

Comenzado el

Estado

Finalizado en

Tiempo empleado

Puntos

Calificación

Pregunta 1

Incorrecta

Puntúa 0,00 sobre 1,00



La distribución a posteriori de los parámetros es:

Seleccione una:

- ☐ a. siempre normal en el enfoque frecuentista si el modelo es correcto
- ☐ b. el supuesto más importante en el enfoque bayesiano
- ☒ c. la herramienta con la cual se evalúa la bondad de ajuste de un modelo en el enfoque frecuentista **✗**
- ☐ d. la principal fuente de información para hacer inferencia en el enfoque bayesiano

Respuesta incorrecta.

La respuesta correcta es: la principal fuente de información para hacer inferencia en el enfoque bayesiano

Pregunta 2

Correcta

Puntúa 1,00 sobre 1,00



Se propone el siguiente modelo para los datos:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$$

donde **Y = gasto**, el gasto medio por admisión (en dólares) y **X₁ = salario**, salario promedio por empleado (en dólares).

¿Cuál es el significado de β_1 en el modelo? Indique la opción correcta.

Seleccione una:

- ☐ a. Valores positivos de β_1 significan que X_1 influye en Y, mientras que valores negativos de β_1 significan que X_1 no influye en Y.
- ☐ b. β_1 es el porcentaje de variabilidad de Y que queda explicado por el modelo.
- ☐ c. β_1 es la variable aleatoria que estima la pendiente de la relación lineal entre el valor esperado de Y y X_1 .
- ☒ d. El valor esperado de Y aumenta (o disminuye, según el signo) β_1 unidades (dólares) cuando X_1 aumenta una unidad (dólar). ✓

Respuesta correcta

La respuesta correcta es: El valor esperado de Y aumenta (o disminuye, según el signo) β_1 unidades (dólares) cuando X_1 aumenta una unidad (dólar).

Pregunta 3

Correcta

Puntúa 1,00 sobre 1,00



La **inferencia estadística** es un proceso que consiste en:

Seleccione una:

- ☐ a. Calcular probabilidades de eventos observaciones, suponiendo conocidos los parámetros poblacionales.
- ☐ b. Detectar variables latentes o factores como combinaciones de las variables originales.
- ☒ c. Estimar, basado en los datos observados, parámetros y patrones inherentes de la población. ✓
- ☐ d. Predecir futuras observaciones.

Respuesta correcta

La respuesta correcta es: Estimar, basado en los datos observados, parámetros y patrones inherentes de la población.

Pregunta 4

Correcta

El enfoque estadístico bayesiano, a diferencia del frecuentista:

Puntúa 1,00 sobre

1,00



Seleccione una:

- ☐ a. supone fijos (no aleatorios) a los parámetros
- ☒ b. supone aleatorios a los parámetros ✓
- ☐ c. no hace ningún supuesto distribucional
- ☐ d. supone aleatorias a las observaciones

Respuesta correcta

La respuesta correcta es: supone aleatorios a los parámetros

Pregunta 5

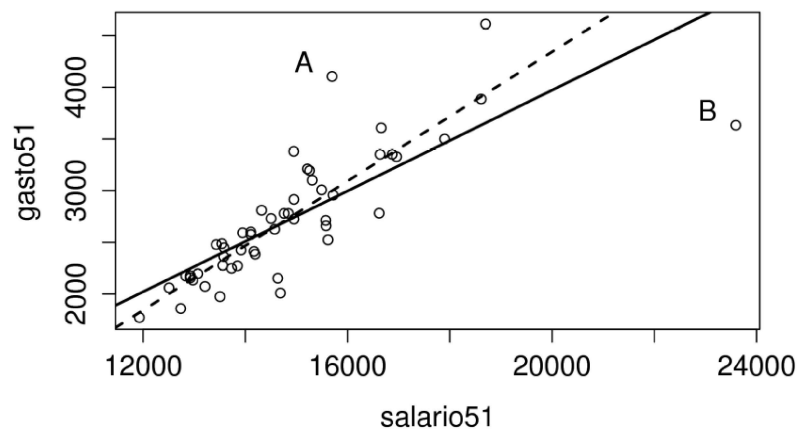
Correcta

Puntúa 1,00 sobre

1,00



La siguiente figura describe un scatterplot de 51 datos con 2 ajustes lineales distintos.



