valor más pequeño del cociente

$$h(X_1, \dots, X_N) \equiv \frac{\sum_{i=1}^{N} X_i^2}{N \sum_{i=1}^{N} (X_i - \bar{X})^2}$$

Tarea: demuestre que para un tamaño de muestra fijo N la funcion h se minimiza si $\bar{X} = 0$.

La conclusión es que $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores lineales e insesgados para β_0 y β_1, cuyas varianzas dependen de los valores de la variable explicativa X. El siguiente teorema establece que $\hat{\beta}_0$ y $\hat{\beta}_1$ son estimadores; que dentro de la clase o conjunto de los estimadores lineales insesgados para β_0 y β_1, tienen la varianza más pequeña

TEOREMA (GAUSS-MARKOV)

Sean $Y_1, \dots, Y_N$ v.a. tales que

(i) $\mathbb{E}(Y_i) = \beta_0 + \beta_1 x_i$; $i = 1, 2, \dots, N$,

(ii) $VAR(Y_i) = \sigma^2$; $i = 1, 2, \dots, N$,

(iii) $COV(Y_i, Y_j) = 0 \quad \forall i \neq j \quad (i, j = 1, 2, \dots, N)$, donde $x_1, \ldots, x_N$ son constantes conocidas. Entonces los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$ son los Mejores Estimadores Lineales Insesgados (MELI) de β_0 y β_1, en el sentido de varianza mínima.

<u>NOTA</u>: En el idioma Inglés se dice que $\hat{\beta}_0$ y $\hat{\beta}_1$ son los BLUE (Best Linear Unbiased Estimators)

NOTA: El teorema no establece que $\hat{\beta}_0$ y $\hat{\beta}_1$ son óptimos en general. Lo que establece es que $\hat{\beta}_0$ y $\hat{\beta}_1$ son los mejores, dentro de una clase particular de estimadores, la clase ó conjunto de los estimadores lineales e insesgados.

Si quisieramos comparar $\hat{\beta}_0$ y $\hat{\beta}_1$ con los estimadores que se obtienen con otros métodos estadísticos, podríamos pensar en estimar β_0 y β_1 vía momentos o vía Máxima Verosimilitud. El método de momentos no funcionará, toda vez que este supone que las variables aleatorias que se observan son idénticamente distribuidas, que en nuestro caso no se tiene ($\mathbb{E}(Y_i)$ no es igual a $\mathbb{E}(Y_j)$).

Entonces pensaremos como estimar β_0 y β_1 vía el método de Máxima Verosimilitud. Lo anterior requiere tener la distribución conjunta para las variables que se observan.

Como la variable aleatoria ε describe el efecto de todos los factores, distintos de la variable explicativa, que no están fijos en el estudio, no parecería verosimil que tome valores en un conjunto finito ó numerable, así que suena razonable asumir que ε es una v.a. continua.

Gauss estudió el comportamiento empírico de los errores en la estimación de distancias a objetos en la boveda celeste, a partir de ese análisis sugirió que la distribución debería ser unimodal y simétrica con media 0. En particular si se piensa en ε como la suma infinita de efectos que no están bajo control en el estudio (y que no tienen relación con X) podemos conectar esta concepción de ε con el Teorema Central del Límite, que establece que bajo algunas condiciones, las sumas de variables aleatorias tienen distribución "cercana" a ser Normal.

MODELO DE REGRESION LINEAL NORMAL

$$Y = \beta_0 + \beta_1 X + \varepsilon,$$

donde ε es una v.a. con distribución $N(0, \sigma^2)$.

Los datos son generados con la relación

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad ; \quad i = 1, 2, \ldots, N,$$

donde $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ son variables aleatorias idénticamente distribuidas e independientes según el modelo Normal$(0, \sigma^2)$.

Esta estructura conduce a la siguiente formulación alternativa del Modelo Lineal Normal, el cual está vinculado con la discusión en las primeras páginas de estas notas

" $Y_1, Y_2, \ldots, Y_N$ son v.a. independientes y tales que $Y_i | X_i = x_i \sim$ Normal$(\beta_0 + \beta_1 x_i, \sigma^2)$ "

Bajo esta premisa, la verosimilitud (la función de densidad conjunta de $Y_1, \ldots, Y_N$ condicional a $X_1 = x_1; \ldots; X_N = x_N$ es

$$L(\beta_0, \beta_1, \sigma^2) = f_{Y_1 \ldots Y_N | X_1 \ldots X_N}(y_1, \ldots, y_N | x_1, \ldots, x_N, \beta_0, \beta_1, \sigma^2)$$

$$= \prod_{i=1}^{N} f_{Y_i|X_i}(y_i | x_i, \beta_0, \beta_1, \sigma^2)$$

$$= \prod_{i=1}^{N} \left\{ (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}\left[y_i - (\beta_0 + \beta_1 x_i)\right]^2} \right\}$$

$$= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \beta_0 - \beta_1 x_i)^2}$$