

## Regresión múltiple y otras técnicas multivariadas | Semestre 2019-2

### Tarea 10: repaso para el segundo examen parcial

Fecha de entrega: 1 de mayo

1. Enunciar los supuestos del modelo de regresión múltiple.
2. Enuncie correctamente el Teorema de Gauss-Markov para el estimador de  $\beta$  en el modelo de regresión múltiple.
3. Mostrar que el estadístico  $F$  utilizado para contrastar las hipótesis

$$H_0 : \beta_1 = \dots = \beta_p = 0 \quad \text{vs.} \quad H_1 : \beta_i \neq 0 \quad \text{para alguna } i,$$

se puede escribir como

$$F = \frac{R^2(n-p-1)}{p(1-R^2)},$$

donde  $R^2$  es el coeficiente de determinación del modelo.

4. Suponer que se ha ajustado un modelo de regresión lineal con  $p = 2$  variables explicativas y  $n = 25$  observaciones y que los resultados muestran que  $R^2 = 0.90$ .
  - a) Contrastar la hipótesis de significancia de la regresión. Utilizar  $\alpha = 0.05$ .
  - b) ¿Cuál es el mínimo valor de  $R^2$  que nos lleva a concluir que la regresión es significativa?
5. Suponer que se ajusta el modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon.$$

En cada caso, indicar qué matriz  $\mathbf{A}$  y qué vector  $\mathbf{b}$  se deben utilizar para contrastar las siguientes hipótesis

- a)  $\beta_1 = \beta_2 = \beta_3 = \beta_4$ .
  - b)  $\beta_1 = \beta_2, \beta_3 = \beta_4$ .
  - c)  $\beta_1 - 2\beta_2 = 4\beta_3, \beta_1 + 2\beta_2 = 0$ .
6. Se ajustó con R un modelo lineal para explicar el ingreso por trabajo en los hogares a partir del gasto, un índice de características de la vivienda y un índice de equipamiento de las viviendas (bienes). Los resultados se muestran a continuación:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	273.09923	3843.94478	0.071	0.9434
Gasto	0.90400	0.02202	41.059	<2e-16 ***

Vivienda	-25.67979	48.07923	-0.534	0.5938
Bienes	44.90692	17.99172	2.496	0.0132 *

---

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

Residual standard error: 2524 on 241 degrees of freedom  
Multiple R-squared: 0.898, Adjusted R-squared: 0.8967  
F-statistic: 707.3 on 3 and 241 DF, p-value: < 2.2e-16

- ¿El modelo es significativo?
- Calcular intervalos de confianza simultáneos con el método de Hotelling-Scheffé para  $\beta_0$ ,  $\beta_1$ ,  $\beta_2$  y  $\beta_3$ .
- ¿Qué variables son significativas para explicar el ingreso?
- ¿Qué porcentaje de la varianza del ingreso es explicada por el modelo?
- ¿Cómo interpretaría la estimación del coeficiente del gasto?
- ¿Sería mejor ajustar un modelo sin intercepto?
- ¿Se podría afirmar que el índice de vivienda tiene un efecto negativo en el ingreso?
- ¿Qué cambios propondría para mejorar el modelo?
- Construir la tabla ANOVA con los resultados del ajuste del modelo.