Estadística: Modelos Lineales

Final Enero 2.003, Tipo: $\boxed{\text{A}}$

Apellidos y Nombre: ______________________

DNI: ______________________

Grupo: ______________________

Profesor : ______________________

Sección 1. Instrucciones

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.

Intento medir conocimientos y no agudeza visual. Inevitablemente, en un examen de este tipo hay que prestar mucha atención. Cada curso hay más de una persona que echa 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.

1. Ajustas el modelo $Y = \beta_0 + \beta_1 X_1 + \epsilon$ y obtienes: $\sigma^2 = 0{,}45$,

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{pmatrix} 0{,}83 & 0{,}34 \\ 0{,}34 & 0{,}04 \end{pmatrix}$$

y $\hat{\boldsymbol{\beta}}' = (0{,}15 \quad 2{,}14)$. Te piden ahora que realices una predicción del valor esperable de Y cuando $X_1 = 4{,}2$ y de su varianza. Tu estimación será aproximadamente:

(a) $\hat{Y} = 9{,}1380$ y $\hat{\sigma}_{\hat{Y}} = 8{,}6972$.

(b) $\hat{Y} = 6{,}2245$ y $\hat{\sigma}_{\hat{Y}} = 5{,}3425$.

(c) $\hat{Y} = 3{,}2347$ y $\hat{\sigma}_{\hat{Y}} = 2{,}3419$.

(d) $\hat{Y} = 7{,}3485$ y $\hat{\sigma}_{\hat{Y}} = 6{,}3491$.

2. El vector de residuos $\hat{\boldsymbol{\epsilon}}$:

(a) Es ortogonal al vector $\boldsymbol{y}$.

(b) Es ortogonal a todos los regresores.

(c) Es ortogonal a la columna de “unos”, si está presente.

(d) Verifica *siempre* (haya o no columna de “unos”) que $\sum_{i=1}^{N} \hat{\epsilon}_i = 0$

3. En un problema de estimación condicionada,

(a) SSE_h (la suma de cuadrados en el modelo estimado con la condición) nunca será menor que SSE del modelo sin la condición, y esto tanto si ésta es “correcta” como si no.

(b) Las estimaciones de los parámetros no afectados por la condición siempre serán insesgadas.

(c) Los estimadores tendrán menor varianza que los estimadores MCO si, y solo si, la condición impuesta es “correcta” (verificada por los parámetros del modelo).

(d) Todo falso.

4. La igualdad $SST = SSR + SSE$ de la que se deduce, entre otras cosas, que $0 \leq R^2 \leq 1$ está garantizada si hay:

(a) Normalidad de las perturbaciones.

(b) Columna de “unos” entre los regresores.

(c) Homoscedasticidad de las perturbaciones.

(d) Un número suficiente de grados de libertad; de otro modo, SSE podría ser cero y la igualdad no se verificaría.

5. Si tenemos tres factores (o tratamientos), cada uno con tres niveles, y tenemos la certeza de que no hay interacción, el número de experimentos a realizar si adoptamos un diseño en cuadrado latino asciende a:

(a) 27

(b) 9

(c) 12

(d) Todo falso.

6. Si tenemos tres factores (o tratamientos), cada uno con tres niveles, y queremos replicar 4 veces, el número de experimentos a realizar asciende a:

(a) 108

(b) 27

(c) 36

(d) Todo falso.

7. Podemos estimar un modelo de regresión logística:

   (a) Por máxima verosimilitud.

   (b) Resolviendo un sistema lineal análogo a las ecuaciones normales.

   (c) Todo falso.

   (d) Linealizando el modelo, a base de tomar logaritmos y estimar el modelo lineal resultante.

8. El teorema conocido como de Gauss-Markov asegura:

   (a) Que de entre todos los estimadores lineales e insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.

   (b) Que de entre todos los estimadores lineales e insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático, con la única excepción del estimador *ridge*.