Comparemos los resultados que obtendríamos al ajustar un modelo lineal simple (por mínimos cuadrados, instrucción `lm` del R) con las 49 observaciones originales (sin A y sin B) con el resultado de ajustar un modelo de regresión lineal simple con las 49 observaciones originales (sin A y sin B) más la observación A, que llamaremos datos aumentados. Indique la opción correcta:

Seleccione una:

- ☐ a. para el conjunto de datos aumentados, **disminuye** el $\hat{\sigma}^2$ y en consecuencia la longitud del intervalo de confianza para la respuesta media de Y para un valor x_h prefijado **disminuye** en comparación con el que se obtiene para el conjunto de datos originales.
- ☐ b. para el conjunto de datos aumentados, **disminuye** el $\hat{\sigma}^2$ y en consecuencia la longitud del intervalo de confianza para la respuesta media de Y para un valor x_h prefijado **aumenta** en comparación con el que se obtiene para el conjunto de datos originales.

- ☒ c. para el conjunto de datos aumentados, **aumenta** el $\hat{\sigma}^2$ y en consecuencia la longitud del intervalo de confianza para la respuesta media de Y para un valor x_h prefijado **aumenta** en comparación con el que se obtiene para el conjunto de datos originales. ✓
- ☐ d. para el conjunto de datos aumentados, **aumenta** el $\hat{\sigma}^2$ y en consecuencia la longitud del intervalo de confianza para la respuesta media de Y para un valor x_h prefijado **disminuye** en comparación con el que se obtiene para el conjunto de datos originales.

Respuesta correcta

La respuesta correcta es: para el conjunto de datos aumentados, **aumenta** el $\hat{\sigma}^2$ y en consecuencia la longitud del intervalo de confianza para la respuesta media de Y para un valor x_h prefijado **aumenta** en comparación con el que se obtiene para el conjunto de datos originales.

Pregunta 6

Correcta

Puntúa 1,00 sobre 1,00



El **p-valor** de un parámetro asociado a una variable en un modelo estadístico de regresión lineal es:

Seleccione una:

- ☐ a. La probabilidad que las observaciones de la variable asociada al parámetro sean distintas de cero (estadísticamente).
- ☒ b. La probabilidad, suponiendo que el parámetro es cero, de observar un valor del estadístico tan o mas extremo que el observado. ✓
- ☐ c. La probabilidad, suponiendo que el parámetro es distinto de cero, de observar un valor del estadístico a lo sumo tan extremo como el observado.
- ☐ d. La probabilidad que la variable asociada al parámetro sirva para explicar a la respuesta.

Respuesta correcta

La respuesta correcta es: La probabilidad, suponiendo que el parámetro es cero, de observar un valor del estadístico tan o mas extremo que el observado.

Pregunta 7

Correcta

Puntúa 1,00 sobre 1,00



La **función de verosimilitud** es una función que:

Seleccione una:

- ☐ a. Tiene como “fijos” a los parámetros poblacionales y como “variables” a las observaciones y toma un valor máximo en la muestra observada.

- ☐ b. Tienen como “fijas” a las observaciones y como “variables” a los parámetros y toma un valor máximo en los verdaderos valores de los parámetros poblacionales.
- ☒ c. Tiene como “fijas” a las observaciones y como “variables” a los parámetros. ✓
- ☐ d. Tiene como “fijos” a los parámetros y como “variables” a las observaciones.

Respuesta correcta

La respuesta correcta es: Tiene como “fijas” a las observaciones y como “variables” a los parámetros.

Pregunta 8

Correcta

Puntúa 1,00 sobre 1,00



En un test de hipótesis:

Seleccione una:

- ☐ a. Al aumentar el tamaño de muestra disminuyen tanto el error tipo 1 como el error tipo 2.
- ☒ b. A un nivel de significancia dado, un aumento del tamaño de muestra aumenta la potencia. ✓
- ☐ c. Al aumentar el nivel de significancia, con el tamaño de muestra fijo, siempre aumenta la potencia.
- ☐ d. Al aumentar el tamaño de muestra disminuye el error tipo 1 pero aumenta el error tipo 2.

Respuesta correcta

La respuesta correcta es: A un nivel de significancia dado, un aumento del tamaño de muestra aumenta la potencia.

