

4 de Junio de 1.999

Puedes responder: i) A todas las cuestiones de la Sección 1; o ii) Idem. de la Sección 2. En ambos casos puedes obtener puntuación total. No puedes sin embargo escoger cosas de ambas Secciones para acumular puntos.

Estadística III (P28)

Exam, Tipo: A

Name: _____

Student Number: _____

TA: _____

Date: _____

Las preguntas respondidas correctamente, valen un punto salvo que se indique lo contrario. Las preguntas incorrectas cuentan como -0.5 puntos. Las preguntas en blanco no cuentan. Puede haber más de una respuesta correcta y en tal caso has de señalarlas todas.

1. Se adopta el modelo

$$\log\left(\frac{P(Y_i = 1)}{1 - P(Y_i = 1)}\right) = \beta_0 + \vec{x}_i' \vec{\beta}$$

para explicar la incidencia de una cierta enfermedad en función de diversas variables recogidas en \vec{x} . Para estimar dicho modelo, se toman 500 sujetos enfermos y 500 pacientes sanos, investigando a continuación los valores que para cada uno de ellos toman las variables en \vec{x}_i . Con tal diseño experimental:

- (a) Podremos, en general, estimar la probabilidad de que un sujeto con valores dados de \vec{x} contraiga la enfermedad.
- (b) Podremos, en general, contrastar si las variables incluidas en \vec{x} influyen en la propensión a contraer dicha enfermedad.
- (c) La estimación de β_0 será arbitraria y carente de sentido.
- (d) Todo falso.

2. “Emplear un criterio como validación cruzada para escoger entre diferentes modelos de regresión lineal, es una idea atractiva, pero extremadamente intensiva en computación: para cada modelo tenemos que realizar varias (típicamente entre 5 y 10) regresiones, dejando de lado cada vez una porción de la muestra.”

- (a) Cierto.
- (b) Falso.

3. Señala cuál o cuáles de los siguientes estadísticos están relacionados con la influencia de una observación:

- (a) R^2
- (b) SIC
- (c) C_p
- (d) Distancia de Cook.
- (e) DFIT

4. Observa detenidamente la siguiente salida, proporcionando los resultados de una regresión.

```
>lm(formula = Y ~ X1 + X2 + X3 + X4)
Residuals:
    Min       1Q   Median       3Q      Max
-1992.67  -156.96   -52.65   112.76  1746.24
Coefficients:
            Estimate Std. Error t value
(Intercept)  144.59611    55.16005    2.621
X1           1.78411    12.34601    0.144
X2          -0.11953     0.12560   -0.952
X3          -0.05906     0.35387   -0.167
X4           0.68074     4.29712    0.158
Residual standard error: 472.7 on 106 degrees of freedom
Multiple R-Squared:  0.994, Adjusted R-squared:  0.9938
F-statistic: 4381 on 4 and 106 degrees of freedom,
p-value: 0
```

Es evidente que:

- (a) Hay fuerte multicolinealidad.
- (b) Los regresores, en su conjunto, no tienen gran cosa que decir acerca del regresando.
- (c) Los regresores son ortogonales.
- (d) Hay un problema de heterocedasticidad.
- (e) Se hace precisa una transformación no lineal del regresando.

5. ¿En cuál o cuáles de entre las siguientes situaciones podría ayudar una transformación de Box-Cox del regresando?

- (a) Cuando hay fuerte multicolinealidad.
- (b) Cuando la distribución de los residuos parece separarse de la normalidad.
- (c) Cuando hay sospecha fundada de que algunos de los regresores son irrelevantes.
- (d) Cuando la perturbación no parece tener media cero, como debe de acuerdo con los supuestos habituales.
- (e) Todo falso.

6. Una matriz de proyección es siempre:

- (a) Definida positiva.
- (b) Definida negativa.
- (c) Triangular superior.
- (d) Diagonalizable, con valores propios estrictamente positivos.
- (e) Idempotente.