   (c) Que de entre todos los estimadores insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.

   (d) Que de entre todos los estimadores lineales de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.

9. ¿Cuál de los siguientes supuestos es imprescindible para que la proyección de $\boldsymbol{y}$ sobre el subespacio generado por las columnas de $\boldsymbol{X}$ sea única?

   (a) Normalidad en las perturbaciones.

   (b) $\boldsymbol{X}$ de rango completo.

   (c) $\boldsymbol{X}$ de dimensiones $N \times p$ con $N > p$.

   (d) Ninguno de los anteriores: la proyección de $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ existe y es única siempre.

10. En un modelo de regresión lineal, la omisión de regresores que deberían formar parte de la especificación:

    (a) Afecta (sesgándola) a la estimación de β_0, pero no a la de los demás parámetros presentes en el modelo.

    (b) Sesga, en general, los estimadores de los parámetros incluidos.

    (c) Puede no afectar al sesgo de los estimadores de los parámetros incluidos, si acontece que los regresores omitidos son ortogonales a los incluidos.

    (d) Disminuye, en general, las varianzas de los estimadores, respecto a las que hubieran sido en el caso de ajustar el modelo completo.

11. Cuando omitimos la columna de "unos" en la matriz de diseño, el hiperplano ajustado:

    (a) Proporciona mejor ajuste a la $\boldsymbol{y}$ (en términos de SSE) que incluyendo la columna de unos.

    (b) Pasa *siempre* por el origen de coordenadas.

    (c) Necesariamente tiene pendiente positiva.

    (d) Pasa a una distancia del origen de coordenadas siempre igual a $||\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}||$.

12. Llamamos *restricciones de identificación* a:

    (a) Ecuaciones que, suplementando a las ecuaciones normales, permiten obtener una solución única para los $\hat{\boldsymbol{\beta}}$.

    (b) Restricciones que podemos usar en sustitución de un DNI o pasaporte cuando viajamos fuera de la Unión Europea.

    (c) Ecuaciones que, junto con las ecuaciones normales, dan un sistema compatible indeterminado.

    (d) Ecuaciones no lineales, que sustituyen a las normales y pueden resolverse mediante técnicas iterativas.

13. Al estimar un modelo lineal mediante regresión *ridge*, es decir, $\hat{\boldsymbol{\beta}}_{(\boldsymbol{k})} = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}\ldots$

    (a) Sólamente es posible mejorar el ECM (error cuadrático medio) en los casos en que la multicolinealidad es perfecta.

    (b) Un incremento de k acarrea un incremento del sesgo del estimador.

    (c) Un incremento de k acarrea una disminución de la varianza del estimador.

    (d) Un incremento de k acarrea una disminución del error cuadrático medio.

    (e) Siempre hay un k que hace que el estimador *ridge* sea, en términos de ECM, mejor que el mínimo cuadrático ordinario.

14. En un modelo de Análisis de Varianza cruzado completo, el equilibrio junto con las restricciones de identificación garantiza:

    (a) La ortogonalidad por bloques: las estimaciones de los parámetros correspondientes a un efecto no se alteran con la inclusión u omisión de otros efectos.

    (b) La ortogonalidad del vector de residuos a los regresores.

    (c) La homoscedasticidad de la perturbación.

    (d) La normalidad de los estimadores.

15. ¿En qué condiciones puede $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ (proyección de $\boldsymbol{y}$ sobre el subespacio que generan las columnas de $\boldsymbol{X}$) no ser única?

    (a) Cuando haya acusada multicolinealidad entre las columnas de $\boldsymbol{X}$.

    (b) Siempre es única.

    (c) No tiene por qué ser única en ningún caso.

    (d) Todo falso.