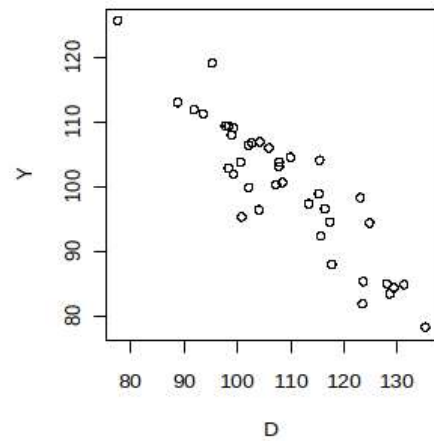
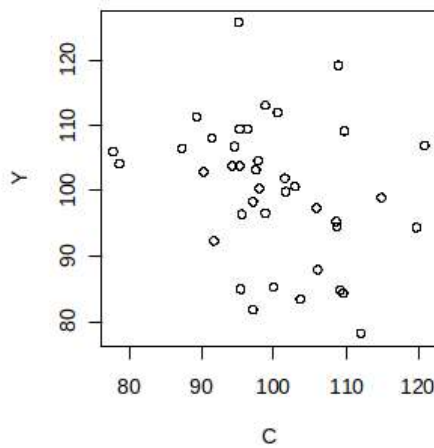
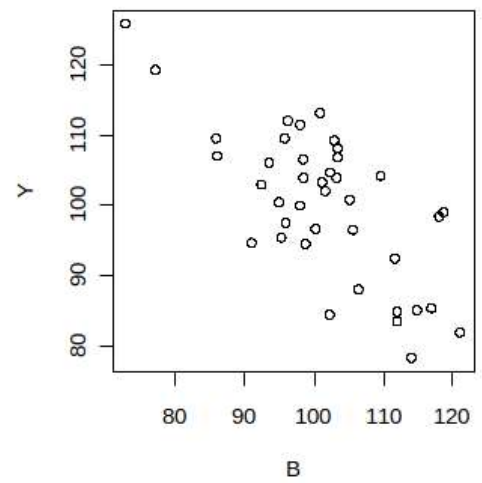
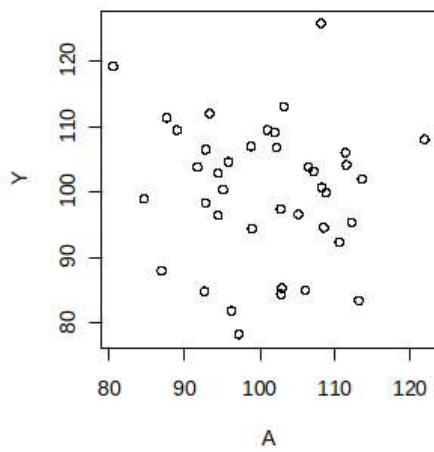
Pregunta 9

Correcta

Puntúa 1,00 sobre 1,00



Se miden 5 variables sobre 40 individuos. Éstas son la respuesta Y y cuatro covariables A, B, C y D. Los scatter plots de Y versus cada covariable son



Los coeficientes de correlación muestral de Pearson entre Y las cuatro A, B, C y D covariables son, respectivamente (indique la opción correcta):

Seleccione una:

- ☐ a. **A:** 0.9, **B:** 0.7, **C:** 0.3, **D:** 0
- ☒ b. **A:** 0, **B:** -0.7, **C:** -0.3, **D:** -0.9 ✓
- ☐ c. **A:** -0.9, **B:** -0.3, **C:** 0.3, **D:** 0
- ☐ d. **A:** 0, **B:** 0.7, **C:** 0.3, **D:** 0.9

Respuesta correcta

La respuesta correcta es: **A:** 0, **B:** -0.7, **C:** -0.3, **D:** -0.9

Pregunta 10

Correcta

Puntúa 1,00 sobre 1,00



El **sobre-ajuste** (overfitting) es un problema producto de:

Seleccione una:

- ☐ a. Usar un modelo poco flexible, que no permite captar el patrón general de los datos.
- ☒ b. Un exceso de variabilidad de los estimadores de un modelo altamente complejo. ✓

- ☐ c. Aplicar un modelo simple, pero “erróneo”, que no representa bien la verdadera relación de los datos.
- ☐ d. Los sesgos del modelo ajustado.

Respuesta correcta

La respuesta correcta es: Un exceso de variabilidad de los estimadores de un modelo altamente complejo.

Pregunta 11

Correcta

Puntúa 1,00 sobre 1,00



La **relación entre Estimador y Estimación** es:

Seleccione una:

- ☒ a. El estimador es una variable aleatoria, mientras que la estimación es la realización de la misma. ✓
- ☐ b. La estimación es la estandarización del estimador, que se obtiene al dividir al estimador por su desvío estándar (estimado).
- ☐ c. La estimación es el procedimiento matemático que se aplica a la muestra, y el estimador es el resultado numérico de ese cálculo.
- ☐ d. Tanto estimador como estimación son los valores numéricos observados provenientes de aplicar un cálculo matemático a la muestra de datos.

