EXAMEN DE ESTADÍSTICA

2º Fisioterapia   Modelo A   **28 de junio de 2018**

Duración: 2 horas.

(2,5 pts.) 1. Se ha medido el tiempo de espera para que un medicamento A haga efecto en un grupo de 150 pacientes, obteniéndose los siguientes resultados:

Tiempo	Pacientes
$(0, 5]$	5
$(5, 10]$	15
$(10, 15]$	32
$(15, 20]$	36
$(20, 30]$	42
$(30, 60]$	20

Se pide:

a) ¿Existen datos atípicos en la muestra? Justificar la respuesta.

b) ¿Cuál fue el tiempo de espera mínimo para el 20 % de los pacientes que más esperaron?

c) ¿Cuál fue el tiempo medio de espera? ¿Es muy representativa esta medida? Justificar la respuesta

d) Estudiar el apuntamiento de la distribución.

e) Si tomamos otro grupo de pacientes tratados con el medicamento A cuya media fue 18 min con una desviación típica de 15 min, ¿en cuál de los dos grupos sería más alto un tiempo de espera de 25 min?

Usar las siguientes sumas para los cálculos: $\sum x_i = 3105$ min, $\sum x_i^2 = 83650$ min^2, $\sum (x_i - \bar{x})^3 = 206851{,}65$ min^3 y $\sum (x_i - \bar{x})^4 = 8140374{,}96$ min^4.

Solución

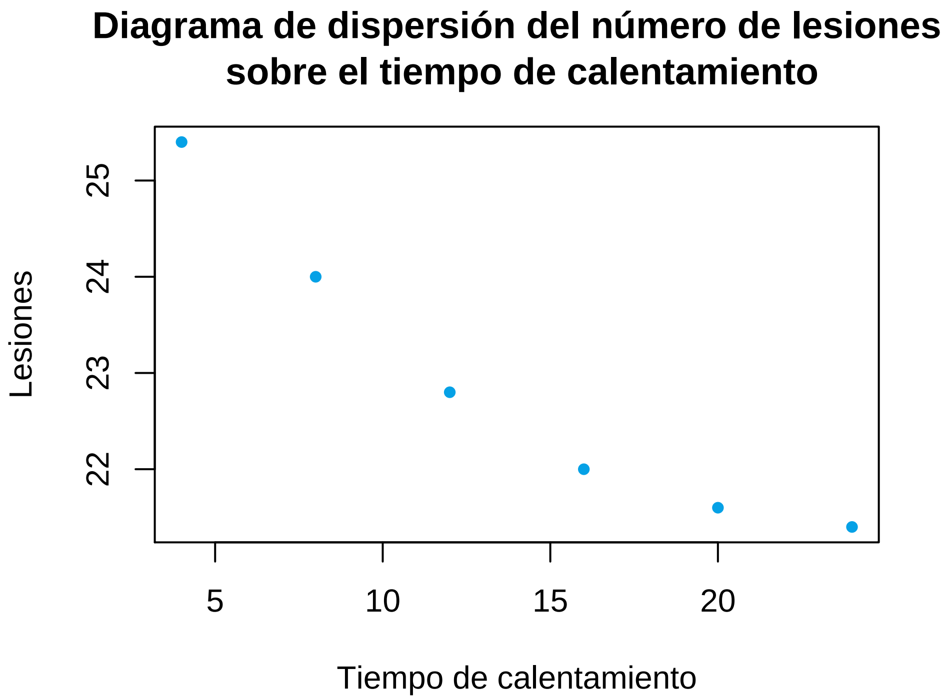
(2,5 pts.) 2. Se ha medido el número medio de horas de estudio semanales X y la nota en un test de acceso a una clínica Y, de 8 aspirantes, obteniendo los siguientes resultados:

$\sum x_i = 15{,}9$, $\sum \log(x_i) = 3{,}852$, $\sum y_j = 41{,}5$, $\sum \log(y_j) = 11, 511$,
$\sum x_i^2 = 42{,}23$, $\sum \log(x_i)^2 = 5{,}559$, $\sum y_j^2 = 274{,}25$, $\sum \log(y_j)^2 = 20{,}742$,
$\sum x_i y_j = 103{,}3$, $\sum x_i \log(y_j) = 28{,}047$, $\sum \log(x_i) y_j = 32{,}616$.

a) Calcular las ecuaciones de los modelos de regresión logarítmico y exponencial de la nota en función del número medio de horas de estudio.

b) Utilizar el mejor de los dos modelos anteriores para predecir la nota si se estudia 3,2 horas a la semana.