16. Los "residuos borrados" *(deleted residuals)* $d_i\ldots$

    (a) ... se definen por $d_i = \hat{\epsilon}_i/\sqrt{\hat{\sigma}^2(1-p_{ii})}$

    (b) ... se definen por $d_i = \hat{\epsilon}_i/\sqrt{\hat{\sigma}^2_{(i)}(1-p_{ii})}$

    (c) ...permiten emplear el criterio de validación cruzada para seleccionar un modelo sin necesidad de realizar mas que una regresión.

    (d) ... se definen por $d_i = \hat{\epsilon}_i/(\hat{\sigma}^2(1-p_{ii}))$

17. En una regresión logística expresamos como función lineal de los predictores:

    (a) Una variable que toma valores 0 y 1.

    (b) El logaritmo de una razón de probabilidades condicionadas.

    (c) Una razón de probabilidades.

    (d) La media de la variable respuesta, que coincide con P_i, la probabilidad de que tome el valor 1.

    (e) Todo falso.

18. ¿Cuando es inadecuada la inclusión de una columna de "unos" entre los regresores?

    (a) Cuando el número de grados de libertad es relativamente elevado.

    (b) Cuando las escalas de los regresores sean muy similares.

    (c) Cuando tanto regresando como regresores se toman en desviaciones respecto a su media. En tal caso, β_0 sería siempre 0.

    (d) Cuando las escalas de los regresores sean muy diferentes.

19. A la luz de la discusión hecha en clase sobre los efectos derivados de incluir regresores irrelevantes en una ecuación de regresión, cabría concluir que el efecto pernicioso de "pasarse" se aminora cuando:

    (a) El tamaño de la muestra crece.

    (b) El número de regresores crece.

    (c) Los regresores añadidos son combinación lineal exacta unos de otros.

    (d) Todo falso.

20. Cuando tomamos una variable y la reescalamos (expresándola, por ejemplo, en unidades que son mil veces mayores que las primitivas) en un modelo que incluye la columna de "unos", se verifica que:

    (a) El ajuste es idéntico al primitivo; el coeficiente estimado de la variable reescalada será ahora mil veces más pequeño.

    (b) El ajuste es idéntico al primitivo; el coeficiente estimado de la variable reescalada será ahora mil veces más grande.

    (c) R^2 no varía.

    (d) SSE y SSR no varían, pero SST puede hacerlo.

21. En un modelo ANOVA cruzado, completo, equilibrado, la omisión indebida de uno de los efectos que hubiéramos debido incluir:

    (a) Hace que las estimaciones de todos los demás resulten sesgadas, como sucede en general en Regresión al omitir regresores relevantes.

    (b) No tiene otro efecto que "engordar" la suma de cuadrados de los residuos

    (c) Disminuye los grados de libertad.

    (d) Todo falso.

22. Considera una matriz como $\sigma^2((\boldsymbol{X'X})^{-1} - \boldsymbol{G}(\boldsymbol{X'X})^{-1}\boldsymbol{G'})$ (una tal matriz apareció en la discusión hecha en clase acerca de las consecuencias de introducir restricciones en la estimación). Puedes asegurar por simple inspección:

    (a) Que es simétrica.

    (b) Que tiene en la diagonal principal elementos menores que los de $\sigma^2(\boldsymbol{X}'X)^{-1}$.

    (c) Que tiene en la diagonal principal elementos mayores que los de $\sigma^2(\boldsymbol{X}'X)^{-1}$.

    (d) Que es idempotente.

    (e) Todo falso.

23. Hemos ajustado un modelo de regresión lineal a $N = 534$ observaciones y $p = 30$ parámetros. Computamos los residuos externamente studentizados, que como sabemos siguen una distribución t_{N-p-1} bajo los supuestos habituales más el de normalidad. Observamos un residuo muy separado del resto, del que sospechamos que puede corresponder a una observación anómala, de régimen diferente a la de las demás. Para decidir con criterio estadístico si es el caso, comparamos dicho residuo con el cuantil 0.05 de una t de Student con 503 grados de libertad. Con este modo de proceder, podemos esperar:

    (a) Un 5 % de "falsas alarmas": residuos declarados anómalos cuando son perfectamente normales.

