

Examen Parcial. 18 de Marzo de 2009

Estadística I. GDO - ADE

Nombre:

Grupo pequeño nº:

Puntuación sobre 25 puntos. Tiempo de realización: 1 h 15 min

Todas las respuestas finales y los cálculos intermedios deben escribirse en esta misma hoja en los espacios reservados para ello. NO SE PUEDE DESGRAPAR LAS HOJAS.

Para las cuentas en sucio se puede usar la parte de atrás de las hojas.

El resultado final ha de expresarse con 4 decimales.

1. A continuación se presenta una muestra de 10 valores de dos variables estadísticas continuas, X e Y .

X:	3.9	4.2	1.9	3.5	1.9	2.2	1.7	5.0	3.8	2.6
Y:	45	65	64	38	64	55	83	46	64	52

Se ha usado el programa R para obtener un resumen numérico de cada una de estas variables y del análisis de regresión de Y sobre X . Se han obtenido los siguientes resultados:

```
> numSummary(datos[,c("x", "y")], statistics=c("mean", "sd",
+ "quantiles"), quantiles=c(0,.25,.5,.75,1))
  mean      sd   0%   25%   50%   75%  100%   n
x  3.07  1.154749  1.7  1.975  3.05  3.875    5  10
y 57.60 13.074148 38.0 47.500 59.50 64.000   83  10

> summary(RegModel.1)
```