Respuesta correcta

La respuesta correcta es: El estimador es una variable aleatoria, mientras que la estimación es la realización de la misma.

Pregunta 12

Correcta

Puntúa 1,00 sobre 1,00



Se propone el siguiente modelo para los datos:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad (1)$$

donde **Y = gasto**, el gasto medio por admisión (en dólares) y **X₁ = salario**, salario promedio por empleado (en dólares).

El ajuste del modelo (1) a los datos con el R figura a continuación

```
> ajuste <- lm(gasto ~ salario)
> summary(ajuste)
```

```
Call:
lm(formula = gasto ~ salario)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2142.4	351.7	-6.091	1.96e-07	***
salario	0.3283	0.0238	13.756	< 2e-16	***

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

Residual standard error: 255 on 47 degrees of freedom

Multiple R-squared: 0.801, Adjusted R-squared: 0.7968

F-statistic: 189.2 on 1 and 47 DF, p-value: < 2.2e-16

El modelo de regresión lineal explica el _____ % de la variabilidad de Y, la respuesta. (Indique la opción que completa el espacio en blanco).

Seleccione una:

- ☐ a. 18.92
- ☐ b. 32.83
- ☒ c. 80.1 ✓
- ☐ d. 79.68

Respuesta correcta

La respuesta correcta es: 80.1

Pregunta 13

Correcta

Puntúa 1,00 sobre 1,00



Al agregar a un conjunto de datos una nueva observación con alto leverage (indique la opción correcta):

Seleccione una:

- ☐ a. siempre se mantiene constante la significatividad de los coeficientes del modelo
- ☐ b. disminuye la significatividad de los coeficientes del modelo sólo cuando la observación sigue el patrón lineal de los datos
- ☒ c. aumenta la significatividad de los coeficientes del modelo sólo cuando la observación sigue el patrón lineal de los datos ✓
- ☐ d. siempre aumenta la significatividad de los coeficientes del modelo

Respuesta correcta

La respuesta correcta es: aumenta la significatividad de los coeficientes del modelo sólo cuando la observación sigue el patrón lineal de los datos

Pregunta 14

Correcta

Puntúa 1,00 sobre 1,00



Se propone el siguiente modelo para los datos:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \quad (1)$$

donde Y = **gasto**, el gasto medio por admisión (en dólares) y X_1 = **salario**, salario promedio por empleado (en dólares).

El ajuste del modelo (1) a los datos con el R figura a continuación

```
> ajuste <- lm(gasto ~ salario)
> summary(ajuste)
```

```
Call:
lm(formula = gasto ~ salario)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2142.4	351.7	-6.091	1.96e-07	***
salario	0.3283	0.0238	13.756	< 2e-16	***

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

Residual standard error: 255 on 47 degrees of freedom

Multiple R-squared: 0.801, Adjusted R-squared: 0.7968

F-statistic: 189.2 on 1 and 47 DF, p-value: < 2.2e-16

Indique el valor estimado de β_1 (o sea $\hat{\beta}_1$):

Seleccione una:

- ☐ a. 0.02387
- ☐ b. -2142.4
- ☐ c. 0.801
- ☒ d. 0.3283 ✓

Respuesta correcta

La respuesta correcta es: 0.3283

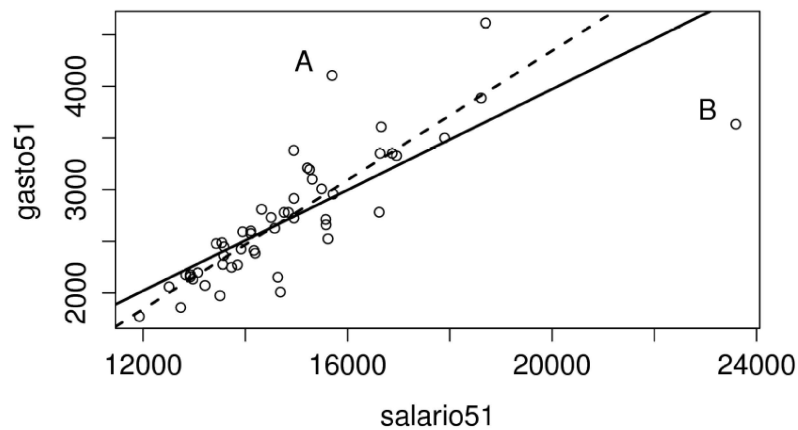
Pregunta 15

Correcta

Puntúa 1,00 sobre 1,00



La siguiente figura describe un scatterplot de 51 datos con 2 ajustes lineales distintos.



