Práctico 1

Friday, August 24, 2018 6:20 PM

Resumen de conceptos clave

En este práctico vemos técnicas de reducción de datos. En la reducción de datos, lo que interesa es sintetizar información sobre parámetro desconocido (θ), a partir de los datos de una muestra.

1)* Estadístico
2)* Familia paramétrica
3)* Familias exponenciales
4)* Estadístico suficiente (noción)
5)* Ppio. de suficiencia
6)* Estadístico suficiente (def.)
7)* Teorema de factorización.
8)* Estadístico mínimo suficiente.
9)* Estadístico auxiliar.

—— x ——

1) Un estadístico es una función de una muestra $X_1, .., X_n$. Se nota como $T(X)$.
Ejemplo: · media muestral $\bar{X}_n = \frac{X_1 + ... + X_n}{n}$
· varianza muestral
· mínimo muestral $= \min\{X_1, .., X_n\}$
· máximo muestral $= \max\{X_1, .., X_n\}$

$X = x$
realización.

2) Flia paramétrica $f(x|\theta)$, $\theta \in \Theta$.
Ej: $N(X|\mu, \sigma^2)$.
· más grande: flias exponenciales.

3) Una flia exponencial es una flia paramétrica

de la sgte forma:

$$f(x|\theta) = h(x)c(\theta).\exp\left(\sum_{i=1}^{k} w_i(\theta)t_i(x)\right).$$

- Casos particulares: normal, gamma, beta (cont.)
- binomial, Poisson (cont.)

· Ejemplo:

$$f(x|p) = P_r(X=x) = \binom{n}{x} p^x (1-p)^{n-x} = \underbrace{\binom{n}{x}}_{h(x)} e^{\overbrace{x}^{t_1(x)}\overbrace{\log p}^{w_1(p)}} . e^{\overbrace{(n-x)}^{t_2(x)}\overbrace{\log(1-p)}^{w_2(p)}}.$$

4) Un estadístico suficiente para θ captura toda la info. sobre θ que está contenida en la muestra.

5) Si $T(X)$ es suficiente para θ, entonces cualquier inferencia sobre θ, debe depender de X sólo a través del valor $T(X)$.

· Es decir: x e y son dos muestras con $T(x)=T(y)$ $\Rightarrow$ la inferencia sobre θ debe ser $=$ ya sea que observemos $X=x$, ó $X=y$.

6) $T(X)$ es un estadístico suficiente para θ si la distribución condicional de la muestra X dado el valor de $T(X)$ no depende de θ.

7) Teo. factorización: Sea $f(x|\theta)$ la pdf ó pmf de una muestra X. Un estadístico $T(X)$ es suficiente para θ sii $\exists$ $g(t|\theta)$, $h(x)$ tq. $\forall$ muestras x y valores del parámetro $\theta \in \Theta$:

$$f(x|\theta) = g(T(x)|\theta) \cdot h(x).$$

Ejercicio 1

· $X \sim N(0,\sigma^2)$. ¿Es $|X|$ un estadístico suf.?

• **Idea**: No necesitamos conocer la pdf de $|X|$ para aplicar el Teo. factorización.

• **Sol**:

$$f(x|\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} = \underbrace{\frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{|x|^2}{2\sigma^2}}}_{g(|x| \,|\, \sigma^2)} \cdot \underbrace{1}_{h(x)}.$$

$\Rightarrow$ Sí, $|X|$ es un est. suficiente.

* **Obs**: ("Half-normal" distribution – media-normal).

· Vamos a calcular la distribución de $Y = |X|$, suponiendo que $X \sim N(\mu, \sigma^2)$.