    (b) Mas de un 5 % de "falsas alarmas"

    (c) Menos de un 5 % de "falsas alarmas"

    (d) Todo falso.

24. En un modelo de regresión, el método de estimación en componentes principales utilizando tantas componentes principales como regresores hay, daría:

    (a) El mismo resultado que el método de estimación *ridge* con $k \to \infty$.

    (b) Un resultado igual pero con más sesgo que si utilizáramos menos componentes.

    (c) El mismo resultado que el método de estimación *ridge* con $k = 0$

    (d) El mismo resultado que el método de estimación MCO.

    (e) Todo falso.

25. Al estimar una recta de regresión obtienes $\boldsymbol{\beta_i} = 0{,}74$ y $\sigma_{\boldsymbol{\beta_i}} = 0{,}12$. Los grados de libertad $N - p$ son 324. La práctica adquirida debe permitirte, "a ojo", sin necesidad de consultar tablas, concluir:

    (a) Que $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$ es significativo a un $\alpha = 0{,}05$ e incluso a un $\alpha = 0{,}01$.

    (b) Que difícilmente la estimación de ese parámetro será insesgada.

    (c) Que el parámetro no es significativamente diferente de cero.

    (d) Todo falso.

26. De entre R^2 y $\overline{R}^2$ (R^2 corregida), será un criterio más conservador (= menos proclive a introducir regresores en nuestro modelo)

    (a) R^2

    (b) $\overline{R}^2$

    (c) Depende del número de grados de libertad.

    (d) Depende del número de regresores presentes en el modelo.

27. Si la hipótesis $H_0 : A\boldsymbol{\beta} = \boldsymbol{c}$ es cierta, el hacer estimación mínimo cuadrática forzando a los estimadores a verificar $A\hat{\boldsymbol{\beta}} = \boldsymbol{c}$

    (a) Producirá un SSE_h *mayor o igual* que el SSE en el modelo no restringido.

    (b) Producirá un SSE_h *menor o igual* que el SSE en el modelo no restringido.

    (c) Incrementará las varianzas de los estimadores $\hat{\boldsymbol{\beta}}$.

    (d) Sesgará los estimadores $\hat{\boldsymbol{\beta}}$

    (e) Incrementará R^2 respecto del obtenido en el modelo no restringido.

28. La introducción de *factores de bloque* en el Análisis de Varianza tiene por objeto:

    (a) Corregir el efecto de las múltiples hipótesis simultáneas que se consideran.

    (b) Simplificar el análisis.

    (c) Evitar que la inhomogeneidad de las unidades experimentales perjudique el análisis.

    (d) Reducir los grados de libertad de los estimadores.

    (e) Todo falso.

29. Un modelo está jerárquicamente bien estructurado. . .

    (a) Cuando el número de grados de libertad es máximo.

    (b) Cuando incluye columna de "unos".

    (c) Cuando verifica restricciones de marginalidad: la ausencia de un efecto implica la de todas las interacciones en las que dicho efecto tomaría parte.

    (d) Cuando el diseño es ortogonal y equilibrado.

30. ¿Cuales de los siguientes tipos de residuos son homoscedásticos?

    (a) Los internamente *studentizados*, r_i.

    (b) Los externamente *studentizados*, t_i.

    (c) Los borrados. d_i.

    (d) Los brutos (o MCO), $\hat{\epsilon}_i$.

    (e) Cualesquiera, si realmente se verifica: $E(\epsilon\epsilon') = \sigma^2 I$.

    (f) Ningunos; incluso si realmente se verifica: $E(\epsilon\epsilon') = \sigma^2 I$, los residuos, de cualquier especie, no pueden ser homoscedásticos.

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

Sección 1. Instrucciones

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.

Intento medir conocimientos y no agudeza visual. Inevitablemente, en un examen de este tipo hay que prestar mucha atención. Cada curso hay más de una persona que echa 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.