Los dos ajustes corresponden a mínimos cuadrados (lm en R) y a la regresión robusta (lmrob) ¿Cuál recta estimada corresponde a la regresión robusta?

Seleccione una:

- ☒ a. La recta de línea punteada. ✓
- ☐ b. La recta de línea continua.

Respuesta correcta

La respuesta correcta es: La recta de línea punteada.

Pregunta 16

Correcta

Puntúa 1,00 sobre 1,00



Interesa poder explicar la esperanza del gasto medio por admisión hospitalaria usando covariables explicativas. Para eso se recabaron, durante 1982, los datos de 49 distritos elegidos al azar en Estados Unidos en los que se midieron varias variables. Entre ellas:

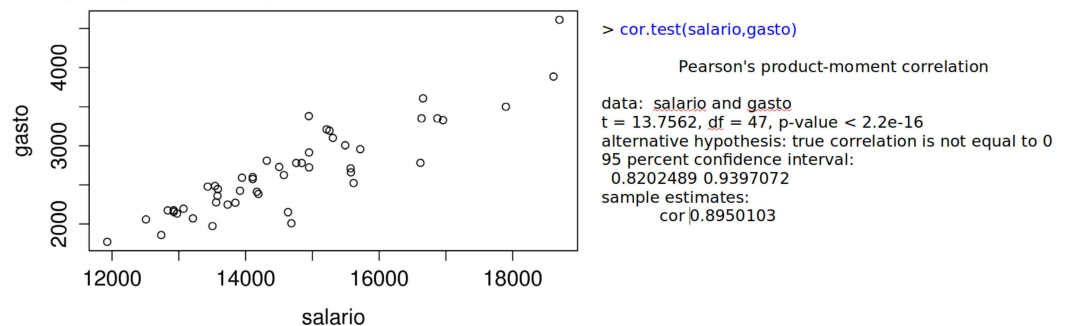
gasto : gasto medio por admisión (en dólares)

tiempo : tiempo promedio de internación (en días)

salario : salario promedio por empleado (en dólares).

A continuación figura el scatter plot de **gasto** versus **salario**. Además se calculó el coeficiente de correlación entre **gasto** (el gasto medio por admisión) y **salario** (el salario promedio por empleado).

El gráfico y la salida del R figuran a continuación.



En base a la salida del R se quiere hacer un test para la verdadera correlación poblacional entre **salario** y **gasto**. Indique cuáles son las hipótesis del test

Seleccione una:

- ☐ a. $H_0 : r = 0.895$ vs. $H_1 : r \neq 0.895$
- ☐ b. $H_0 : r = 0$ vs. $H_1 : r \neq 0$

- ☒ c. $H_0: \rho = 0$ vs. $H_1: \rho \neq 0$ ✓
- ☐ d. $H_0: \rho = 0.895$ vs. $H_1: \rho \neq 0.895$