Vocabulario:

Si $X \sim N(\mu, \sigma^2) \Rightarrow Y = |X|$: folded normal

Si $X \sim N(0, \sigma^2) \Rightarrow Y = |X|$: half-normal

· Sea $Z \sim N(0,1) \Rightarrow \phi(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-z^2/2}$; $z \in \mathbb{R}$ (densidad)

(distribución) $\Phi(z) = \int_{-\infty}^{z} \phi(x)\,dx = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$. $z \in \mathbb{R}$

Tenemos que:

$$\begin{aligned} F(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(|X| \leq y) \\ &= \mathbb{P}(|\mu + \sigma Z| \leq y) \\ &= \mathbb{P}(-y \leq \mu + \sigma Z \leq y) \\ &= \mathbb{P}\left(\frac{-y-\mu}{\sigma} \leq Z \leq \frac{y-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(\frac{-y-\mu}{\sigma}\right) \end{aligned}$$

Recordando que: $\Phi(-z) = 1 - \Phi(z)$

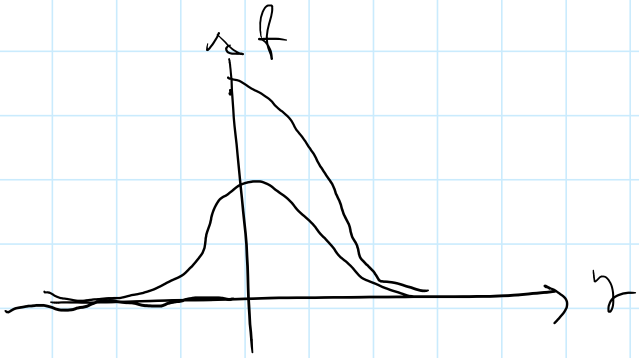
$$= \Phi\left(\frac{y-\mu}{\sigma}\right) + \Phi\left(\frac{y+\mu}{\sigma}\right) - 1.$$

$$F'(y) = f(y) = \frac{1}{\sigma}\left[\phi\left(\frac{y-\mu}{\sigma}\right) + \phi\left(\frac{y+\mu}{\sigma}\right)\right].$$

$$f(y) = f_Y(y) = \frac{1}{\sigma}\left[f\left(\frac{y-\mu}{\sigma}\right) + f\left(\frac{y+\mu}{\sigma}\right)\right].$$

$\Rightarrow$ Caso para $\mu = 0$: $f(y|\sigma^2) = \frac{1}{\sigma} \cdot 2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{y^2}{2\sigma^2}}$

$$\boxed{f(y|\sigma^2) = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sigma} \cdot e^{-y^2/2\sigma^2}}, \; y \geq 0$$



Ejercicio 2

Calculamos la pdf conjunta: $\begin{cases} 1 & \text{si } x_n \geq i\theta \\ 0 & \text{en otro caso} \end{cases}$

$$f(x_1, \dots, x_n | \theta) = e^{i\theta - x_1} \cdot \mathbb{1}_{(i\theta, +\infty)}(x_1) \cdots e^{i\theta - x_n} \cdot \mathbb{1}_{(i\theta, +\infty)}(x_n)$$

$$= \underbrace{e^{\frac{\theta n(n+1)}{2}}}_{} \cdot \underbrace{e^{-\sum_i x_n}}_{h(x)} \cdot \underbrace{\prod_i \mathbb{1}_{(i\theta, +\infty)}(x_i)}_{}$$

$$i\theta < x_i \iff \frac{x_i}{i} > \theta \cdot \quad \mathbb{1}_{(\theta, +\infty)}(T(x))$$

Ej. $n = 3$: x_1, x_2, x_3 t.q. $\frac{x_1}{1} < \frac{x_2}{2} < \frac{x_3}{3}$. $\quad g(T(x)|\theta)$

$\theta \quad \frac{x_1}{1} \quad \frac{x_2}{2} \quad \frac{x_3}{3}$

$$\begin{cases} g(T(x)|\theta) = e^{\theta(n+1)n/2} \cdot \mathbb{1}_{(\theta, +\infty)}(T(x)) \\ h(x) = \exp\left(-\sum x_i\right). \end{cases}$$

$\Rightarrow T = \min_i (X_i / i)$ es un est. suficiente.