1. Ajustas el modelo $Y = \beta_0 + \beta_1 X_1 + \epsilon$ y obtienes: $\sigma^2 = 0{,}45$,

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{pmatrix} 0{,}83 & 0{,}34 \\ 0{,}34 & 0{,}04 \end{pmatrix}$$

y $\hat{\boldsymbol{\beta}}' = (0{,}15 \quad 2{,}14)$. Te piden ahora que realices una predicción del valor esperable de Y cuando $X_1 = 4{,}2$ y de su varianza. Tu estimación será aproximadamente:

- **(a)** $\hat{Y} = 9{,}1380$ **y** $\hat{\sigma}_{\hat{Y}} = 8{,}6972$.
- (b) $\hat{Y} = 6{,}2245$ y $\hat{\sigma}_{\hat{Y}} = 5{,}3425$.
- (c) $\hat{Y} = 3{,}2347$ y $\hat{\sigma}_{\hat{Y}} = 2{,}3419$.
- (d) $\hat{Y} = 7{,}3485$ y $\hat{\sigma}_{\hat{Y}} = 6{,}3491$.

2. El vector de residuos $\hat{\boldsymbol{\epsilon}}$:

- (a) Es ortogonal al vector $\boldsymbol{y}$.
- **(b) Es ortogonal a todos los regresores.**
- **(c) Es ortogonal a la columna de “unos”, si está presente.**
- (d) Verifica *siempre* (haya o no columna de “unos”) que $\sum_{i=1}^{N} \hat{\epsilon}_i = 0$

3. En un problema de estimación condicionada,

- **(a)** SSE_h **(la suma de cuadrados en el modelo estimado con la condición) nunca será menor que** SSE **del modelo sin la condición, y esto tanto si ésta es “correcta” como si no.**
- (b) Las estimaciones de los parámetros no afectados por la condición siempre serán insesgadas.
- (c) Los estimadores tendrán menor varianza que los estimadores MCO si, y solo si, la condición impuesta es “correcta” (verificada por los parámetros del modelo).
- (d) Todo falso.

4. La igualdad $SST = SSR + SSE$ de la que se deduce, entre otras cosas, que $0 \leq R^2 \leq 1$ está garantizada si hay:

- (a) Normalidad de las perturbaciones.
- **(b) Columna de “unos” entre los regresores.**
- (c) Homoscedasticidad de las perturbaciones.
- (d) Un número suficiente de grados de libertad; de otro modo, SSE podría ser cero y la igualdad no se verificaría.

5. Si tenemos tres factores (o tratamientos), cada uno con tres niveles, y tenemos la certeza de que no hay interacción, el número de experimentos a realizar si adoptamos un diseño en cuadrado latino asciende a:

- (a) 27
- **(b) 9**
- (c) 12
- (d) Todo falso.

6. Si tenemos tres factores (o tratamientos), cada uno con tres niveles, y queremos replicar 4 veces, el número de experimentos a realizar asciende a:

- **(a) 108**
- (b) 27
- (c) 36
- (d) Todo falso.

7. Podemos estimar un modelo de regresión logística:

   (a) **Por máxima verosimilitud.**

   (b) Resolviendo un sistema lineal análogo a las ecuaciones normales.

   (c) Todo falso.

   (d) Linealizando el modelo, a base de tomar logaritmos y estimar el modelo lineal resultante.

8. El teorema conocido como de Gauss-Markov asegura:

   (a) **Que de entre todos los estimadores lineales e insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.**

   (b) Que de entre todos los estimadores lineales e insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático, con la única excepción del estimador *ridge*.

   (c) Que de entre todos los estimadores insesgados de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.

   (d) Que de entre todos los estimadores lineales de los $\boldsymbol{\beta}$, ninguno proporciona varianzas menores que el $\hat{\boldsymbol{\beta}}$ mínimo-cuadrático.