Respuesta correcta

La respuesta correcta es: $H_0: \rho = 0$ vs. $H_1: \rho \neq 0$

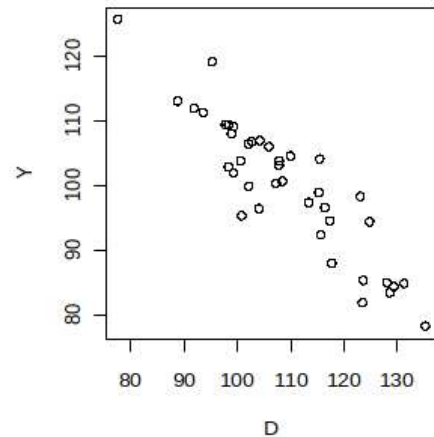
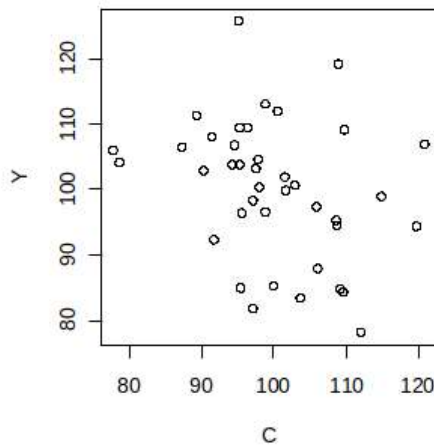
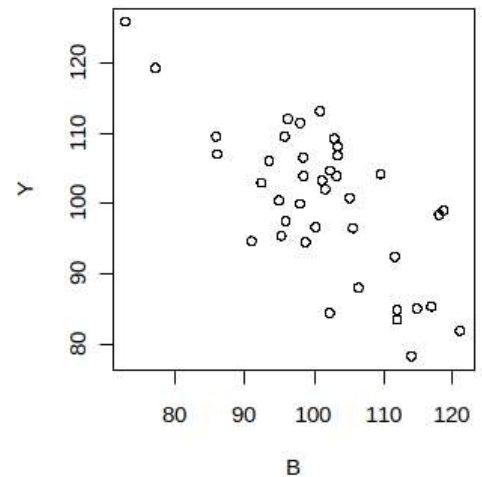
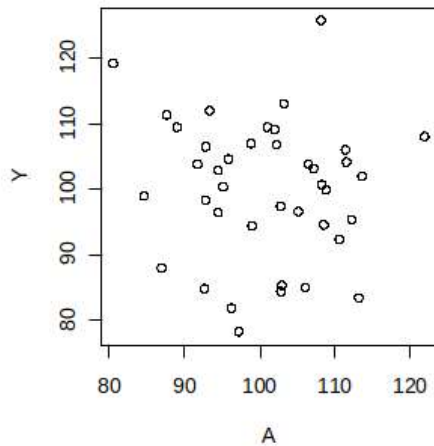
Pregunta 17

Correcta

Puntúa 1,00 sobre 1,00



Se miden 5 variables sobre 40 individuos. Éstas son la respuesta Y y cuatro covariables A, B, C y D. Los scatter plots de Y versus cada covariable son



En primera instancia se propone un modelo de regresión lineal simple para explicar a Y en función de una sola covariable. ¿Cuál modelo elegiría explorar primero?

Seleccione una:

- ☐ a. $Y = \beta_0 + \beta_1 B + \varepsilon$
- ☐ b. $Y = \beta_0 + \beta_1 A + \varepsilon$

- ☒ c. $Y = \beta_0 + \beta_1 D + \varepsilon$ ✓
- ☐ d. $Y = \beta_0 + \beta_1 C + \varepsilon$

Respuesta correcta

La respuesta correcta es: $Y = \beta_0 + \beta_1 D + \varepsilon$

Pregunta 18

Incorrecta

Puntúa 0,00 sobre 1,00



El **sobre-ajuste** (overfitting) se puede controlar mediante:

Seleccione una:

- ☐ a. La utilización de modelos muy flexibles y con muchos parámetros (como las redes neuronales).
- ☐ b. El aumento del tamaño de la muestra con la que se ajusta el modelo.
- ☒ c. La utilización de métodos de estimación robustos (poco sensibles a los datos atípicos). ✗
- ☐ d. El uso de técnicas del tipo “validación cruzada”.

Respuesta incorrecta.

La respuesta correcta es: El uso de técnicas del tipo “validación cruzada”.

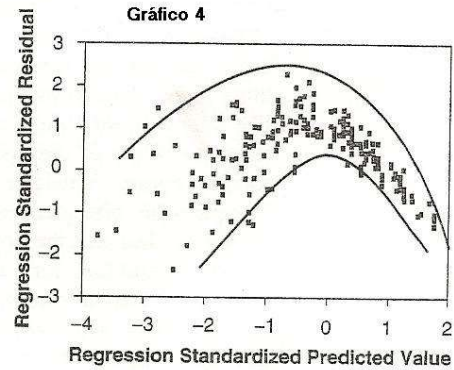
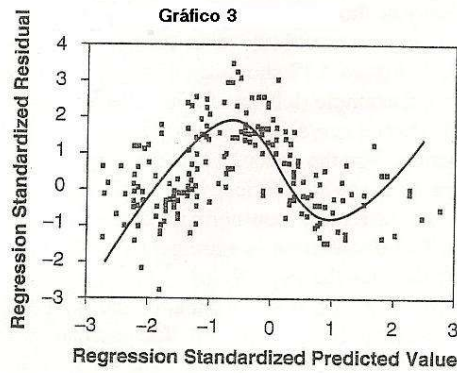
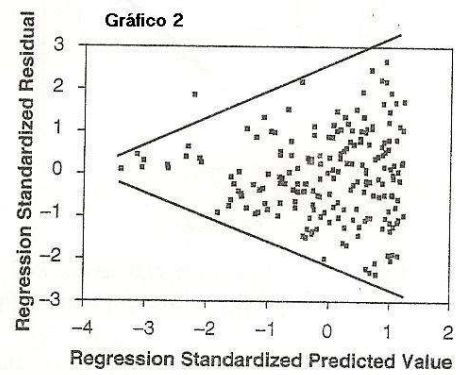
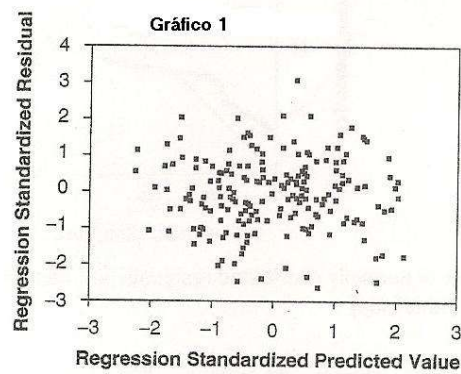
Pregunta 19

