Ingeniería Matemática
FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE
Estadística MA 3402, 14/11/17, Prof. R. Gouet.

Guía de Ejercicios #3

1. Un proceso determinístico $z_0, z_1, \ldots, z_n$ se rige por la recurrencia $z_{i+1} = \lambda z_i$, donde λ es una constante conocida. Los valores de z_i no puede ser observados sin error y las observaciones Y_i se rigen por el modelo
$$Y_i = z_i + \epsilon_i,$$
donde los ϵ_i son errores centrados, no correlacionados con varianza común. Determine los estimadores de mínimos cuadrados de $z_0, z_1, \ldots, z_n$ (los parámetros del modelo). Debe tomar en cuenta la recurrencia entre los z_k al momento de plantear el problema de mínimos cuadrados.

2. Un topógrafo realiza mediciones $Y_1, Y_2, \ldots, Y_k$ de los ángulos $\theta_1, \ldots, \theta_k$ de un polígono de k lados. Suponiendo que las mediciones están sujetas a errores ϵ_i, centrados, no correlacionados y con varianzas σ^2, obtenga los estimadores de mínimos cuadrados de los ángulos θ_i, en los siguientes casos:
   a) ignorando que los ángulos suman $(k-2)\pi$ e
   b) incorporando al modelo la restricción $\sum_{i=1}^{k} \theta_i = (k-2)\pi$.
   c) Compare las varianzas de los estimadores en ambos casos e indique, de acuerdo con la comparación, cual sería el procedimiento más adecuado.

3. Considere el modelo lineal simple $Y_i = \alpha + \beta x_i + \epsilon_i$, $i = 1, \ldots, n$, con las hipótesis habituales. Para estimar β se ha sugerido usar
$$\tilde{\beta} = \frac{1}{n-1} \sum_{i=2}^{n} \frac{Y_i - Y_{i-1}}{x_i - x_{i-1}}.$$
   a) Indique qué justificación o motivación puede tener dicho estimador.
   b) Averigüe si $\tilde{\beta}$ es insesgado.
   c) Compare $\tilde{\beta}$ con el estimador de mínimos cuadrados en términos de las varianzas respectivas.

4. Se sabe que los puntos (x_i, y_i) con $x_i > 0, i = 1, \ldots, n$ están sobre la hipérbola de ecuación $xy = \alpha$. Los valores x_i se pueden fijar a voluntad y se consideran no aleatorios pero los y_i sólo se pueden observar con error. Observamos entonces $Z_i = y_i + \epsilon_i, i = 1, \ldots, n$, donde los ϵ_i son errores iid centrados con varianza común σ^2. Por lo tanto, tenemos el siguiente modelo
$$Z_i = \frac{\alpha}{x_i} + \epsilon_i,\ i = 1, \ldots, n.$$
   a) Encuentre el estimador de mínimos cuadrados para α.
   b) Determine el estimador insesgado para σ^2 basado en los residuos.

5. Considere el modelo lineal simple $Y_i = \alpha + \beta x_i + \epsilon_i$, $i = 1, \ldots, n$ (errores centrados, no correlacionados y varianzas σ^2). Por un error del encargado del sistema de información, se han extraviado los valores de la variable X para las observaciones $m+1, \ldots, n$ pero se conservan los valores de Y. Se propone "reconstruir" los datos extraviados sustituyendo los x_i perdidos por la media $\overline{x}$ de los x_i, para $i = 1, \ldots, m$.

a) Calcule los estimadores de mínimos cuadrados $\widetilde{\alpha}$ y $\widetilde{\beta}$ de α y β, obtenidos a partir del conjunto de datos reconstituidos.

b) Compare $\widetilde{\alpha}$ y $\widetilde{\beta}$ con los correspondientes estimadores basados en los datos completos, en términos de sesgo y varianza.

c) Compare $\widetilde{\alpha}$ y $\widetilde{\beta}$ con los correspondientes estimadores basados en los primeros m datos (es decir, se ignoran los Y_i para los cuales se extraviaron los x_i), en términos de sesgo y varianza.

6. Considere el modelo lineal $Y = X\beta + \epsilon$, con errores centrados, de varianza constante $\sigma^2 < \infty$ y no correlacionados. Suponga la matriz X no aleatoria, con n filas y p columnas y tal que $rg(X) \leq p$. Sea $\hat{\beta}$ cualquier solución del sistema de ecuaciones normales $X'X\beta = X'Y$.