9. ¿Cuál de los siguientes supuestos es imprescindible para que la proyección de $\boldsymbol{y}$ sobre el subespacio generado por las columnas de $\boldsymbol{X}$ sea única?

   (a) Normalidad en las perturbaciones.

   (b) $\boldsymbol{X}$ de rango completo.

   (c) $\boldsymbol{X}$ de dimensiones $N \times p$ con $N > p$.

   (d) **Ninguno de los anteriores: la proyección de $\boldsymbol{y}$ sobre $R(\boldsymbol{X})$ existe y es única siempre.**

10. En un modelo de regresión lineal, la omisión de regresores que deberían formar parte de la especificación:

    (a) Afecta (sesgándola) a la estimación de β_0, pero no a la de los demás parámetros presentes en el modelo.

    (b) **Sesga, en general, los estimadores de los parámetros incluidos.**

    (c) **Puede no afectar al sesgo de los estimadores de los parámetros incluidos, si acontece que los regresores omitidos son ortogonales a los incluidos.**

    (d) **Disminuye, en general, las varianzas de los estimadores, respecto a las que hubieran sido en el caso de ajustar el modelo completo.**

11. Cuando omitimos la columna de “unos” en la matriz de diseño, el hiperplano ajustado:

    (a) Proporciona mejor ajuste a la $\boldsymbol{y}$ (en términos de SSE) que incluyendo la columna de unos.

    (b) **Pasa *siempre* por el origen de coordenadas.**

    (c) Necesariamente tiene pendiente positiva.

    (d) Pasa a una distancia del origen de coordenadas siempre igual a $||\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}||$.

12. Llamamos *restricciones de identificación* a:

    (a) **Ecuaciones que, suplementando a las ecuaciones normales, permiten obtener una solución única para los $\hat{\boldsymbol{\beta}}$.**

    (b) Restricciones que podemos usar en sustitución de un DNI o pasaporte cuando viajamos fuera de la Unión Europea.

    (c) Ecuaciones que, junto con las ecuaciones normales, dan un sistema compatible indeterminado.

    (d) Ecuaciones no lineales, que sustituyen a las normales y pueden resolverse mediante técnicas iterativas.

13. Al estimar un modelo lineal mediante regresión *ridge*, es decir, $\hat{\boldsymbol{\beta}}_{(\boldsymbol{k})} = (\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I})^{-1}\boldsymbol{X}'\boldsymbol{y}\ldots$

    (a) Sólamente es posible mejorar el ECM (error cuadrático medio) en los casos en que la multicolinealidad es perfecta.

    (b) Un incremento de k acarrea un incremento del sesgo del estimador.

    (c) Un incremento de k acarrea una disminución de la varianza del estimador.

    (d) Un incremento de k acarrea una disminución del error cuadrático medio.

    (e) Siempre hay un k que hace que el estimador *ridge* sea, en términos de ECM, mejor que el mínimo cuadrático ordinario.

14. En un modelo de Análisis de Varianza cruzado completo, el equilibrio junto con las restricciones de identificación garantiza:

    (a) La ortogonalidad por bloques: las estimaciones de los parámetros correspondientes a un efecto no se alteran con la inclusión u omisión de otros efectos.

    (b) La ortogonalidad del vector de residuos a los regresores.

    (c) La homoscedasticidad de la perturbación.

    (d) La normalidad de los estimadores.

15. ¿En qué condiciones puede $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ (proyección de $\boldsymbol{y}$ sobre el subespacio que generan las columnas de $\boldsymbol{X}$) no ser única?

    (a) Cuando haya acusada multicolinealidad entre las columnas de $\boldsymbol{X}$.

    (b) Siempre es única.

    (c) No tiene por qué ser única en ningún caso.

    (d) Todo falso.

