

1. Suponer que se ajusta un modelo RLS a las observaciones  $(x_i, y_i)$ ,  $i = 1, \dots, n$ . Mostrar que

$$SCE = \frac{S_{xx}S_{yy} - S_{xy}^2}{S_{xx}}.$$

donde:

- $SCE = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$ ,
- $S_{yy} = \sum_{i=1}^n (y_i - \bar{y}_n)^2$ .

2. Mostrar la desigualdad de Bonferroni. Si  $E_1, \dots, E_k$  son eventos en un espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , entonces

$$Pr\left\{\bigcap E_j\right\} \geq 1 - \sum Pr\{E_j^c\}.$$

3. Considerar los datos de ingreso y escolaridad utilizados en los ejemplos de intervalos de confianza de las notas. Reportar intervalos simultáneos de confianza 95 % para las medias del ingreso por hora para 9, 15 y 19 años de escolaridad a) con el método de Bonferroni y b) con el método de Hotelling-Scheffé.
4. El conjunto de datos `airquality`, de paquete `datasets` de R contiene información sobre la calidad del aire en Nueva York registrada de Mayo a Septiembre de 1973 (se puede consultar más información con el comando `help(airquality)`). Para responder este ejercicio, descartar las observaciones con valores perdidos.
- a) Ajustar un modelo RLS para explicar el nivel de ozono como función del  $\log_2$  de la velocidad del viento. Reportar las estimaciones de los parámetros.
  - b) Mostrar una gráfica de dispersión de los datos utilizados para ajustar el modelo del inciso anterior, la recta de regresión ajustada y bandas de confianza 95 %. Anexar el código relacionado con el cómputo de las bandas de confanza.
5. (Sheater) Un estadístico colaboró en un proyecto de investigación con dos entomólogos. El análisis involucró el ajuste de modelos de regresión a grandes conjuntos de datos. Entre los tres escribieron y sometieron un manuscrito a una revista de entomología. El escrito contenía varias gráficas de dispersión mostrando la recta de regresión ajustada y las *bandas de confianza* 95 % para la verdadera recta de regresión calculadas con los IC individuales, así como los datos observados. Uno de los revisores del manuscrito hizo la siguiente observación:

*No puedo entender cómo el 95 % de las observaciones cae fuera de las bandas de confianza 95 % que se muestran en las figuras*

Dar una respuesta breve del porqué es posible que el 95 % de las observaciones caigan fuera de las bandas de confianza 95 % que se muestran en las figuras.

6. (Ross) Suponer que se tiene el siguiente conjunto de datos donde  $x$  representa la humedad de una mezcla fresca de un determinado producto y  $y$  la densidad del producto terminado.

$x$	5	6	7	10	12	15	18	20
$y$	7.4	9.3	10.6	15.4	18.1	22.2	24.1	24.8

Ajustar un modelo RLS a los datos anteriores y responder lo siguiente.

- Reportar la estimación puntual de  $\sigma^2$  e interpretar el resultado en cuanto a la utilidad del modelo RLS ajustado.
- Reportar el IC 90 % para  $\sigma^2$  con los cuantiles simétricos y su longitud.
- Indicar cuáles son los cuantiles que proporcionan el IC 90 % para  $\sigma^2$  de menor longitud.
- Reportar el IC 90 % para  $\sigma^2$  de menor longitud y compararlo con el intervalo del inciso a).

## Referencias

- Ross, S. (1987), *Introduction to Probability and Statistics for Engineers and Scientists*. Wiley: New York.
- Sheather, S. J. (2009), *A Modern Approach to Regression with R*. Springer: New York.