CO6341: Métodos de la Estadística

Problemas de Estimación

- **Problema de Inferencia Estadística**: Se analizan datos que provienen de una distribución de probabilidades desconocida y se debe realizar algún tipo de inferencia sobre la distribución.
- **Problemas reales:** Existe un número infinito de distribuciones que podrían haber generado los datos. Analizando los datos se intenta conocer la distribución desconocida para realizar inferencia acerca de la distribución. Se quiere determinar la *verosimilitud relativa* que cada distribución posible tiene de ser la correcta.
- **Parámetros**: Valores que no conocemos y que definen la distribución que generó los datos.

Ejemplo: La distribución exponencial se utiliza a menudo para representar la distribución del tiempo que transcurre antes de la ocurrencia de un suceso, por ejemplo, la duración de un bombillo.

Sea X una variable aleatoria con parámetro β desconocido.

$$X \sim exp(\beta)$$

La función de densidad de probabilidades (f.d.p.) de X viene dada por:

$$f(X|\beta) = \begin{cases} \beta e^{-\beta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

donde $\beta > 0$ y $E[X] = 1/\beta$, $V[X] = 1/\beta^2$

Problemas de estimación (Cont.)

- A partir de valores observados de la duración de un conjunto de bombillos, es posible hacer inferencia sobre el parámetro desconocido β por alguna de estas opciones:
  - Especificar el mejor valor para β
  - Especificar un intervalo
  - Especificar una cota superior
- **Otro ejemplo:** Valores de una distribución normal con μ y σ^2 desconocidos. Supongamos que se tienen observaciones de las estaturas de los individuos de una muestra aleatoria de una población que se asume normal (μ,σ^2) donde μ y σ^2 son los parámetros de la distribución. Es posible hacer inferencia sobre los valores de μ y los valores de σ^2.
- **Espacio paramétrico**: Es el conjunto Ω de todos los valores posibles del parámetro θ o de un vector de parámetros $(\theta_1, \theta_2, \ldots, \theta_n)$. En el caso exponencial Ω es el conjunto de todos los reales positivos; en el caso de la distribución normal, $-\infty < \mu < \infty$ y $\sigma^2 > 0$.

Distribución a priori (inicial) y distribución a posteriori (final)

- Distribución a priori

  Supóngase que se tienen o se tendrán a futuro observaciones de $f(x|\theta)$. En muchos casos antes de disponer de observaciones de $f(x|\theta)$ el experimentador o estadístico tiene conocimientos previos acerca de dónde es probable que se encuentre el valor de θ en el espacio paramétrico Ω. Para ello se construye una distribución de probabilidad θ en el conjunto Ω. Esta distribución se denomina distribución inicial o distribución a priori de θ.

  1. Hay un grupo de estadísticos (cada vez mayor!) que piensa que siempre se puede elegir una distribución inicial para θ que tiene caracter subjetivo. Ejemplo: Suponer una distribución de probabilidades que represente la creencia subjetiva del valor más probable para θ. Este grupo se adhiere a la Estadística Bayesiana.
  2. Hay otro grupo de estadísticos que opina que es necesario conocer extensa información previa para proponer una distribución inicial sobre θ. Ejemplo: Sea $\theta=$ la proporción de artículos defectuosos en un lote de manufacturas. θ es desconocido pero existe información de lotes anteriores que puede ser utilizada para construir una f.d.p para θ. En ambos casos, la estimación usando métodos Bayesianos es aplicable.

  Los métodos de estimación por *máxima verosimilitud* no están basados en asignar una distribución inicial para θ.

Distribuciones a priori discretas y contínuas

- Si θ sólo puede tomar un número finito de valores la f.d.p. $\xi(\theta)$ será una dist. discreta.
- Si θ puede tomar cualquier valor en la recta real entonce la f.d.p $\xi(\theta)$ será una dist. contínua.

Ejemplos

1 Sea $\theta =$ probabilidad de obtener una cara al lanzar una moneda (θ desconocida). Supongamos que se sabe que la moneda es equilibrada ó tiene dos caras. Entonces $\theta = 1/2$ ó $\theta = 1$.

