Estadística: Modelos Lineales

Final Enero 2.006, Tipo: A

Apellidos y Nombre: ____________________

DNI: ____________________

Grupo: ____________________

Profesor : ____________________

Sección 1. Instrucciones

1. Salvo que se indique lo contrario, las preguntas bien contestadas valen un punto. Puede haber más de una respuesta correcta, y para obtener puntuación has de señalarlas todas. Preguntas que no estén bien contestadas puntúan -0.5 veces su valor.

2. Intento medir conocimientos y no agudeza visual. Inevitablemente, en un examen de este tipo hay que prestar mucha atención. Cada curso hay personas que echan a perder una nota potencialmente buena por responder temeraria o atolondradamente.

   ¡Por favor, fíjate bien en todos los detalles!

   Te ayudará proceder por exclusión de absurdos. Si una pregunta te parece ambigua, anota brevemente la razón al margen y no la contestes.

Sección 2. Cuestiones de elección múltiple

3. Si A es cuadrada, simétrica, con valores propios $\lambda_1, \ldots, \lambda_p$, entonces:
   (a) $\mathrm{tr}(A) = \sum_{i=1}^{p} \lambda_i$.
   (b) $\mathrm{tr}(A) = \prod_{i=1}^{p} \lambda_i$.
   (c) $\mathrm{tr}(A) \geq 0$.
   (d) Nada de lo anterior.

4. Los valores propios de una matriz simétrica son siempre. . .
   (a) Positivos.
   (b) Reales.
   (c) Simétricos.
   (d) Nada de lo anterior.

5. Una matriz simétrica, semidefinida positiva e idempotente tiene valores propios que son siempre. . .
   (a) Cero.
   (b) Cero o uno.
   (c) Uno o menos uno.
   (d) Nada de lo anterior.

6. Sea $\boldsymbol{X}$ la matrizde diseño $N \times p$ de un modelo lineal $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. La condición de que rango$(\boldsymbol{X}) = p < N$:
   (a) Excluye la posibilidad de multicolinealidad exacta.
   (b) Implica que el número de grados de libertad en la estimación será mayor que cero.
   (c) Garantiza la unicidad de la proyección de $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$.
   (d) Nada de lo anterior.

7. Las ecuaciones normales $\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}$:
   (a) Siempre tienen una solución única.
   (b) Siempre tienen al menos una solución.
   (c) Siempre tienen múltiples soluciones; de hecho, infinitas.
   (d) Nada de lo anterior.

8. Entre los supuestos habituales en el modelo lineal ordinario está que:
   (a) El vector de perturbaciones $\boldsymbol{\epsilon}$ es ortogonal a cada columna de la matriz $\boldsymbol{X}$.
   (b) El vector de perturbaciones tiene media $\mathbf{0}$.
   (c) El vector de perturbaciones está incorrelado con las $\boldsymbol{y}$.
   (d) Nada de lo anterior.

9. ¿Cuáles de las siguientes son matrices de proyección?
   (a) $\boldsymbol{X}'\boldsymbol{X}$.
   (b) $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}X'$.
   (c) $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$.
   (d) Nada de lo anterior.

10. Geométricamente, SSE puede interpretarse como:
   (a) El coseno que forma el vector $\hat{\boldsymbol{\epsilon}}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.
   (b) La norma al cuadrado del vector $\hat{\boldsymbol{\epsilon}}$.
   (c) La norma al cuadrado del vector $\boldsymbol{y} - \overline{\boldsymbol{y}}$.
   (d) Nada de lo anterior.

11. Considera el modelo lineal $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ con los supuestos habituales. Sea $\hat{\boldsymbol{\beta}}$ el estimador MCO y $\hat{\boldsymbol{\beta}}^*$ el estimador obtenido imponiendo la condición $\boldsymbol{A\beta} = \boldsymbol{c}$. Sean SSE y SSE* las sumas de cuadrados de los residuos correspondientes. Se verifica en todo caso que:

    (a) SSE $\geq$ SSE*.

    (b) SSE $\leq$ SSE*.

    (c) SSE $\geq$ SSE* si los parámetros $\boldsymbol{\beta}$ realmente verifican $\boldsymbol{A\beta} = \boldsymbol{c}$

    (d) Nada de lo anterior.