7. Un modelo ANOVA con dos tratamientos cruzados, con tres y cuatro niveles respectivamente, sin interacción, posee:

- (a) Seis parámetros libres.
- (b) Siete parámetros libres.
- (c) Cinco parámetros libres.
- (d) Ocho parámetros libres.
- (e) Nueve parámetros libres.

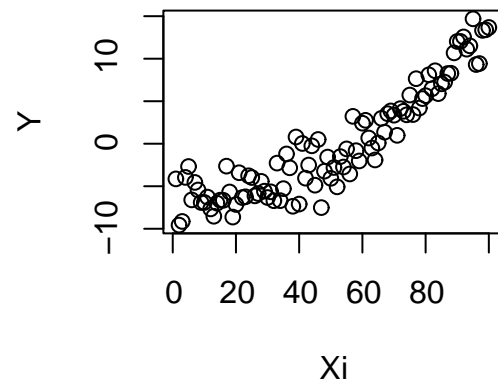
8. Para contrastar la hipótesis $H_0 : \beta_1 = \dots = \beta_{p-1}$ en un modelo de regresión lineal con normalidad, conocido el tamaño muestral N y número de regresores p , basta saber:

- (a) R^2 .
- (b) SSR.
- (c) SST.
- (d) Todos los $\hat{\beta}$'s.
- (e) La varianza estimada de los residuos.

9. Si para contrastar una hipótesis que prescribe un valor para cada uno de k parámetros (por ejemplo, $H_0 : \beta_i = 0$ para $i = 1, \dots, k$) contrastásemos al nivel de significación α todas y cada una de las siguientes: $H_{0i} : \beta_i = 0$, para $i = 1, \dots, k$, rechazando H_0 si alguna de las H_{0i} fuera rechazada, el nivel de significación:

- (a) Sería en general mayor que α .
- (b) Sería en general menor que α .
- (c) Estaría acotado superiormente por $k\alpha$.
- (d) Estaría acotado inferiormente por $k\alpha$.
- (e) Todo falso.

10. Observa el siguiente gráfico. En él, se representan los residuos ordinarios de un cierto modelo en ordenadas, frente a los valores de un regresor X_i en abscisas.



A la vista del mismo,

- (a) Es evidente la ausencia de homoscedasticidad.
- (b) Parece que es preciso transformar el regresando Y sustituyéndolo por alguna potencia entera (Y^2, Y^3, \dots).
- (c) Parece que es preciso transformar el regresor X_i sustituyéndolo por alguna potencia entera (X_i^2, X_i^3, \dots).
- (d) Es evidente que un modelo lineal en los parámetros está fuera de la realidad; se impone hacer regresión no lineal.
- (e) Todo falso.