   Sea $\phi = c'\beta = \sum_{i=1}^{n} c_i\beta_i$. Se dice que ϕ es estimable (linealmente) si existe b tal que $\tilde{\phi} := b'Y$ es estimador insesgado de ϕ. Sea $\hat{\phi} = c'\hat{\beta}$.

   a) Compruebe que siempre existe solución del sistema de ecuaciones normales.

   b) Muestre que ϕ es estimable si y solo si existe $b \in \mathbb{R}^n$ tal que $b'X = c'$ y que toda función ϕ es estimable cuando $rg(X) = p$.

   c) Demuestre que $\hat{\phi}$ es estimador insesgado de ϕ.

   d) Muestre que no toda ϕ es estimable. En particular, no todas las componentes de β son estimables.

   e) Sea $\hat{b}$ la proyección de b en $< X >$ (el sev engendrado por las columnas de X). Pruebe que $\hat{b}'Y$ es estimador insesgado de ϕ.

   f) Sea $\tilde{b} \in < X >$ tal que $\tilde{b}'Y$ es insesgado para ϕ. Verifique que $\tilde{b} = \hat{b}$.

   g) Demuestre que $Var(\hat{b}'Y) \leq Var(b'Y)$ y compruebe que $\hat{\phi} = \hat{b}'Y$ para concluir que $Var(\hat{\phi}) \leq Var(\tilde{\phi})$ (esto es el teorema de Gauß-Markov para el modelo de rango incompleto.)

7. Sea el modelo lineal

$$Y_i = \alpha + \beta x_i + \gamma z_i + \delta(x_i/z_i) + \epsilon_i, \ i = 1, \ldots, 8$$

   donde los ϵ_i son va centradas no, correlacionadas, con varianza constante. Investigue qué parámetros del modelo son (linealmente) estimables, de acuerdo con la definición del ejercicio anterior.

8. Considere la matriz de datos X tal que sus columnas $X_1, ..., X_p$ son ortogonales.
   Muestre que la matriz de covarianzas del estimador de mínimos cuadrados es diagonal.

9. Generalizaremos el modelo lineal para el caso en que X es variable también, sólo para el caso unidimensinal. Considere X, Y variables aleatorias con esperanzas μ_x, μ_y y varianzas σ_x^2, σ_y^2 respectivamente, también $Cov(X,Y) = \sigma_{x,y}$. Muestre que $\widehat{\alpha}, \widehat{\beta}$ tales que minimizan $\mathbb{E}((Y - (\alpha + \beta X))^2)$ están dados por

$$\widehat{\beta} = \frac{\sigma_{x,y}}{\sigma_x^2}$$
$$\widehat{\alpha} = \mu_y - \beta\mu_x$$

   También, reduzca lo que más pueda el término

$$\frac{Var(Y) - Var(Y - \widehat{Y})}{Var(Y)}$$

10. Considere el modelo $Y_i = \alpha + \beta x_i + \epsilon_i$, donde $\{x_1, ..., x_n\} \subseteq [-1, 1]$. Suponga que usted puede diseñar el experimento de tal forma que tiene la libertad de situar $x_i \in [-1, 1]$ como usted desee,

   ¿Cómo deben disponerse los puntos $x_1, ..., x_n$ para minimizar $Var(\widehat{\beta})$?

11. Considere el modelo $Y = X\beta_x + \epsilon$. Se definen las variables $u_{ij} := k_i x_{ij}$ y el nuevo modelo lineal $Y = U\beta_u + \epsilon$. Muestre que

$$\widehat{Y}_x = \widehat{Y}_u$$

   Y exprese $\widehat{\beta}_u$ en términos de $\widehat{\beta}_x$.

12. Considere el modelo $Y = X\beta + \epsilon$ con 20 observaciones, donde su supone independencia y homocedasticidad del vector ϵ. Asuma que obtuvo

$$\widehat{\sigma}^2 = 1{,}5 \quad , \quad \widehat{\beta} = (1, 2, 1)^t$$

   con $(X^t X)^{-1} = (2, 1, 0; 1, 2, 1; 0, 1, 2)$.

   Diseñe un test de hipótesis para

$$H_0 : \beta_1 = 4 - 2\beta_3 \wedge \beta_2 = 0{,}5 + \beta_3$$

   para $\alpha = 0{,}05$.