12. Geométricamente, cuando entre las columnas de $\boldsymbol{X}$ está la de "unos", R^2 puede interpretarse como:

    (a) El coseno que forma el vector $\hat{\boldsymbol{\epsilon}}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.

    (b) La norma al cuadrado del vector $\hat{\boldsymbol{\epsilon}}$.

    (c) El coseno que forma el vector $\boldsymbol{y}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.

    (d) Nada de lo anterior.

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las instrucciones

```
library(MASS)
data(UScrime)
mod <- lm(y ~ ., data=UScrime)
summary(mod)
```

generan lo siguiente:

```
Call:
lm(formula = y ~ ., data = UScrime)

Residuals:
     Min       1Q   Median       3Q      Max
-395.738  -98.088   -6.695  112.989  512.671

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5984.2876  1628.3184  -3.675 0.000893 ***
M               8.7830     4.1714   2.106 0.043443 *
So             -3.8035   148.7551  -0.026 0.979765
Ed             18.8324     6.2088   3.033 0.004861 **
Po1            19.2804    10.6110   1.817 0.078892 .
Po2           -10.9422    11.7478  -0.931 0.358830
LF             -0.6638     1.4697  -0.452 0.654654
M.F             1.7407     2.0354   0.855 0.398995
Pop            -0.7330     1.2896  -0.568 0.573845
NW              0.4204     0.6481   0.649 0.521279
U1             -5.8271     4.2103  -1.384 0.176238
U2             16.7800     8.2336   2.038 0.050161 .
GDP             0.9617     1.0367   0.928 0.360754
Ineq            7.0672     2.2717   3.111 0.003983 **
Prob        -4855.2658  2272.3746  -2.137 0.040627 *
Time           -3.4790     7.1653  -0.486 0.630708
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Residual standard error: 209.1 on 31 degrees of freedom
Multiple R-Squared: 0.8031,Adjusted R-squared: 0.7078
F-statistic: 8.429 on 15 and 31 DF,  p-value: 3.539e-07
```

En relación con la salida anterior, responde a las preguntas a continuación, hasta el fin del bloque.

13. Dejando de lado $\hat{\beta}_0$ (`Intercept`), la variable con mayor coeficiente estimado en valor absoluto es `Prob`. Ello permite afirmar que:

    (a) Será la variable que entrará en primer lugar en un algoritmo *forward stepwise* (redgresión escalonada hacia adelante).

    (b) Es la variable que explica una mayor porción de suma de cuadrados.

    (c) Es la variable más significativa.

    (d) Nada de lo anterior.

14. Supongamos que se verifican los supuestos habituales más normalidad en las perturbaciones. El hecho de que la variable `Ineq` tenga asociado un *p-value* de 0.03983 significa que:

    (a) La hipótesis $H_0 : \beta_{\texttt{Ineq}} = 0$ sería rechazada con un nivel de significación $\alpha = 0{,}05$.

    (b) La probabilidad de que $H_0 : \beta_{\texttt{Ineq}} = 0$ sea cierta es 0.03983.

    (c) Suponiendo que $H_0 : \beta_{\texttt{Ineq}} = 0$ fuera cierta, el valor observado del t-ratio se sitúa entre el 3.983 % más "raro" de los que podríamos observar.

    (d) Nada de lo anterior.

15. La salida anterior permite deducir que el número de observaciones con que se ha contado para realizar la estimación es de:

    (a) 46

    (b) 15

    (c) 31

    (d) Nada de lo anterior.

16. La salida anterior permite deducir que hay multicolinealidad acusada, habida cuenta de que bastantes estadísticos t son significativos, y la regresión en su conjunto, evaluada con ayuda del estadístico R^2, carece de capacidad explicativa de la variable respuesta.

    (a) Falso

    (b) Cierto

17. La salida anterior permite deducir que hay bastantes parámetros estimados no significativos al nivel $\alpha = 0{,}05$. La inclusión, pese a todo, de los regresores correspondientes en el modelo:

    (a) Disminuye los grados de libertad.

    (b) Tiene el potencial de sesgar las estimaciones de los demás parámetros.

    (c) Necesariamente disminuye el valor de la estimación de σ^2_{ϵ}.

    (d) Nada de lo anterior.