11. Si para contrastar la hipótesis $H_0 : \beta_1 = \dots = \beta_{p-1}$ adoptásemos el criterio de calcular todos los estadísticos $t, \hat{\beta}_i/\hat{\sigma}_{\beta_i}$, compararlos con una t de Student de grados de libertad adecuados, y rechazar H_0 cuando $|\hat{\beta}_i/\hat{\sigma}_{\beta_i}| > t_{N-p}^{\alpha/2}$ para algún i , el contraste resultante:
- Tendría un error de tipo I en general igual a α
 - Tendría un error de tipo I en general inferior a α
 - Tendría una potencia muy baja.
 - Tendría probabilidad de dar lugar al rechazo superior a α .
 - Todo falso.
12. Al hacer regresión en componentes principales introducimos un sesgo pero reducimos la varianza de los estimadores.
- Cierto
 - Falso
13. Un modelo ANOVA incompleto, que hace uso de un diseño en cuadrado latino, siendo cuatro el número común de niveles de cada tratamiento, posee:
- Catorce parámetros libres.
 - Doce parámetros libres.
 - Once parámetros libres.
 - Diez parámetros libres.
 - Nueve parámetros libres.
14. Los siguientes tipos de estimación se reducen a resolver un sistema de ecuaciones lineales en los parámetros:
- Regresión MCO.
 - Regresión *ridge*.
 - Regresión *logit*.
 - Modelos ANOVA cruzados.
 - Modelos ANOVA anidados.
15. ¿Cuál o cuáles de los siguientes hechos tomarías como evidencia de que una observación es influyente?
- Un residuo borrado grande, y un residuo estudentizado (interna o externamente) pequeño.
 - Un residuo internamente estudentizado grande.
 - Un residuo externamente estudentizado pequeño.
 - Un residuo borrado pequeño.
 - Un residuo ordinario grande.
16. En igualdad de todo lo demás, la inclusión de un regresor irrelevante en un modelo será menos grave cuando:
- El tamaño de muestra sea muy grande.
 - Un regresor omitido sea ortogonal a todos los incluidos.
 - Las perturbaciones sean homoscedásticas.
 - El número de parámetros presentes en el modelo sea muy elevado.
 - Se trate de un modelo sin ordenada en el origen, y por tanto sin parámetro β_0 .
17. ¿Cuál de las siguientes características en un modelo de Análisis de Varianza es responsable de que los residuos tengan todos la misma varianza?
- El equilibrio.
 - La ortogonalidad entre regresores, una vez consideradas las restricciones.
 - El hecho de que los regresores son variables cualitativas.
 - El hecho de que la matriz de diseño es de rango deficiente.
18. En un modelo de regresión con p regresores (incluida, si está presente, la columna de “unos”), en que la matriz de diseño X es de rango completo, la matriz de proyección $X(X'X)^{-1}X'$ tiene traza igual a:
- N
 - p
 - $N - p$
 - traza $\Sigma_{\hat{\beta}}$.
 - Todo falso.

19. Al estimar los parámetros de un modelo lineal mediante regresión *ridge*, el incremento de λ acarrea:
- Un sesgo que crece monótonamente con λ .
 - Un incremento de la varianza de los estimadores, tanto mayor cuanto mayor es λ .
 - Una disminución en la varianza hasta un cierto valor de λ ; luego la varianza vuelve a crecer.
 - Una no linealidad progresivamente creciente en las ecuaciones que es preciso resolver para hallar los estimadores de los parámetros.
 - Todo falso.
20. Vista la naturaleza de los inconvenientes que derivan de sobreparametrizar un modelo, hay que concluir que son más graves con una muestra pequeña que con una muestra muy grande.
- Cierto.
 - Falso.
21. ¿Cuáles de entre los siguientes conjuntos de información te permitirían calcular los residuos internamente *studentizados*?
- $\hat{\sigma}^2$ y $X'X$.
 - $\hat{\sigma}^2$ y $X(X'X)^{-1}X'$.
 - $\hat{\sigma}^2$ y traza[$X(X'X)^{-1}X'$].
 - $\hat{\sigma}^2$ y la diagonal de $X(X'X)^{-1}X'$.
 - Todo falso.
22. Al seleccionar un modelo, sobreparametrizar tendría el desgraciado efecto de:
- Sesgar los estimadores de (en general) todos los parámetros.
 - Sesgar los estimadores de los parámetros incluidos innecesariamente.
 - Sesgar los estimadores de los parámetros relevantes.
 - Sesgar $\hat{\sigma}^2 = \frac{SSE}{N-p}$, estimador habitual de la varianza.
 - Todo falso.
23. Los estimadores mínimo cuadrático ordinarios MCO son:
- Insesgados.
 - Insesgados y de varianza mínima entre los lineales insesgados.
 - De error cuadrático medio mínimo entre los lineales insesgados.
 - De error cuadrático medio mínimo, en todo caso.
 - De sesgo mínimo entre todos los lineales.
24. Señala la o las parejas de estadísticos que juegan papeles equiparables en diferentes tipos de modelos:
- R^2 y SSE .
 - AIC y C_p .
 - C_p y *deviance*.
 - Deviance* y SSE .
 - SSR y SIC.
25. El carácter cualitativo de los regresores es *per se* una violación del supuesto de normalidad de las perturbaciones, y fuerza a considerar alternativas al modelo lineal ordinario. Una de ellas es el modelo logit.
- Cierto
 - Falso
26. Al hacer regresión en componentes principales podríamos obtener una estimación (quizá sesgada) de una función lineal de los parámetros $\vec{c}'\vec{\beta}$ no estimable insesgadamente.
- Cierto
 - Falso
27. ¿En cuales de las siguientes circunstancias el vector de residuos MCO, $\hat{\epsilon}$, se distribuye como $N(\vec{0}, H)$, siendo H una matriz de rango complejo?
- Nunca, no puede.
 - En el caso de modelos en que, como sucede en ANOVA equilibrado, los elementos p_{ii} en la diagonal de la matriz de proyección son idénticos.
 - Siempre.
 - Si, y sólo si, la matriz de diseño es de rango completo.
 - Todo falso.

