Universidad Técnica Federico Santa María
Departamento de Matemática.

Certamen N° 2 MAT-266 Análisis de Regresión

Profesor: Felipe Osorio. 23 de marzo de 2012
Ayudante: Claudio Henríquez.

1. (20 puntos) Considere el modelo $Y_i = \beta x_i + \epsilon_i$, donde $\{\epsilon_i\}$ son variables iid $N(0, \sigma^2 x_i^2)$. Obtenga el estimador máximo verosímil para β y su varianza.

2. (20 puntos) Considere el modelo de regresión lineal simple
$$Y_t = \alpha + \beta(x_t - \overline{x}) + \epsilon_t, \qquad t = 1, \dots, T,$$
donde $\{\epsilon_t\}$ son variables aleatorias iid $\mathcal{N}(0, \sigma^2)$ y $\overline{x} = \frac{1}{T}\sum_{t=1}^{T} x_t$. Muestre que
$$h_i = \frac{1}{T} + \frac{(x_i - \overline{x})}{\sum_{t=1}^{T}(x_t - \overline{x})^2}.$$

3. (20 puntos) Sea el modelo lineal $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, con $\boldsymbol{\epsilon} \sim \mathcal{N}_n(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. Muestre que

   a) $\widehat{Y}_i = (1 - h_i)\boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_{(i)} + h_i Y_i$,

   b) $1 + \boldsymbol{x}_i^T(\boldsymbol{X}_{(i)}^T\boldsymbol{X}_{(i)})^{-1}\boldsymbol{x}_i = 1/(1 - h_i)$,

   c) y usando el resultado en *b)*,
$$\frac{y_i - \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}_{(i)}}{\sqrt{1 + \boldsymbol{x}_i^T(\boldsymbol{X}_{(i)}^T\boldsymbol{X}_{(i)})^{-1}\boldsymbol{x}_i}} = \frac{y_i - \widehat{y}_i}{\sqrt{1 - h_i}}$$

4. (20 puntos) Considere las regresiones de Y sobre x para los datos a continuación, especificadas por $\mathrm{E}(Y) = \beta_0 x$ y $\mathrm{E}(Y) = \beta_1 x + \beta_2 x^2$. Obtenga $\widehat{\beta}_0$, $\widehat{\beta}_1$ y $\widehat{\beta}_2$. ¿Cuál de esos modelos es preferido?

Y	5	7	7	10	16	20
x	1	2	3	4	5	6

5. (20 puntos) Considere el conjunto de datos

x_i	1	2	3	4	5
Y_i	0.5	1.2	0.9	1.6	5.2

Se ajustó el modelo lineal

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad i = 1, \ldots, 5,$$

donde $\{\epsilon_i\}$ son variables aleatorias iid $N(0, \sigma^2)$, obteniendo los siguientes resultados:

Variable	Estimación	Error Est.	t	valor-p
Intercepto	-1.060	1.330	-0.797	0.484
x	0.980	0.401	2.444	0.092

además $s = 1{,}268$, $R^2 = 0{,}665$ y $F = 5{,}973$, se calculó también:

i	1	2	3	4	5
e_i	0.580	0.300	-0.980	-1.260	1.360
r_i	0.723	0.283	-0.864	-1.188	1.696
t_i	0.650	0.234	-0.814	-1.332	6.800
h_i	0.600	0.300	0.200	0.300	0.600

a) ¿Existe algún outlier? Justifique su respuesta.

b) Utilice estadísticas apropiadas para estudiar el rol de la observación 5.