18. La salida anterior permite deducir que SSE toma el valor aproximado:

    (a) 43723

    (b) 2011249

    (c) 1355407

    (d) 345234

19. La salida anterior permite deducir que SST toma el valor aproximado:

    (a) 6883733

    (b) 124031

    (c) 94727131

    (d) 1130234

20. Supongamos ahora que queremos contrastar que los coeficientes de `U1` y `U2` son simultáneamente cero. Realizamos una nueva regresión así:

```
> mod2 <-  lm(y ~ . - U1 - U2,data=UScrime)
> summary(mod2)

Call:
lm(formula = y ~ . - U1 - U2, data = UScrime)

Residuals:
     Min       1Q   Median       3Q      Max
-399.449  -91.328   -7.218  129.387  524.928

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5281.6045  1624.7957  -3.251  0.00265 **
M               6.8797     3.9869   1.726  0.09378 .
So             49.0098   141.2059   0.347  0.73073
Ed             14.9124     6.1061   2.442  0.02012 *
Po1            21.4693    10.8448   1.980  0.05613 .
Po2           -12.0878    11.9179  -1.014  0.31784
LF             -0.7848     1.1854  -0.662  0.51253
M.F             1.3747     1.6208   0.848  0.40248
Pop            -0.7415     1.2946  -0.573  0.57067
NW              0.4412     0.6443   0.685  0.49828
GDP             1.3150     1.0558   1.246  0.22171
Ineq            7.5103     2.3369   3.214  0.00292 **
Prob        -4949.5291  2342.5358  -2.113  0.04226 *
Time           -2.1380     7.3364  -0.291  0.77256
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Residual standard error: 216 on 33 degrees of freedom
Multiple R-Squared: 0.7761,Adjusted R-squared: 0.688
F-statistic: 8.801 on 13 and 33 DF,  p-value: 2.335e-07
```

    Con ayuda de los resultados de esta regresión y de la precedente, podemos fácilmente calcular que el valor de estadístico Q_h para el contraste de $H_0 : \beta_{\texttt{U1}} = \beta_{\texttt{U2}} = 0$ es aproximadamente:

    (a) 2.1069

    (b) 1.0324

    (c) 8.3415

    (d) 6.1715

21. Comparando las dos regresiones anteriores, constatamos que al eliminar los dos regresores `U1` y `U2`, R^2

    (a) R^2 no ha aumentado; necesariamente tenía que ocurrir esto.

    (b) $\overline{R}^2$ (`Adjusted R-Squared` en la salida precedente) no ha aumentado; necesariamente tenía que ocurrir esto.

    (c) SSE no ha aumentado; necesariamente tenía que ocurrir esto.

    (d) Nada de lo anterior.

FIN DEL BLOQUE DE PREGUNTAS

22. Cuando hay multicolinealidad exacta,

    (a) La proyección del vector $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ es única, pero no hay estimación única de los $\boldsymbol{\beta}$.

    (b) La proyección del vector $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ no es única, y por tanto no hay estimación única de los $\boldsymbol{\beta}$.

    (c) La suma de cuadrados de los residuos no está bien definida, y por tanto no tiene sentido hacer estimación mínimo-cuadrática ordinaria.

    (d) Nada de lo anterior.

23. Sea R_i^2 el R^2 resultante de regresar la columna X_i de la matriz $\boldsymbol{X}$ sobre las restantes. Entonces,

$$\frac{1}{1-R_i^2}$$

    es un estadístico. . .

    (a) . . .conocido como VIF(i), que permite diagnosticar la presencia de multicolinealidad aproximada en la matriz de diseño $\boldsymbol{X}$.

    (b) . . .cuyos valores “grandes” son indicativos de multicolnealidad aproximada.

    (c) . . .cuyos valores “pequeños” son indicativos de multicolnealidad aproximada.

    (d) Nada de lo anterior.

24. Consideremos el estimador *ridge*

$$\hat{\boldsymbol{\beta}}^{(k)} = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

    Al aumentar el valor k:

    (a) Aumento el sesgo de$\hat{\boldsymbol{\beta}}^{(k)}$.

    (b) Disminuye el sesgo de$\hat{\boldsymbol{\beta}}^{(k)}$.

    (c) Aumentan las varianzas de$\hat{\boldsymbol{\beta}}^{(k)}$.

    (d) Disminuyen las varianzas de $\hat{\boldsymbol{\beta}}^{(k)}$.

    (e) Nada de lo anterior.