28. Un cambio de escala de los regresores (uno o varios):
- Altera SSE.
 - Altera el o los estimadores $\hat{\beta}$ asociados a los regresores cuya escala ha cambiado.
 - Altera R^2 proporcionalmente.
 - Altera R^2 en proporción al cuadrado del cambio de escala introducido.
 - Todo falso.
29. En un modelo de regresión “escaso”, los estimadores de los parámetros correspondientes a variables incluidas:
- Son, en general, sesgados.
 - Pueden ser insesgados (si los regresores indebidamente excluidos son ortogonales a los regresores presentes).
 - Pueden ser insesgados, pero sólo si los regresores presentes son todos ortogonales entre sí.
 - Todo falso.
30. En un modelo ortogonal por bloques (como los modelos ANOVA estudiados en clase al imponer las restricciones de identificación), en presencia de normalidad,
- Los parámetros en cada bloque son independientes entre sí.
 - Los parámetros en cada bloque son independientes de los pertenecientes a otros bloques.
 - Omitir un bloque completo (por ej., todos los parámetros recogiendo el efecto de un tratamiento) incrementaría SSE exactamente en el mismo valor que la suma de cuadrados atribuible a dicho tratamiento.
 - Omitir un bloque completo (por ej., todos los parámetros recogiendo el efecto de un tratamiento) incrementaría SSE en una cantidad diferente según qué otros regresores quedaran presentes en el modelo.
31. ¿Cuáles de los siguientes problemas pueden acarrear la inestimabilidad de algún o algunos parámetros?
- La multicolinealidad exacta..
 - El incumplimiento de la hipótesis de homoscedasticidad.
 - El incumplimiento de la hipótesis de normalidad.
 - El incumplimiento de la hipótesis de incorrelación entre los residuos.
 - Todo falso.
32. Un modelo ANOVA con dos tratamientos cruzados, con tres y cuatro niveles respectivamente, *con interacción*, posee:
- Catorce parámetros libres.
 - Todo falso.
 - Doce parámetros libres.
 - Once parámetros libres.
 - Diez parámetros libres.
33. ¿En cuáles de los siguientes casos un $\hat{\beta}_i / \hat{\sigma}_{\hat{\beta}_i}$ no sigue una distribución *t* de Student?
- Cuando $\hat{\beta}_i$ ha sido estimado mediante regresión *ridge*.
 - Cuando $\hat{\beta}_i$ ha sido estimado en un modelo *logit* por máxima verosimilitud.
 - Cuando $\hat{\beta}_i$ ha sido estimado mediante regresión mínimo cuadrática ordinaria, pero $H_0 : \beta_i = 0$ no es cierta.
 - Todo falso.
34. Imponer restricciones *a priori* sobre los valores de los parámetros, tiene como efecto:
- Nunca incrementa la varianza de los estimadores.
 - Siempre* crea algún sesgo en los estimadores.
 - Siempre* crea algún sesgo en los estimadores, pero la reducción en varianza que logra puede compensar.
 - Todo falso.