16. Los "residuos borrados" *(deleted residuals)* $d_i\ldots$

    (a) ... se definen por $d_i = \hat{\epsilon}_i / \sqrt{\hat{\sigma}^2(1-p_{ii})}$

    (b) ... se definen por $d_i = \hat{\epsilon}_i / \sqrt{\hat{\sigma}^2_{(i)}(1-p_{ii})}$

    (c) ...permiten emplear el criterio de validación cruzada para seleccionar un modelo sin necesidad de realizar mas que una regresión.

    (d) ... se definen por $d_i = \hat{\epsilon}_i / (\hat{\sigma}^2(1-p_{ii}))$

17. En una regresión logística expresamos como función lineal de los predictores:

    (a) Una variable que toma valores 0 y 1.

    (b) El logaritmo de una razón de probabilidades condicionadas.

    (c) Una razón de probabilidades.

    (d) La media de la variable respuesta, que coincide con P_i, la probabilidad de que tome el valor 1.

    (e) Todo falso.

18. ¿Cuando es inadecuada la inclusión de una columna de "unos" entre los regresores?

    (a) Cuando el número de grados de libertad es relativamente elevado.

    (b) Cuando las escalas de los regresores sean muy similares.

    (c) Cuando tanto regresando como regresores se toman en desviaciones respecto a su media. En tal caso, β_0 sería siempre 0.

    (d) Cuando las escalas de los regresores sean muy diferentes.

19. A la luz de la discusión hecha en clase sobre los efectos derivados de incluir regresores irrelevantes en una ecuación de regresión, cabría concluir que el efecto pernicioso de "pasarse" se aminora cuando:

    (a) El tamaño de la muestra crece.

    (b) El número de regresores crece.

    (c) Los regresores añadidos son combinación lineal exacta unos de otros.

    (d) Todo falso.

20. Cuando tomamos una variable y la reescalamos (expresándola, por ejemplo, en unidades que son mil veces mayores que las primitivas) en un modelo que incluye la columna de "unos", se verifica que:

    (a) El ajuste es idéntico al primitivo; el coeficiente estimado de la variable reescalada será ahora mil veces más pequeño.

    (b) El ajuste es idéntico al primitivo; el coeficiente estimado de la variable reescalada será ahora mil veces más grande.

    (c) R^2 no varía.

    (d) SSE y SSR no varían, pero SST puede hacerlo.

21. En un modelo ANOVA cruzado, completo, equilibrado, la omisión indebida de uno de los efectos que hubiéramos debido incluir:

    (a) Hace que las estimaciones de todos los demás resulten sesgadas, como sucede en general en Regresión al omitir regresores relevantes.

    (b) No tiene otro efecto que "engordar" la suma de cuadrados de los residuos

    (c) Disminuye los grados de libertad.

    (d) Todo falso.

22. Considera una matriz como $\sigma^2((\boldsymbol{X'X})^{-1} - \boldsymbol{G}(\boldsymbol{X'X})^{-1}\boldsymbol{G'})$ (una tal matriz apareció en la discusión hecha en clase acerca de las consecuencias de introducir restricciones en la estimación). Puedes asegurar por simple inspección:

    (a) Que es simétrica.

    (b) Que tiene en la diagonal principal elementos menores que los de $\sigma^2(\boldsymbol{X}'X)^{-1}$.

    (c) Que tiene en la diagonal principal elementos mayores que los de $\sigma^2(\boldsymbol{X}'X)^{-1}$.

    (d) Que es idempotente.

    (e) Todo falso.

23. Hemos ajustado un modelo de regresión lineal a $N = 534$ observaciones y $p = 30$ parámetros. Computamos los residuos externamente studentizados, que como sabemos siguen una distribución t_{N-p-1} bajo los supuestos habituales más el de normalidad. Observamos un residuo muy separado del resto, del que sospechamos que puede corresponder a una observación anómala, de régimen diferente a la de las demás. Para decidir con criterio estadístico si es el caso, comparamos dicho residuo con el cuantil 0.05 de una t de Student con 503 grados de libertad. Con este modo de proceder, podemos esperar:

    (a) Un 5 % de "falsas alarmas": residuos declarados anómalos cuando son perfectamente normales.