25. Al hacer estimación mínimo-cuadrica condicionada por $A\boldsymbol{\beta} = \boldsymbol{c}$, en el caso de que los verdaderos parámetros no veriquen la restricción impuesta, hay riesgo de:

    (a) Incrementar las varianzas de los estimadores $\hat{\boldsymbol{\beta}}$.

    (b) Reducir los grados de libertad.

    (c) Sesgar las estimaciones de los parámetros $\boldsymbol{\beta}$.

    (d) Nada de lo anterior.

26. Un estadístico como

$$\frac{N-p}{p-1} \times \frac{R^2}{1-R^2}$$

    permitiría contratar la hipótesis de que:

    (a) La dependencia de $\boldsymbol{y}$ sobre X es lineal.

    (b) La varianza de la perturbación es nula.

    (c) Todos los parámetros, salvo β_0 son nulos.

    (d) Nada de lo anterior.

27. Construimos intervalos de confianza para cinco parámetros $\beta_1, \ldots, \beta_5$ de forma que cada intervalo cubre el respectivo parámetro con confianza 0.90. La confianza con que los cinco intervalos cubren simultáneamente los cinco valores de los parámetros es:

    (a) Mayor o igual que 0.85

    (b) Menor o igual que 0.35

    (c) Mayor o igual que 0.75

    (d) Nada de lo anterior.

28. Una observación muy influyente, tendrá:

    (a) Un residuo MCO grande.

    (b) Un borrado pequeño.

    (c) Un residuo borrado grande.

    (d) Un VIF(i) grande.

29. En el análisis tanto de residuos como de influencia aparece unos términos denotados por p_{ii}. Se trata de:

    (a) Los elementos diagonales de $\boldsymbol{X}'\boldsymbol{X}$.

    (b) Los elementos diagonales de $\boldsymbol{X}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}'$.

    (c) Los valores propios de $\boldsymbol{X}'\boldsymbol{X}$

    (d) nada de lo anterior.

30. El método de regresion en componentes principales se emplea preferentemente en situaciones en que hay:

    (a) Muy pocas observaciones.

    (b) Fuerte multicolinealidad.

    (c) Posibilidad de dependencia no lineal de la respuesta sobre los regresores.

    (d) Observaciones influyentes.

31. Sea un modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, en que la matriz $\boldsymbol{X}$ es $N \times p$. La estimación por mínimos cuadrados ordinarios y mediante componentes principales, tomando un número $k < p$ de componentes, proporcionaran:

    (a) El mismo R^2.

    (b) Estimadores en ambos casos insesgados de los parámetros.

    (c) Estimadores en ambos casos sesgados de los parámetros.

    (d) Nada de lo anterior.

32. El estadístico C_p permite seleccionar modelos. Escogeremos preferentemente aquéllos que tengan:

    (a) Un valor pequeño de C_p.

    (b) Un valor grande de C_p.

    (c) Un C_p cercano a 1.

    (d) Nada de lo anterior.

33. En un modelo ANOVA con dos tratamientos, el diseño equilibrado (= mismo número de observaciones para cada combinación de niveles de los dos tratamientos) proporciona:

    (a) Ortogonalidad por bloques.

    (b) Homoscedasticidad de las perturbaciones.

    (c) Aditividad del modelo.

    (d) Nada de lo anterior.

34. Al ajustar un modelo de regresión con N observaciones, tenemos la sospecha de que una concreta es anómala. Para contrastar dicha hipótesis, podríamos tomar su residuo externamente *studentizado* y compararlo con el cuantil adecuado de una distribución:

    (a) Máximo de N variables t de Student con grados de libertad apropiados.

    (b) Beta con grados de libertad adecuados.

    (c) t de Student y grados de libertad adecuados.

    (d) Nada de lo anterior.

35. Al ajustar un modelo de regresión con N observaciones, el mayor de los residuos externamente *studentizados* nos parece sospechosamente grande. Para valorar si corresponde a una observación anómala, lo compararíamos con el cuantil adecuado de una distribución:

    (a) Máximo de N variables t de Student con grados de libertad apropiados.

    (b) Beta con grados de libertad adecuados.

    (c) t de Student y grados de libertad adecuados.

    (d) Nada de lo anterior.