Respuestas al examen de tipo **A**

Las preguntas respondidas correctamente, valen un punto salvo que se indique lo contrario. Las preguntas incorrectas cuentan como -0.5 puntos. Las preguntas en blanco no cuentan. Puede haber más de una respuesta correcta y en tal caso has de señalarlas todas.

1. Se adopta el modelo

$$\log\left(\frac{P(Y_i = 1)}{1 - P(Y_i = 1)}\right) = \beta_0 + \vec{x}_i' \vec{\beta}$$

para explicar la incidencia de una cierta enfermedad en función de diversas variables recogidas en \vec{x} . Para estimar dicho modelo, se toman 500 sujetos enfermos y 500 pacientes sanos, investigando a continuación los valores que para cada uno de ellos toman las variables en \vec{x}_i . Con tal diseño experimental:

- (a) Podremos, en general, estimar la probabilidad de que un sujeto con valores dados de \vec{x} contraiga la enfermedad.
 - (b) Podremos, en general, contrastar si las variables incluidas en \vec{x} influyen en la propensión a contraer dicha enfermedad.
 - (c) La estimación de β_0 será arbitraria y carente de sentido.
 - (d) Todo falso.
2. "Emplear un criterio como validación cruzada para escoger entre diferentes modelos de regresión lineal, es una idea atractiva, pero extremadamente intensiva en computación: para cada modelo tenemos que realizar varias (típicamente entre 5 y 10) regresiones, dejando de lado cada vez una porción de la muestra."
- (a) Cierto.
 - (b) Falso.

3. Señala cuál o cuáles de los siguientes estadísticos están relacionados con la influencia de una observación:

- (a) R^2
- (b) SIC
- (c) C_p
- (d) Distancia de Cook.
- (e) DFIT

4. Observa detenidamente la siguiente salida, proporcionando los resultados de una regresión.

```
>lm(formula = Y ~ X1 + X2 + X3 + X4)
Residuals:
    Min       1Q   Median       3Q      Max
-1992.67  -156.96   -52.65   112.76  1746.24
Coefficients:
            Estimate Std. Error t value
(Intercept)  144.59611    55.16005   2.621
           X1    1.78411    12.34601   0.144
           X2   -0.11953     0.12560  -0.952
           X3   -0.05906     0.35387  -0.167
           X4    0.68074     4.29712   0.158
Residual standard error: 472.7 on 106 degrees of freedom
Multiple R-Squared:  0.994, Adjusted R-squared:  0.9938
F-statistic: 4381 on 4 and 106 degrees of freedom,
p-value: 0
```

Es evidente que:

- (a) Hay fuerte multicolinealidad.
 - (b) Los regresores, en su conjunto, no tienen gran cosa que decir acerca del regresando.
 - (c) Los regresores son ortogonales.
 - (d) Hay un problema de heterocedasticidad.
 - (e) Se hace precisa una transformación no lineal del regresando.
5. ¿En cuál o cuáles de entre las siguientes situaciones podría ayudar una transformación de Box-Cox del regresando?
- (a) Cuando hay fuerte multicolinealidad.
 - (b) Cuando la distribución de los residuos parece separarse de la normalidad.
 - (c) Cuando hay sospecha fundada de que algunos de los regresores son irrelevantes.
 - (d) Cuando la perturbación no parece tener media cero, como debe de acuerdo con los supuestos habituales.
 - (e) Todo falso.

6. Una matriz de proyección es siempre:

- (a) Definida positiva.
- (b) Definida negativa.
- (c) Triangular superior.
- (d) Diagonalizable, con valores propios estrictamente positivos.
- (e) Idempotente.

7. Un modelo ANOVA con dos tratamientos cruzados, con tres y cuatro niveles respectivamente, sin interacción, posee:

- (a) Seis parámetros libres.
- (b) Siete parámetros libres.
- (c) Cinco parámetros libres.
- (d) Ocho parámetros libres.
- (e) Nueve parámetros libres.