Solución



a)

b) $\bar{x} = 14$ min, $s_x^2 = 46{,}6667$ min^2.
$\overline{\log(x)} = 2{,}4828$ log(min), $s_{\log(x)}^2 = 0{,}3659$ log(min)^2.
$\bar{y} = 22{,}8667$ lesiones, $s_y^2 = 2{,}0356$ lesiones^2.
$\overline{\log(y)} = 3{,}1278$ log(lesiones), $s_{\log(y)}^2 = 0{,}0038$ log(lesiones)^2.
$s_{x\log(y)} = -0{,}4029$, $s_{\log(x)y} = -0{,}8593$
Coeficiente de determinanción exponencial: $r^2 = 0{,}9274$
Coeficiente de determinación logarítmico: $r^2 = 0{,}9913$
Por tanto, el modelo de regresión exponencial es mejor para predecir el número de lesiones en función del tiempo de calentamiento.
Modelo de regresión exponencial: $y = e^{3{,}2486+-0{,}0086x}$.
Predicción: $y(20) = 21{,}6712$ lesiones.

c) El modelo logarítmico es mejor para predecir el tiempo de calentamiento en función del número de lesiones.
Modelo de regresión logarítmico: $x = 28{,}6971 + -2{,}3483\log(y)$.
Predicción: $x(10) = 23{,}29$ min.

d) De acuerdo al coeficiente de determinación ambas predicciones son muy fiables pero la última lo es menos ya que es para un valor que no está incluido en el rango de valores de la muestra.

(1,5 pts.) 3. El 40 % de los atletas de una población tienen una madre muy deportista y un 30 % tienen un padre muy deportista. Si el 50 % de los atletas de esta población tienen algún progenitor muy deportista, se pide:

a) Calcular la probabilidad de que, eligiendo uno de estos atletas al azar, su madre sea muy deportista, si su padre lo es.

b) Calcular la probabilidad de que, eligiendo uno de estos atletas al azar, su padre sea muy deportista, si su madre no lo es.

c) ¿Son independientes los sucesos correspondientes a tener una madre muy deportista y tener un padre muy deportista? Razonar la respuesta.

Solución

(1,5 pts.) 4. El número medio de lesiones en un torneo internacional de tenis es de 2.

Se pide:

a) Calcular la probabilidad de que en uno de tales torneos haya más de 2 lesiones.

b) Si un circuito consta de seis torneos internacionales, ¿cuál es la probabilidad de que en alguno de ellos no se produzca ninguna lesión?

Solución

a) Sea X el número de pacientes que llegan en 1 horas. $X \sim P(2)$ y $P(X > 4) = 0{,}0527$.

b) Sea Y el número de horas en un día en las que algún paciente no puede ser atendido. $Y \sim B(6, 0{,}0527)$ y $P(Y > 0) = 0{,}2771$. Se necesitan 5 empleados para que esta probabilidad sea menor del 10 %.
Se necesitan 5 empleados para que esta probabilidad sea menor del 10 %, ya que $P(X > 5) = 0{,}0527$ y $P(Y > 0) = 0{,}0954$, siendo ahora $Y \sim B(6, 0{,}0166)$.

(2 pts.) 5. Un test diagnóstico para detectar el dopaje en atletas da un resultado positivo cuando la concentración de una determinada sustancia en sangre es mayor de 4 μg/ml. Si la distribución de la concentración de a sustancia en sangre en atletas dopados sigue un modelo de distribución normal con media 4,5 μg/ml y desviación típica 0,2 μg/ml, y en atletas no dopados sigue un modelo de distribución normal con media 3 μg/ml y desviación típica 0,3 μg/ml,

a) ¿cuál es la sensibilidad y la especificidad del test?

b) Si hay un 10 % de atletas dopados en una competición, ¿cuales son los valores predictivos positivo y negativo del test? Interpretarlos.

Solución