Respuestas al examen de tipo $\boxed{\text{A}}$

Sección 1. Instrucciones

1. Salvo que se indique lo contrario, las preguntas bien contestadas valen un punto. Puede haber más de una respuesta correcta, y para obtener puntuación has de señalarlas todas. Preguntas que no estén bien contestadas puntúan -0.5 veces su valor.

2. Intento medir conocimientos y no agudeza visual. Inevitablemente, en un examen de este tipo hay que prestar mucha atención. Cada curso hay personas que echan a perder una nota potencialmente buena por responder temeraria o atolondradamente.

   ¡Por favor, fíjate bien en todos los detalles!

   Te ayudará proceder por exclusión de absurdos. Si una pregunta te parece ambigua, anota brevemente la razón al margen y no la contestes.

Sección 2. Cuestiones de elección múltiple

3. Si A es cuadrada, simétrica, con valores propios $\lambda_1, \ldots, \lambda_p$, entonces:
   - **(a)** $\mathrm{tr}(A) = \sum_{i=1}^{p} \lambda_i$**.**
   - (b) $\mathrm{tr}(A) = \prod_{i=1}^{p} \lambda_i$.
   - (c) $\mathrm{tr}(A) \geq 0$.
   - (d) Nada de lo anterior.

4. Los valores propios de una matriz simétrica son siempre...
   - (a) Positivos.
   - **(b) Reales.**
   - (c) Simétricos.
   - (d) Nada de lo anterior.

5. Una matriz simétrica, semidefinida positiva e idempotente tiene valores propios que son siempre...
   - (a) Cero.
   - **(b) Cero o uno.**
   - (c) Uno o menos uno.
   - (d) Nada de lo anterior.

6. Sea $\boldsymbol{X}$ la matrizde diseño $N \times p$ de un modelo lineal $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$. La condición de que rango$(\boldsymbol{X}) = p < N$:
   - **(a) Excluye la posibilidad de multicolinealidad exacta.**
   - **(b) Implica que el número de grados de libertad en la estimación será mayor que cero.**
   - (c) Garantiza la unicidad de la proyección de $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$.
   - (d) Nada de lo anterior.

7. Las ecuaciones normales $\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}$:
   - (a) Siempre tienen una solución única.
   - **(b) Siempre tienen al menos una solución.**
   - (c) Siempre tienen múltiples soluciones; de hecho, infinitas.
   - (d) Nada de lo anterior.

8. Entre los supuestos habituales en el modelo lineal ordinario está que:
   - (a) El vector de perturbaciones $\boldsymbol{\epsilon}$ es ortogonal a cada columna de la matriz $\boldsymbol{X}$.
   - **(b) El vector de perturbaciones tiene media 0.**
   - (c) El vector de perturbaciones está incorrelado con las $\boldsymbol{y}$.
   - (d) Nada de lo anterior.

9. ¿Cuáles de las siguientes son matrices de proyección?
   - (a) $\boldsymbol{X}'\boldsymbol{X}$.
   - **(b)** $\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}X'$**.**
   - (c) $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$.
   - (d) Nada de lo anterior.

10. Geométricamente, SSE puede interpretarse como:
    - (a) El coseno que forma el vector $\hat{\boldsymbol{\epsilon}}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.
    - **(b) La norma al cuadrado del vector $\hat{\boldsymbol{\epsilon}}$.**
    - (c) La norma al cuadrado del vector $\boldsymbol{y} - \overline{\boldsymbol{y}}$.
    - (d) Nada de lo anterior.

11. Considera el modelo lineal $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}$ con los supuestos habituales. Sea $\hat{\boldsymbol{\beta}}$ el estimador MCO y $\hat{\boldsymbol{\beta}}^*$ el estimador obtenido imponiendo la condición $\boldsymbol{A\beta} = \boldsymbol{c}$. Sean SSE y SSE* las sumas de cuadrados de los residuos correspondientes. Se verifica en todo caso que:

    (a) SSE $\geq$ SSE*.

    (b) SSE $\leq$ SSE*.

    (c) SSE $\geq$ SSE* si los parámetros $\boldsymbol{\beta}$ realmente verifican $\boldsymbol{A\beta} = \boldsymbol{c}$

    (d) Nada de lo anterior.

12. Geométricamente, cuando entre las columnas de $\boldsymbol{X}$ está la de "unos", R^2 puede interpretarse como:

    (a) El coseno que forma el vector $\hat{\boldsymbol{\epsilon}}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.