Correcta

Puntúa 1,00 sobre 1,00



La figura siguiente muestra los gráficos de residuos estandarizados versus los valores predichos para cuatro conjuntos de datos. En algunos de ellos se satisfacen los supuestos para aplicar el método de regresión lineal, en otros no.



¿Cuál descripción caracteriza más apropiadamente al conjunto de datos del Gráfico 2?

Seleccione una:

- ☐ a. No se satisface el supuesto de linealidad.
- ☒ b. No se satisface el supuesto de homoscedasticidad. ✓
- ☐ c. Se satisfacen los supuestos.
- ☐ d. No se satisfacen ni la linealidad ni la homoscedasticidad.

Respuesta correcta

La respuesta correcta es: No se satisface el supuesto de homoscedasticidad.

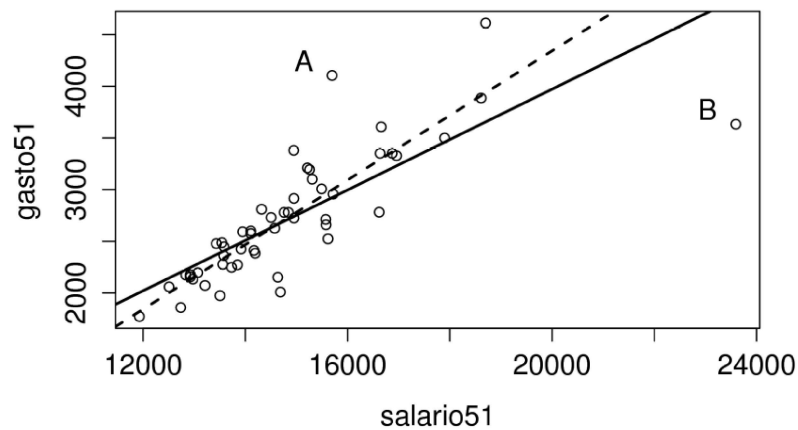
Pregunta 20

Correcta

Puntúa 1,00 sobre 1,00



La siguiente figura describe un scatterplot de 51 datos con 2 ajustes lineales distintos.



Indique que frase describe mejor a la observación correspondiente a B:

Seleccione una:

- ☒ a. es una observación influyente que no sigue la estructura lineal de los demás datos ✓
- ☐ b. es una observación no influyente que sigue la estructura lineal de los demás datos
- ☐ c. es una observación influyente que sigue la estructura lineal de los demás datos
- ☐ d. es una observación no influyente que no sigue la estructura lineal de los demás datos

Respuesta correcta

La respuesta correcta es: es una observación influyente que no sigue la estructura lineal de los demás datos

Pregunta 21

Correcta

Puntúa 1,00 sobre 1,00



Se propone el siguiente modelo para los datos:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \quad (1)$$

donde **Y = gasto**, el gasto medio por admisión (en dólares) y **X₁ = salario**, salario promedio por empleado (en dólares).

Los ε_i en el modelo (1) representan

Seleccione una:

- ☐ a. los errores que son observables.
- ☐ b. los residuos que no son observables.
- ☒ c. los errores que no son observables. ✓

- ☐ d. los residuos que son observables.

Respuesta correcta

La respuesta correcta es: los errores que no son observables.