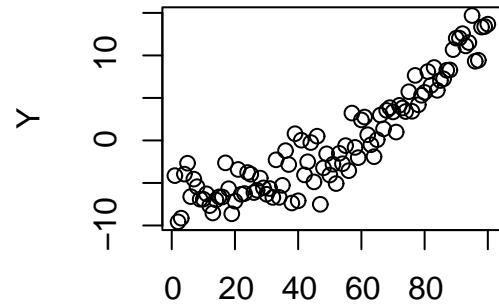
8. Para contrastar la hipótesis $H_0 : \beta_1 = \dots = \beta_{p-1}$ en un modelo de regresión lineal con normalidad, conocido el tamaño muestral N y número de regresores p , basta saber:

- (a) R^2 .
- (b) SSR.
- (c) SST.
- (d) Todos los $\hat{\beta}$'s.
- (e) La varianza estimada de los residuos.

9. Si para contrastar una hipótesis que prescribe un valor para cada uno de k parámetros (por ejemplo, $H_0 : \beta_i = 0$ para $i = 1, \dots, k$) contrastásemos al nivel de significación α todas y cada una de las siguientes: $H_{0i} : \beta_i = 0$, para $i = 1, \dots, k$, rechazando H_0 si alguna de las H_{0i} fuera rechazada, el nivel de significación:

- (a) Sería en general mayor que α .
- (b) Sería en general menor que α .
- (c) Estaría acotado superiormente por $k\alpha$.
- (d) Estaría acotado inferiormente por $k\alpha$.
- (e) Todo falso.

10. Observa el siguiente gráfico. En él, se representan los residuos ordinarios de un cierto modelo en ordenadas, frente a los valores de un regresor X_i en abscisas.



Xi

A la vista del mismo,

- (a) Es evidente la ausencia de homoscedasticidad.
- (b) Parece que es preciso transformar el regresando Y sustituyéndolo por alguna potencia entera (Y^2, Y^3, \dots).
- (c) Parece que es preciso transformar el regresor X_i sustituyéndolo por alguna potencia entera (X_i^2, X_i^3, \dots).
- (d) Es evidente que un modelo lineal en los parámetros está fuera de la realidad; se impone hacer regresión no lineal.
- (e) Todo falso.

11. Si para contrastar la hipótesis $H_0 : \beta_1 = \dots = \beta_{p-1}$ adoptásemos el criterio de calcular todos los estadísticos $t, \hat{\beta}_i / \hat{\sigma}_{\beta_i}$, compararlos con una t de Student de grados de libertad adecuados, y rechazar H_0 cuando $|\hat{\beta}_i / \hat{\sigma}_{\beta_i}| > t_{N-p}^{\alpha/2}$ para algún i , el contraste resultante:

- (a) Tendría un error de tipo I en general igual a α
- (b) Tendría un error de tipo I en general inferior a α
- (c) Tendría una potencia muy baja.
- (d) Tendría probabilidad de dar lugar al rechazo superior a α .
- (e) Todo falso.