    (b) La norma al cuadrado del vector $\hat{\boldsymbol{\epsilon}}$.

    (c) El coseno que forma el vector $\boldsymbol{y}$ con el hiperplano generado por las columnas de $\boldsymbol{X}$.

    (d) Nada de lo anterior.

COMIENZO DE UN BLOQUE DE PREGUNTAS

Las instrucciones

```
library(MASS)
data(UScrime)
mod <- lm(y ~ ., data=UScrime)
summary(mod)
```

generan lo siguiente:

```
Call:
lm(formula = y ~ ., data = UScrime)

Residuals:
     Min       1Q   Median       3Q      Max
-395.738  -98.088   -6.695  112.989  512.671

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5984.2876  1628.3184  -3.675 0.000893 ***
M               8.7830     4.1714   2.106 0.043443 *
So             -3.8035   148.7551  -0.026 0.979765
Ed             18.8324     6.2088   3.033 0.004861 **
Po1            19.2804    10.6110   1.817 0.078892 .
Po2           -10.9422    11.7478  -0.931 0.358830
LF             -0.6638     1.4697  -0.452 0.654654
M.F             1.7407     2.0354   0.855 0.398995
Pop            -0.7330     1.2896  -0.568 0.573845
NW              0.4204     0.6481   0.649 0.521279
U1             -5.8271     4.2103  -1.384 0.176238
U2             16.7800     8.2336   2.038 0.050161 .
GDP             0.9617     1.0367   0.928 0.360754
Ineq            7.0672     2.2717   3.111 0.003983 **
Prob        -4855.2658  2272.3746  -2.137 0.040627 *
Time           -3.4790     7.1653  -0.486 0.630708
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Residual standard error: 209.1 on 31 degrees of freedom
Multiple R-Squared: 0.8031,Adjusted R-squared: 0.7078
F-statistic: 8.429 on 15 and 31 DF,  p-value: 3.539e-07
```

En relación con la salida anterior, responde a las preguntas a continuación, hasta el fin del bloque.

13. Dejando de lado $\hat{\beta}_0$ (`Intercept`), la variable con mayor coeficiente estimado en valor absoluto es `Prob`. Ello permite afirmar que:

    (a) Será la variable que entrará en primer lugar en un algoritmo *forward stepwise* (redgresión escalonada hacia adelante).

    (b) Es la variable que explica una mayor porción de suma de cuadrados.

    (c) Es la variable más significativa.

    (d) Nada de lo anterior.

14. Supongamos que se verifican los supuestos habituales más normalidad en las perturbaciones. El hecho de que la variable `Ineq` tenga asociado un *p-value* de 0.03983 significa que:

    (a) La hipótesis $H_0 : \beta_{\texttt{Ineq}} = 0$ sería rechazada con un nivel de significación $\alpha = 0{,}05$.

    (b) La probabilidad de que $H_0 : \beta_{\texttt{Ineq}} = 0$ sea cierta es 0.03983.

    (c) Suponiendo que $H_0 : \beta_{\texttt{Ineq}} = 0$ fuera cierta, el valor observado del t-ratio se sitúa entre el 3.983 % más "raro" de los que podríamos observar.

    (d) Nada de lo anterior.

15. La salida anterior permite deducir que el número de observaciones con que se ha contado para realizar la estimación es de:

    (a) 46

    (b) 15

    (c) 31

    (d) Nada de lo anterior.

16. La salida anterior permite deducir que hay multicolinealidad acusada, habida cuenta de que bastantes estadísticos t son significativos, y la regresión en su conjunto, evaluada con ayuda del estadístico R^2, carece de capacidad explicativa de la variable respuesta.

    (a) Falso

    (b) Cierto

17. La salida anterior permite deducir que hay bastantes parámetros estimados no significativos al nivel $\alpha = 0{,}05$. La inclusión, pese a todo, de los regresores correspondientes en el modelo:

   (a) **Disminuye los grados de libertad.**

   (b) Tiene el potencial de sesgar las estimaciones de los demás parámetros.

   (c) Necesariamente disminuye el valor de la estimación de σ^2_{ϵ}.

   (d) Nada de lo anterior.