Call:

```
lm(formula = y ~ x, data = datos)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-16.9801  -7.8113  -0.7286   8.1756  17.0528
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    76.305     10.998   6.938 0.000120 ***
x              -6.093      3.374  -1.806 0.108576
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.69 on 8 degrees of freedom

Multiple R-squared: 0.2896, Adjusted R-squared: 0.2008

F-statistic: 3.261 on 1 and 8 DF, p-value: 0.1086

- a) (2 pts.) Indica razonadamente cuál de las variables, X o Y , es más dispersa.

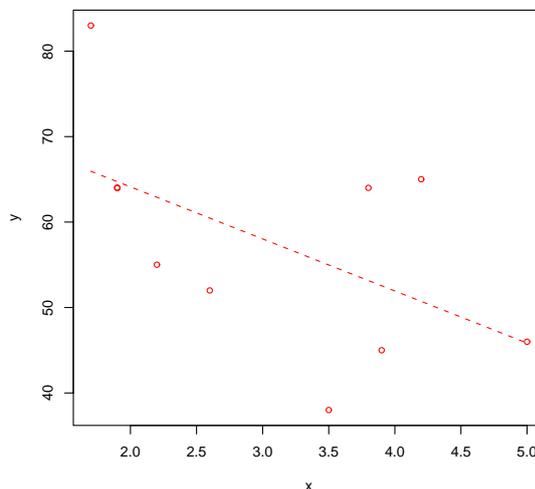
Solución.

$$CV_x = \frac{s_x}{\bar{x}} = \frac{1.154749}{3.07} = 0.3761 \quad CV_y = \frac{s_y}{\bar{y}} = \frac{13.074148}{57.60} = 0.2270.$$

Por lo tanto X es más dispersa al tener mayor coeficiente de variación.

- b) (1.5 pts.) Representa los datos en un diagrama de dispersión.

Solución.



- c) (2 pts.) Obtén, indicando cómo lo haces, el coeficiente de correlación e interpreta su significado de acuerdo a la relación lineal entre las dos variables.

Solución. El coeficiente de correlación es la raíz del coeficiente de determinación. El coeficiente de determinación viene recogido en la salida de R como Multiple R-squared. Por tanto:

$$r_{(x,y)} = \sqrt{r^2_{(x,y)}} = \sqrt{0.2896} = -0.5381.$$

El coeficiente de correlación ha de ser negativo, porque la relación entre ambas variables es negativa (a la vista del diagrama de dispersión y de la ecuación de la recta de regresión), por tanto tomamos la raíz negativa del coeficiente de determinación. Como es un valor no demasiado alto y negativo, podemos decir que existe una relación lineal moderada o débil y negativa entre x e y , es decir, a mayores valores de x , menores valores de y .

- d) (1.5 pts.) Obtén, indicando cómo lo haces, la ecuación de la recta de regresión de y sobre x , y represéntala en el diagrama realizado anteriormente.

Solución. Los coeficientes de la recta de regresión vienen recogidos en la salida de R en la columna Estimate del apartado Coefficients como Intercept (ordenada en el origen) y x (pendiente).

La ecuación de la recta de regresión es: $y = -6.093x + 76.305$.

2. En un grupo de alumnos de 2º de Bachillerato por la rama de Ciencias Sociales de un instituto de Getafe, las expectativas para el año siguiente son: el 30 % quiere estudiar el grado en A.D.E., el 25 % quiere estudiar el grado en Economía y el resto quiere estudiar otros grados. Además, se sabe que de entre los que quieren estudiar A.D.E., el 80 % quiere hacerlo en la Carlos III, y ese porcentaje es del 70 % entre los que estudiarán Economía y del 50 % entre el resto de los alumnos.

- a) **(3 pts.)** Calcula la probabilidad que tiene un alumno elegido al azar en ese grupo de estudiar en la Carlos III.

Solución. Por la ley de la probabilidad total:

$$\begin{aligned} P(UC3M) &= P(UC3M|ADE) \cdot P(ADE) + P(UC3M|EC) \cdot P(EC) \\ &+ P(UC3M|otros) \cdot P(otros) \\ &= 0.8 \cdot 0.3 + 0.7 \cdot 0.25 + 0.5 \cdot 0.45 = 0.6400. \end{aligned}$$

- b) **(3 pts.)** Sabiendo que un alumno de ese grupo acaba estudiando en la Carlos III, ¿cuál es la probabilidad de que estudie el grado en A.D.E.?

Solución. Por el teorema de Bayes:

$$\begin{aligned} P(ADE|UC3M) &= \frac{P(UC3M|ADE) \cdot P(ADE)}{P(UC3M)} \\ &= \frac{0.8 \cdot 0.3}{0.64} = 0.3750. \end{aligned}$$

3. Una empresa aseguradora recibe partes de accidentes de vehículos a razón de 3 al día. Además, se sabe que el número de partes que llegan al día sigue una distribución de Poisson.

- a) **(3 pts.)** Calcula la probabilidad que en un día determinado lleguen más de 2 partes.

Solución. Sea $X = \text{“número de partes que llegan al día”}$; $X \sim \mathcal{P}(3)$.

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) = 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\ &= 1 - \left(\frac{e^{-3} \cdot 3^0}{0!} + \frac{e^{-3} \cdot 3^1}{1!} + \frac{e^{-3} \cdot 3^2}{2!} \right) \\ &= 1 - e^{-3} (1 + 3 + 4.5) = 1 - 0.4232 = 0.5768. \end{aligned}$$

- b) **(3 pts.)** Calcula la probabilidad de que en dos días consecutivos lleguen 5 partes.

Solución. Sea $Y = \text{“número de partes que llegan en dos días”}$; $Y \sim \mathcal{P}(3 \cdot 2) = \mathcal{P}(6)$.

$$P(Y = 5) = \frac{e^{-6} \cdot 6^5}{5!} = 0.1606.$$

- c) **(3 pts.)** Para cada parte que llega, la cuantía que la compañía aseguradora tiene que pagar al asegurado sigue una distribución normal de media 305 euros y desviación típica 500 euros. Calcula la probabilidad de que en un parte determinado haya que pagarle al asegurado más de 1000 euros.

Solución. Sea $C = \text{“cuantía que se paga por parte”}$; $C \sim N(305, 500)$.

$$\begin{aligned} P(C > 1000) &= P\left(\frac{C-305}{500} > \frac{1000-305}{500}\right) \quad \left\{ Z = \frac{C-305}{500} \sim N(0, 1) \right\} \\ &= P(Z > 1.39) = 0.0823. \end{aligned}$$

- d) **(3 pts.)** La probabilidad de que la cuantía a pagar para un parte determinado supere los 305 euros es 0.5. Calcula la probabilidad de que entre los 10 siguientes partes que lleguen haya 4 con una cuantía de pago superior a los 305 euros.

Solución. Sea $W = \text{“número de partes, entre los 10 siguientes, con una cuantía de pago superior a los 305 euros.”}$

Suponiendo que los distintos partes son independientes entre sí, tenemos que $W \sim \mathcal{B}(10, 0.5)$.

$$P(W = 4) = \binom{10}{4} 0.5^4 \cdot 0.5^6 = 0.2051.$$