12. Al hacer regresión en componentes principales introducimos un sesgo pero reducimos la varianza de los estimadores.
- (a) Cierto
 (b) Falso
13. Un modelo ANOVA incompleto, que hace uso de un diseño en cuadrado latino, siendo cuatro el número común de niveles de cada tratamiento, posee:
- (a) Catorce parámetros libres.
 (b) Doce parámetros libres.
 (c) Once parámetros libres.
 (d) Diez parámetros libres.
 (e) Nueve parámetros libres.
14. Los siguientes tipos de estimación se reducen a resolver un sistema de ecuaciones lineales en los parámetros:
- (a) Regresión MCO.
 (b) Regresión *ridge*.
 (c) Regresión *logit*.
 (d) Modelos ANOVA cruzados.
 (e) Modelos ANOVA anidados.
15. ¿Cuál o cuáles de los siguientes hechos tomarías como evidencia de que una observación es influyente?
- (a) Un residuo borrado grande, y un residuo estudentizado (interna o externamente) pequeño.
 (b) Un residuo internamente estudentizado grande.
 (c) Un residuo externamente estudentizado pequeño.
 (d) Un residuo borrado pequeño.
 (e) Un residuo ordinario grande.
16. En igualdad de todo lo demás, la inclusión de un regresor irrelevante en un modelo será menos grave cuando:
- (a) El tamaño de muestra sea muy grande.
 (b) Un regresor omitido sea ortogonal a todos los incluidos.
 (c) Las perturbaciones sean homoscedásticas.
 (d) El número de parámetros presentes en el modelo sea muy elevado.
 (e) Se trate de un modelo sin ordenada en el origen, y por tanto sin parámetro β_0 .
17. ¿Cuál de las siguientes características en un modelo de Análisis de Varianza es responsable de que los residuos tengan todos la misma varianza?
- (a) El equilibrio.
 (b) La ortogonalidad entre regresores, una vez consideradas las restricciones.
 (c) El hecho de que los regresores son variables cualitativas.
 (d) El hecho de que la matriz de diseño es de rango deficiente.
18. En un modelo de regresión con p regresores (incluida, si está presente, la columna de "unos"), en que la matriz de diseño X es de rango completo, la matriz de proyección $X(X'X)^{-1}X'$ tiene traza igual a:
- (a) N
 (b) p
 (c) $N - p$
 (d) traza $\Sigma_{\hat{\beta}}$.
 (e) Todo falso.
19. Al estimar los parámetros de un modelo lineal mediante regresión *ridge*, el incremento de λ acarrea:
- (a) Un sesgo que crece monótonamente con λ .
 (b) Un incremento de la varianza de los estimadores, tanto mayor cuanto mayor es λ .
 (c) Una disminución en la varianza hasta un cierto valor de λ ; luego la varianza vuelve a crecer.
 (d) Una no linealidad progresivamente creciente en las ecuaciones que es preciso resolver para hallar los estimadores de los parámetros.
 (e) Todo falso.
20. Vista la naturaleza de los inconvenientes que derivan de sobreparametrizar un modelo, hay que concluir que son más graves con una muestra pequeña que con una muestra muy grande.
- (a) Cierto.
 (b) Falso.

21. ¿Cuáles de entre los siguientes conjuntos de información te permitirían calcular los residuos internamente *studentizados*?
- (a) $\hat{\sigma}^2$ y $X'X$.
 - (b) $\hat{\sigma}^2$ y $X(X'X)^{-1}X'$.
 - (c) $\hat{\sigma}^2$ y $\text{traza}[X(X'X)^{-1}X']$.
 - (d) $\hat{\sigma}^2$ y la diagonal de $X(X'X)^{-1}X'$.
 - (e) Todo falso.
22. Al seleccionar un modelo, sobreparametrizar tendría el desgraciado efecto de:
- (a) Sesgar los estimadores de (en general) todos los parámetros.
 - (b) Sesgar los estimadores de los parámetros incluidos innecesariamente.
 - (c) Sesgar los estimadores de los parámetros relevantes.
 - (d) Sesgar $\hat{\sigma}^2 = \frac{SSE}{N-p}$, estimador habitual de la varianza.
 - (e) Todo falso.
23. Los estimadores mínimo cuadrático ordinarios MCO son:
- (a) Inesgados.
 - (b) Inesgados y de varianza mínima entre los lineales inesgados.
 - (c) De error cuadrático medio mínimo entre los lineales inesgados.
 - (d) De error cuadrático medio mínimo, en todo caso.
 - (e) De sesgo mínimo entre todos los lineales.
24. Señala la o las parejas de estadísticos que juegan papeles equiparables en diferentes tipos de modelos:
- (a) R^2 y SSE .
 - (b) AIC y C_p .
 - (c) C_p y *deviance*.
 - (d) *Deviance* y SSE .
 - (e) SSR y SIC.
25. El carácter cualitativo de los regresores es *per se* una violación del supuesto de normalidad de las perturbaciones, y fuerza a considerar alternativas al modelo lineal ordinario. Una de ellas es el modelo logit.
- (a) Cierto
 - (b) Falso
26. Al hacer regresión en componentes principales podríamos obtener una estimación (quizá sesgada) de una función lineal de los parámetros $\vec{c}'\vec{\beta}$ no estimable inesgadamente.
- (a) Cierto
 - (b) Falso
27. ¿En cuales de las siguientes circunstancias el vector de residuos MCO, $\hat{\epsilon}$, se distribuye como $N(\vec{0}, H)$, siendo H una matriz de rango complejo?
- (a) Nunca, no puede.
 - (b) En el caso de modelos en que, como sucede en ANOVA equilibrado, los elementos p_{ii} en la diagonal de la matriz de proyección son idénticos.
 - (c) Siempre.
 - (d) Si, y sólo si, la matriz de diseño es de rango completo.
 - (e) Todo falso.
28. Un cambio de escala de los regresores (uno o varios):
- (a) Altera SSE.
 - (b) Altera el o los estimadores $\hat{\beta}$ asociados a los regresores cuya escala ha cambiado.
 - (c) Altera R^2 proporcionalmente.
 - (d) Altera R^2 en proporción al cuadrado del cambio de escala introducido.
 - (e) Todo falso.
29. En un modelo de regresión "escaso", los estimadores de los parámetros correspondientes a variables incluidas:
- (a) Son, en general, sesgados.
 - (b) Pueden ser inesgados (si los regresores indebidamente excluidos son ortogonales a los regresores presentes).
 - (c) Pueden ser inesgados, pero sólo si los regresores presentes son todos ortogonales entre sí.
 - (d) Todo falso.