18. La salida anterior permite deducir que SSE toma el valor aproximado:

   (a) 43723

   (b) 2011249

   (c) **1355407**

   (d) 345234

19. La salida anterior permite deducir que SST toma el valor aproximado:

   (a) **6883733**

   (b) 124031

   (c) 94727131

   (d) 1130234

20. Supongamos ahora que queremos contrastar que los coeficientes de `U1` y `U2` son simultáneamente cero. Realizamos una nueva regresión así:

```
> mod2 <-  lm(y ~ . - U1 - U2,data=UScrime)
> summary(mod2)

Call:
lm(formula = y ~ . - U1 - U2, data = UScrime)

Residuals:
     Min       1Q   Median       3Q      Max
-399.449  -91.328   -7.218  129.387  524.928

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -5281.6045  1624.7957  -3.251  0.00265 **
M               6.8797     3.9869   1.726  0.09378 .
So             49.0098   141.2059   0.347  0.73073
Ed             14.9124     6.1061   2.442  0.02012 *
Po1            21.4693    10.8448   1.980  0.05613 .
Po2           -12.0878    11.9179  -1.014  0.31784
LF             -0.7848     1.1854  -0.662  0.51253
M.F             1.3747     1.6208   0.848  0.40248
Pop            -0.7415     1.2946  -0.573  0.57067
NW              0.4412     0.6443   0.685  0.49828
GDP             1.3150     1.0558   1.246  0.22171
Ineq            7.5103     2.3369   3.214  0.00292 **
Prob        -4949.5291  2342.5358  -2.113  0.04226 *
Time           -2.1380     7.3364  -0.291  0.77256
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Residual standard error: 216 on 33 degrees of freedom
Multiple R-Squared: 0.7761,Adjusted R-squared: 0.688
F-statistic: 8.801 on 13 and 33 DF,  p-value: 2.335e-07
```

   Con ayuda de los resultados de esta regresión y de la precedente, podemos fácilmente calcular que el valor de estadístico Q_h para el contraste de $H_0 : \beta_{\mathtt{U1}} = \beta_{\mathtt{U2}} = 0$ es aproximadamente:

   (a) **2.1069**

   (b) 1.0324

   (c) 8.3415

   (d) 6.1715

21. Comparando las dos regresiones anteriores, constatamos que al eliminar los dos regresores `U1` y `U2`, R^2

   (a) **R^2 no ha aumentado; necesariamente tenía que ocurrir esto.**

   (b) $\overline{R}^2$ (`Adjusted R-Squared` en la salida precedente) no ha aumentado; necesariamente tenía que ocurrir esto.

   (c) **SSE no ha aumentado; necesariamente tenía que ocurrir esto.**

   (d) Nada de lo anterior.

FIN DEL BLOQUE DE PREGUNTAS

---

22. Cuando hay multicolinealidad exacta,

    (a) **La proyección del vector $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ es única, pero no hay estimación única de los β.**

    (b) La proyección del vector $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ no es única, y por tanto no hay estimación única de los $\boldsymbol{\beta}$.

    (c) La suma de cuadrados de los residuos no está bien definida, y por tanto no tiene sentido hacer estimación mínimo-cuadrática ordinaria.

    (d) Nada de lo anterior.

23. Sea R_i^2 el R^2 resultante de regresar la columna X_i de la matriz $\boldsymbol{X}$ sobre las restantes. Entonces,

$$\frac{1}{1 - R_i^2}$$

    es un estadístico. . .

    (a) **. . .conocido como VIF(i), que permite diagnosticar la presencia de multicolinealidad aproximada en la matriz de diseño $\boldsymbol{X}$.**

    (b) **. . .cuyos valores "grandes" son indicativos de multicolnealidad aproximada.**

    (c) . . .cuyos valores "pequeños" son indicativos de multicolnealidad aproximada.

    (d) Nada de lo anterior.

24. Consideremos el estimador *ridge*

$$\hat{\boldsymbol{\beta}}^{(k)} = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}.$$

    Al aumentar el valor k:

    (a) **Aumento el sesgo de$\hat{\boldsymbol{\beta}}^{(k)}$.**

    (b) Disminuye el sesgo de$\hat{\boldsymbol{\beta}}^{(k)}$.

    (c) Aumentan las varianzas de$\hat{\boldsymbol{\beta}}^{(k)}$.

