que cuenta con dos operadores se consideran las variables aleatorias

X=“número de paquetes vendidos al día por el operador A”

Y=“número de paquetes vendidos al día por el operador B”

En la tabla siguiente se muestran las correspondientes probabilidades conjuntas

X/Y	0	1	2
0	0,15	0,15	0,10
1	0,05	0,20	0,05

X/Y	0	1	2
2	0,10	0,05	0,15

a) Obtener la función de cuantía marginal de la variable Y

b) De los días en que el operador B vende algún paquete, ¿cuál es la probabilidad de que el operador A no haya vendido ninguno?

c) ¿Qué porcentaje de días venden entre los dos más de 3 paquetes de viajes?

d) ¿Qué porcentaje de días vende más el operador A que el operador B?

e) ¿Son las variables X e Y independientes?

Sol: b) 0.357, c) 0.25, d) 0.20, e) No son independientes.

Ejercicio 19

Un explorador perforará una serie de pozos en un área. La probabilidad de encontrar un pozo productivo es de 0.2.

a) ¿Cuál es la probabilidad de que el primer pozo productivo sea el tercero?

b) ¿Cuál es la probabilidad de que el explorador no vaya a encontrar un pozo productivo si solo puede perforar 10?

c) ¿Cuál es la probabilidad de que el tercer encuentro de petróleo ocurra en el quinto pozo que se perfora?

Sol: a) 0.128, b) 0.107, c) 0.03.

Ejercicio 20

Un motor de recomendación que funciona en una plataforma de streaming musical basándose en el histórico de reproducciones de sus usuarios funciona con tres algoritmos. Cada uno de estos algoritmos recomienda siempre una canción para que el usuario pueda escucharla. No obstante, es el mismo usuario el que elige qué algoritmo ha encontrado la mejor opción para él. Así, el 30% de las veces, el usuario elige el algoritmo 1, el 40% el algoritmo 2 y el resto, el algoritmo 3.

Sin embargo, el usuario puede rectificar a los pocos segundos. Eso indicaría que se equivocó al elegir. Se sabe que el 5% de las veces, el usuario rectifica. Y que, esas rectificaciones ascienden al 2% en los algoritmos 2 y 3.

Se pide calcular la probabilidad de que el usuario rectifique, sabiendo que se elige una canción del algoritmo 1.

Sol: 0.12.