MA3402-1. Estadística 2017.
Profesor: Raul Gouet.
Auxiliares: Diego Marchant y Raimundo Saona.

Auxiliar 12

Lunes 06 de Noviembre

C1. Poder del modelo lineal

Restringirse a dependencias lineales puede sonar muy fuerte, que tiene poco alcanze, que es capaz de modelar muy pocas dependencias. Argumente que las siguientes relaciones entre x e y pueden ser expresadas a través de un modelo lineal

$$
\begin{aligned}
y &= \frac{a}{b+cx}\,,\\
y &= ae^{-bx}\,,\\
y &= ab^{x}\,,\\
y &= \frac{x}{b+cx}\,,\\
y &= \frac{1}{1+e^{bx}}\,.
\end{aligned}
$$

Fuente: *Mathematical Statistics and data analysis*, Rice.

En esta misma línea, si tienen una matriz de datos X, ¿qué variables pueden tomar en cuenta para su modelo lineal?, ¿qué dependencia no podrían modelar?

C2. ¿Dónde realmente están?

Asuma que tiene tres objetos ubicados en línea recta, digamos a distancias p_1, p_2, p_3 de un punto de origen. Ahora bien, no conoce realmente estos valores, por lo que procede a realizar las siguientes mediciones

(1) Mide desde el origen las distancias hasta los objetos obteniendo y_1, y_2, y_3.

(2) Avanza hasta p_1 y mide, desde ahí, las distancias hasta los otros dos objetos obteniendo y_4, y_5.

(3) Por último, desde p_2 mide la distancia hasta p_3 obteniendo y_6.

Considere que todas sus mediciones tienen un error aditivo otorgado por su instrumento de medida.

- ¿Cómo interpreta la frase "suponga que su instrumento esté bien calibrado"?
- ¿Qué podría significar "errores sistemáticos" y "errores aleatorios"?
- Usando notación matricial, escriba el modelo lineal que le corresponde a la situación.
- Deje expresada la solución de mínimos cuadrados en este modelo.

C3. Estimación usando el modelo lineal

Asuma que tiene un modelo lineal estándar donde se cumple la homocedasticidad (con σ^2 conocido) $Y = X\beta + \epsilon$ y considere $\widehat{\beta}$ el vector resultante de mínimos cuadrados. Asuma que $\beta \in \mathbb{R}^2$ y así $X = (1, x)$, considerando x_0 un punto no observado (ie, x_0 no tiene el valor de alguna coordenada de x) y considere la estimación

$$\widehat{y}_0 = (1, x_0)\widehat{\beta}$$

- Calcule la varianza de $\widehat{y}_0$.
- Grafique la forma que tiene esta última en función de $x_0 - \overline{x}_n$.
- Encuentre el x_0 que tendría menor varianza.
- Entregue un intervalo de $(1-\alpha)\,\%$ de confianza, asumiendo hipótesis de normalidad.

A1. ¡Aplique lo aprendido!

Considere dos tratamientos médicos A y B. Científicos han hecho un experimento tratando de relacionar el nivel de respuesta de los pacientes ante los dos tratamientos y han obtenido la siguiente tabla de datos

i	1	2	3	4	5	6	7	8	9	10
x_i	1.9	0.8	1.1	0.1	-0.1	4.4	4.6	1.6	5.5	2.0
y_i	0.7	-1.0	-0.2	-1.2	-0.1	3.4	0.0	0.8	3.7	3.4

- Determine el modelo matricial que le corresponde al problema y explicite el pensamiento que aplica.
- Determine el valor de
$$\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\sigma}^2, Var(\widehat{\beta}_1), Var(\widehat{\beta}_2)\,.$$
¿Qué estimador tomaría para $\widehat{\sigma}^2$?
- Determine el valor de $Corr(\widehat{\beta}_1, \widehat{\beta}_2)$.
- Suponga que quiere estimar $\theta = 3\beta_1 - 2\beta_2 + 5$. Encuentre el estimador $\widehat{\theta}$ insesgado de mínima varianza. Determine su valor y su error cuadrático medio.
- Ahora suponga que se quiere estimar $\theta = 3\beta_1 + c_2\beta_2$, con $c_2 \in \mathbb{R}$. Tomando $\widehat{\theta}$ el estimador insesgado de mínima varianza, ¿qué valor de c_2 minimiza su varianza?
- Asuma que llega un nuevo paciente que tiene un nivel de respuesta al tratamiento A de 2, ¿cuál es su valor estimado para la respuesta al tratamiento B?, ¿qué varianza tiene esta predicción?
- Dada esta data, ¿para qué valor de x usted puede predecir con la mínima varianza?