    (d) **Disminuyen las varianzas de $\hat{\boldsymbol{\beta}}^{(k)}$.**

    (e) Nada de lo anterior.

25. Al hacer estimación mínimo-cuadrica condicionada por $A\boldsymbol{\beta} = \boldsymbol{c}$, en el caso de que los verdaderos parámetros no veriquen la restricción impuesta, hay riesgo de:

    (a) Incrementar las varianzas de los estimadores $\hat{\boldsymbol{\beta}}$.

    (b) Reducir los grados de libertad.

    (c) **Sesgar las estimaciones de los parámetros β.**

    (d) Nada de lo anterior.

26. Un estadístico como

$$\frac{N-p}{p-1} \times \frac{R^2}{1-R^2}$$

    permitiría contratar la hipótesis de que:

    (a) La dependencia de $\boldsymbol{y}$ sobre X es lineal.

    (b) La varianza de la perturbación es nula.

    (c) **Todos los parámetros, salvo β_0 son nulos.**

    (d) Nada de lo anterior.

27. Construimos intervalos de confianza para cinco parámetros $\beta_1, \ldots, \beta_5$ de forma que cada intervalo cubre el respectivo parámetro con confianza 0.90. La confianza con que los cinco intervalos cubren simultáneamente los cinco valores de los parámetros es:

    (a) Mayor o igual que 0.85

    (b) Menor o igual que 0.35

    (c) **Mayor o igual que 0.75**

    (d) Nada de lo anterior.

28. Una observación muy influyente, tendrá:

    (a) Un residuo MCO grande.

    (b) Un borrado pequeño.

    (c) **Un residuo borrado grande.**

    (d) Un VIF(i) grande.

29. En el análisis tanto de residuos como de influencia aparece unos términos denotados por p_{ii}. Se trata de:

    (a) Los elementos diagonales de $\boldsymbol{X}'\boldsymbol{X}$.

    (b) Los elementos diagonales de $\boldsymbol{X}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{X}'$.

    (c) Los valores propios de $\boldsymbol{X}'\boldsymbol{X}$

    (d) nada de lo anterior.

30. El método de regresion en componentes principales se emplea preferentemente en situaciones en que hay:

   (a) Muy pocas observaciones.

   (b) Fuerte multicolinealidad.

   (c) Posibilidad de dependencia no lineal de la respuesta sobre los regresores.

   (d) Observaciones influyentes.

31. Sea un modelo $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, en que la matriz $\boldsymbol{X}$ es $N \times p$. La estimación por mínimos cuadrados ordinarios y mediante componentes principales, tomando un número $k < p$ de componentes, proporcionaran:

   (a) El mismo R^2.

   (b) Estimadores en ambos casos insesgados de los parámetros.

   (c) Estimadores en ambos casos sesgados de los parámetros.

   (d) Nada de lo anterior.

32. El estadístico C_p permite seleccionar modelos. Escogeremos preferentemente aquéllos que tengan:

   (a) Un valor pequeño de C_p.

   (b) Un valor grande de C_p.

   (c) Un C_p cercano a 1.

   (d) Nada de lo anterior.

33. En un modelo ANOVA con dos tratamientos, el diseño equilibrado (= mismo número de observaciones para cada combinación de niveles de los dos tratamientos) proporciona:

   (a) Ortogonalidad por bloques.

   (b) Homoscedasticidad de las perturbaciones.

   (c) Aditividad del modelo.

   (d) Nada de lo anterior.

34. Al ajustar un modelo de regresión con N observaciones, tenemos la sospecha de que una concreta es anómala. Para contrastar dicha hipótesis, podríamos tomar su residuo externamente *studentizado* y compararlo con el cuantil adecuado de una distribución:

   (a) Máximo de N variables t de Student con grados de libertad apropiados.

   (b) Beta con grados de libertad adecuados.

   (c) t de Student y grados de libertad adecuados.

   (d) Nada de lo anterior.

35. Al ajustar un modelo de regresión con N observaciones, el mayor de los residuos externamente *studentizados* nos parece sospechosamente grande. Para valorar si corresponde a una observación anómala, lo compararíamos con el cuantil adecuado de una distribución:

   (a) Máximo de N variables t de Student con grados de libertad apropiados.

   (b) Beta con grados de libertad adecuados.

   (c) t de Student y grados de libertad adecuados.

   (d) Nada de lo anterior.