30. En un modelo ortogonal por bloques (como los modelos ANOVA estudiados en clase al imponer las restricciones de identificación), en presencia de normalidad,
- (a) Los parámetros en cada bloque son independientes entre sí.
 - (b) Los parámetros en cada bloque son independientes de los pertenecientes a otros bloques.
 - (c) Omitir un bloque completo (por ej., todos los parámetros recogiendo el efecto de un tratamiento) incrementaría SSE exactamente en el mismo valor que la suma de cuadrados atribuible a dicho tratamiento.
 - (d) Omitir un bloque completo (por ej., todos los parámetros recogiendo el efecto de un tratamiento) incrementaría SSE en una cantidad diferente según qué otros regresores quedaran presentes en el modelo.
31. ¿Cuáles de los siguientes problemas pueden acarrear la inestimabilidad de algún o algunos parámetros?
- (a) La multicolinealidad exacta..
 - (b) El incumplimiento de la hipótesis de homoscedasticidad.
 - (c) El incumplimiento de la hipótesis de normalidad.
 - (d) El incumplimiento de la hipótesis de incorrelación entre los residuos.
 - (e) Todo falso.
32. Un modelo ANOVA con dos tratamientos cruzados, con tres y cuatro niveles respectivamente, *con interacción*, posee:
- (a) Catorce parámetros libres.
 - (b) Todo falso.
 - (c) Doce parámetros libres.
 - (d) Once parámetros libres.
 - (e) Diez parámetros libres.
33. ¿En cuáles de los siguientes casos un $\hat{\beta}_i/\hat{\sigma}_{\hat{\beta}_i}$ no sigue una distribución t de Student?
- (a) Cuando $\hat{\beta}_i$ ha sido estimado mediante regresión *ridge*.
 - (b) Cuando $\hat{\beta}_i$ ha sido estimado en un modelo *logit* por máxima verosimilitud.
 - (c) Cuando $\hat{\beta}_i$ ha sido estimado mediante regresión mínimo cuadrática ordinaria, pero $H_0 : \beta_i = 0$ no es cierta.
 - (d) Todo falso.
34. Imponer restricciones *a priori* sobre los valores de los parámetros, tiene como efecto:
- (a) Nunca incrementa la varianza de los estimadores.
 - (b) *Siempre* crea algún sesgo en los estimadores.
 - (c) *Siempre* crea algún sesgo en los estimadores, pero la reducción en varianza que logra puede compensar.
 - (d) Todo falso.