

- Fecha esperada de terminación:

1. Recuerde que la matriz  $H = X(X'X)^{-1}X'$  es simétrica e idempotente.
  - a) Muestre que cada uno de los elementos de la diagonal  $h_{ii}$ , llamado algunas veces “apalancamiento” de  $y_i$  en  $\hat{y}_i$ , ( $i = 1, \dots, n$ ) está entre 0 y 1.
  - b) Muestre que si  $X_{n \times q}$ , la suma de los  $h_{ii}$  es  $q$ .
  - c) Muestre que si  $y_i$  se reemplaza por  $y_i^* = y_i + 1$ , entonces  $\hat{y}_i^* = \hat{y}_i + h_{ii}$ .
2. Repita el análisis discutido en clase de los datos del score adaptativo de Gesell y la edad del niño cuando se emite la primer palabra. [Los datos los encuentran en *Comunidad* como `GesellScore.dat`.]
3. El siguiente juego de datos es sobre el endurecimiento de cemento Portland y famoso por su difícil modelación. Draper & Smith (1998) discuten el análisis de los datos en varias secciones de su texto. [Los datos los encuentran en *Comunidad* como `Hald.dat`.]

| variable | concepto  |
|----------|---|
| $x_1$    | Cantidad de tricalcio de aluminato, $3CaO \cdot Al_2O_3$ .                      |
| $x_2$    | Cantidad de tricalcio de silicato, $3CaO \cdot SiO_2$ .                         |
| $x_1$    | Cantidad de tricalcio de aluminio ferrito, $4CaO \cdot Al_2O_3 \cdot Fe_2O_2$ . |
| $x_1$    | Cantidad de dicalcio de silicato, $2CaO \cdot SiO_2$ .                          |
| $y$      | Calor en calorías por gramo de cemento.   |

Los regresores,  $x_1, x_2, x_3, x_4$  son medidos como porcentaje del peso de las *ollas* donde se hace el cemento.

| obs | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $y$   |
|-----|-------|-------|-------|-------|-------|
| 1   | 7     | 26    | 6     | 60    | 78.5  |
| 2   | 1     | 29    | 15    | 52    | 74.3  |
| 3   | 11    | 56    | 8     | 20    | 104.3 |
| 4   | 11    | 31    | 8     | 47    | 87.6  |
| 5   | 7     | 52    | 6     | 33    | 95.9  |
| 6   | 11    | 55    | 9     | 22    | 109.2 |
| 7   | 3     | 71    | 17    | 6     | 102.7 |
| 8   | 1     | 31    | 22    | 44    | 72.5  |
| 9   | 2     | 54    | 18    | 22    | 93.1  |
| 10  | 21    | 47    | 4     | 26    | 115.9 |
| 11  | 1     | 40    | 23    | 34    | 83.8  |
| 12  | 11    | 66    | 9     | 12    | 113.3 |
| 13  | 10    | 68    | 8     | 12    | 109.4 |

- a) Ajuste el modelo de regresión lineal múltiple en los 4 regresores  $x$ 's y comente. Calcule los residuales  $y - \hat{y}$ , los estandarizados, los *studentizados* interna y externamente y compárelos. Así también, calcule los estadísticos de Cook para cada una de las observaciones.
- b) Elimine la octava observación y repita el inciso anterior. Comente.
- c) Indique las diferencias más importantes entre ambos ajustes. Compare las estimaciones de  $\beta$ .