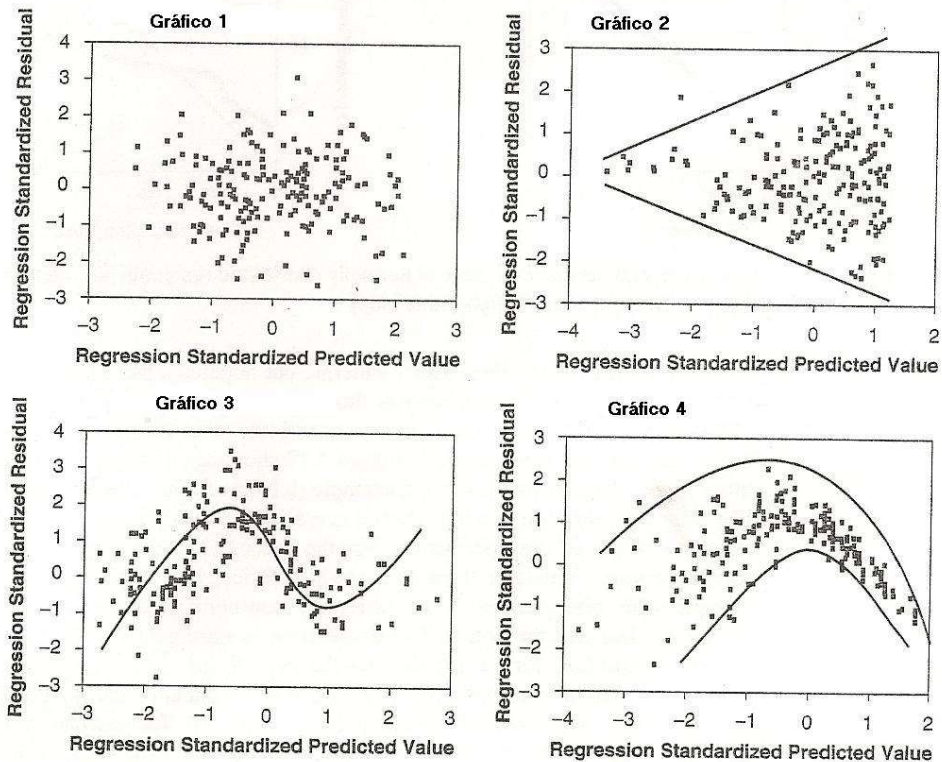
Pregunta 22

Correcta

Puntúa 1,00 sobre 1,00



La figura siguiente muestra los gráficos de residuos estandarizados versus los valores predichos para cuatro conjuntos de datos. En algunos de ellos se satisfacen los supuestos para aplicar el método de regresión lineal, en otros no.



¿Cuál descripción caracteriza más apropiadamente al conjunto de datos del Gráfico 3?

Seleccione una:

- ☐ a. Se satisfacen los supuestos.
- ☐ b. No se satisface el supuesto de homoscedasticidad.
- ☐ c. No se satisfacen ni la linealidad ni la homoscedasticidad.
- ☒ d. No se satisface el supuesto de linealidad. ✓

Respuesta correcta

La respuesta correcta es: No se satisface el supuesto de linealidad.

Pregunta 23

Correcta

Puntúa 1,00 sobre 1,00



Se propone el siguiente modelo para los datos:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i \quad (1)$$

donde **Y = gasto**, el gasto medio por admisión (en dólares) y **X₁ = salario**, salario promedio por empleado (en dólares).

El ajuste del modelo (1) a los datos con el R figura a continuación

```
> ajuste <- lm(gasto ~ salario)
> summary(ajuste)
```

```
Call:
lm(formula = gasto ~ salario)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-2142.4	351.7	-6.091	1.96e-07	***
salario	0.3283	0.0238	13.756	< 2e-16	***

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

Residual standard error: 255 on 47 degrees of freedom

Multiple R-squared: 0.801, Adjusted R-squared: 0.7968

F-statistic: 189.2 on 1 and 47 DF, p-value: < 2.2e-16

El estimador del desvío estándar σ es (indique la opción correcta):

Seleccione una:

- ☐ a. 351.7
- ☒ b. 255 ✓
- ☐ c.
0.02387
- ☐ d. 189.2

Respuesta correcta

La respuesta correcta es: 255

Pregunta 24

Correcta

Puntúa 1,00 sobre 1,00



La figura siguiente muestra los gráficos de residuos estandarizados versus los valores predichos para cuatro conjuntos de datos. En algunos de ellos se satisfacen los supuestos para aplicar el método de regresión lineal, en otros no.