    (b) Mas de un 5 % de "falsas alarmas"

    (c) Menos de un 5 % de "falsas alarmas"

    (d) Todo falso.

24. En un modelo de regresión, el método de estimación en componentes principales utilizando tantas componentes principales como regresores hay, daría:

    (a) El mismo resultado que el método de estimación *ridge* con $k \to \infty$.

    (b) Un resultado igual pero con más sesgo que si utilizáramos menos componentes.

    (c) El mismo resultado que el método de estimación *ridge* con $k = 0$

    (d) El mismo resultado que el método de estimación MCO.

    (e) Todo falso.

25. Al estimar una recta de regresión obtienes $\boldsymbol{\beta_i} = 0{,}74$ y $\sigma_{\boldsymbol{\beta_i}} = 0{,}12$. Los grados de libertad $N - p$ son 324. La práctica adquirida debe permitirte, "a ojo", sin necesidad de consultar tablas, concluir:

    (a) Que $\hat{\boldsymbol{\beta}}_{\boldsymbol{i}}$ es significativo a un $\alpha = 0{,}05$ e incluso a un $\alpha = 0{,}01$.

    (b) Que difícilmente la estimación de ese parámetro será insesgada.

    (c) Que el parámetro no es significativamente diferente de cero.

    (d) Todo falso.

26. De entre R^2 y $\overline{R}^2$ (R^2 corregida), será un criterio más conservador (= menos proclive a introducir regresores en nuestro modelo)

    (a) R^2

    (b) $\overline{R}^2$

    (c) Depende del número de grados de libertad.

    (d) Depende del número de regresores presentes en el modelo.

27. Si la hipótesis $H_0 : A\boldsymbol{\beta} = \boldsymbol{c}$ es cierta, el hacer estimación mínimo cuadrática forzando a los estimadores a verificar $A\hat{\boldsymbol{\beta}} = \boldsymbol{c}$

    (a) Producirá un SSE_h *mayor o igual* que el SSE en el modelo no restringido.

    (b) Producirá un SSE_h *menor o igual* que el SSE en el modelo no restringido.

    (c) Incrementará las varianzas de los estimadores $\hat{\boldsymbol{\beta}}$.

    (d) Sesgará los estimadores $\hat{\boldsymbol{\beta}}$

    (e) Incrementará R^2 respecto del obtenido en el modelo no restringido.

28. La introducción de *factores de bloque* en el Análisis de Varianza tiene por objeto:

    (a) Corregir el efecto de las múltiples hipótesis simultáneas que se consideran.

    (b) Simplificar el análisis.

    (c) Evitar que la inhomogeneidad de las unidades experimentales perjudique el análisis.

    (d) Reducir los grados de libertad de los estimadores.

    (e) Todo falso.

29. Un modelo está jerárquicamente bien estructurado. . .

    (a) Cuando el número de grados de libertad es máximo.

    (b) Cuando incluye columna de "unos".

    (c) Cuando verifica restricciones de marginalidad: la ausencia de un efecto implica la de todas las interacciones en las que dicho efecto tomaría parte.

    (d) Cuando el diseño es ortogonal y equilibrado.

30. ¿Cuales de los siguientes tipos de residuos son homoscedásticos?

    (a) Los internamente *studentizados*, r_i.

    (b) Los externamente *studentizados*, t_i.

    (c) Los borrados. d_i.

    (d) Los brutos (o MCO), $\hat{\epsilon}_i$.

    (e) Cualesquiera, si realmente se verifica: $E(\epsilon\epsilon') = \sigma^2 I$.

    (f) Ningunos; incluso si realmente se verifica: $E(\epsilon\epsilon') = \sigma^2 I$, los residuos, de cualquier especie, no pueden ser homoscedásticos.