Si la probabilidad inicial de que la moneda sea equilibrada es p, entonces la f.d.p. inicial de θ es:

$$\xi(1/2) = p$$
$$\xi(1) = 1 - p$$

2 Sea $\theta =$ proporción de artículos defectuosos (θ es deconocida). Supongamos que se asigna una distribución inicial uniforme a θ en el intervalo (0,1). La f.d.p. inicial de θ es:

$$\xi(\theta) = \begin{cases} 1 & 0 < \theta < 1 \\ 0 & en\ otro\ caso \end{cases}$$

3 Parámetro de una exponencial: Sa va a observar la duración de cierto tipo de lámparas flourescentes. Se asume que la distribución de la duración de cualquier lámpara es exponencial con parámetro β donde:

- β es desconocido
- Por experiencia previa se considera que la distribución inicial de β es gamma con media 0.0002 y desviación típica 0.0001.

Supóngase que la distribución inicial gamma tienen parámetros α_0 y β_0. En general $X \sim \text{gamma}(\alpha, \beta)$ si X tienen una f.d.p. de la forma:

$$f(x|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

donde $\Gamma(\alpha)$ = Función Gamma, la cual se define como:
$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$

Notar que:

- $\Gamma(1) = 1$; si $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$; si n es entero positivo $\Gamma(n) = (n-1)!$
- Además $\int_{-\infty}^{\infty} f(x|\alpha,\beta)dx = 1$ por ser una función de densidad.

La esperanza y la varianza de X se calculan como: $E(X) = \frac{\alpha}{\beta}$ y $V(X) = \frac{\alpha}{\beta^2}$.

Si $E(\beta) = \frac{\alpha_0}{\beta_0} = 0{,}0002$ y $V(\beta)^{1/2} = \frac{\alpha_0^{1/2}}{\beta_0} = 0{,}0001$

$\frac{\alpha_0}{\beta_0^2} = 0{,}00000001 \Rightarrow \frac{0{,}0002}{\beta_0} = 1x10^{-8} \Rightarrow \beta_0 = 20000$ De aquí se deduce

que $\alpha_0 = 4$; por lo tanto la f.d.p. a priori para β es:

$$\xi(\beta) = \begin{cases} \frac{(20000)^4}{3!}\beta^3 e^{-20000\beta} & \beta > 0 \\ 0 & \beta \leq 0 \end{cases}$$

Distribución a posteriori (o distribución final)

Sea $x_1, x_2, \ldots, x_n$ una muestra aleatoria de una distribución con f.d.p. $f(x|\theta)$.

- θ es deconocido con f.d.p. a priori $\xi(\theta)$ con $\theta \in \Omega$.
- La f.d.p. conjunta $f_n(x_1, x_2, \ldots, x_n|\theta) = f(x_1|\theta)\ldots f(x_n|\theta) = f(_n(\mathbf{x}|\theta)$ donde $\mathbf{x}$ es un vector.
- $f_n(\mathbf{x}|\theta)$ es la f.d.p conjunta condicional para un valor dado de θ.
- Al multiplcar $f_n(\mathbf{x}|\theta).\xi(\theta)$ se obtiene la f.d.p. conjunta $n+1$ de $\mathbf{x}$ y θ.
- La distribución marginal de $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ se obtiene integrando la conjunta sobre todos los valores de θ: $g_n(\mathbf{x}) = \int_{\boldsymbol{\Omega}} \mathbf{f_n}(\mathbf{x}|\theta)\xi(\theta)\mathbf{d}\theta$.
- La distribución de probabilidad condicional de θ dado $\mathbf{x}$ se denota como $\xi(\theta|\mathbf{x}) = \frac{\mathbf{f_n}(\mathbf{x}|\theta)\xi(\theta)}{\mathbf{g_n}(\mathbf{x})}, \;\; \theta \in \boldsymbol{\Omega}$. Notar que $g_n(\mathbf{x})$ no depende de $\theta \rightarrow \xi(\theta|\mathbf{x}) \propto \mathbf{f_n}(\mathbf{x}|\theta)\xi(\theta)$. Además $\int_{\Omega} \xi(\theta|\mathbf{x})\mathbf{d}\theta = \mathbf{1}$, ya que $\xi(\theta|\mathbf{x})$ es una f.d.p. para θ. (Teorema de Bayes para parámetros y muestras aleatorias).