Examen de Inferencia Estadística

10 de mayo de 2018

1. Sea X una variable aleatoria absolutamente continua, con función de densidad igual a

$$f(x; \alpha) = e^{-(x-\alpha)} \mathbb{1}_{\mathbb{R}^+}(x - \alpha)$$

donde $\alpha \in \mathbb{R}$. Sean $X_1, \ldots, X_n$ las variables aleatorias muestrales correspondientes a una muestra aleatoria simple de tamaño n.

- Halle el estimador máximo verosímil de α.
- Compruebe que $X_{(1)}$ es un estadístico suficiente. Si le informan además que es completo, halle el estimador UMVU de α.
- Calcule el error cuadrático medio del estimador UMVU.
- Estudie la distribución del estimador UMVU, proporcionando su función de densidad.

2. Sea X una variable aleatoria discreta, cuya función de densidad discreta (o función de probabilidad) es

$$f(x; p) = p^x (1 - p)^{(1-x)} \mathbb{1}_{\{0,1\}}(x)$$

donde $p \in (0, 1)$. Sean $X_1, \ldots, X_n$ las variables aleatorias muestrales correspondientes a una muestra aleatoria simple de tamaño n de X.

- Si suponemos una distribución a priori uniforme en $(0, 1)$, halle el estimador de Bayes de p, correspondiente a la pérdida cuadrática. ¿Es un estimador insesgado?
- Halle también el estimador MLE de p y compárelo con el anterior, valorando pros y contras de ambos estimadores.
- Halle la distribución asintótica del MLE.
- Para una muestra de tamaño $n = 2$, si sabemos que $U = X_1 + X_2$ es un estadístico suficiente y completo para p (para el modelo) halle todos los estimadores insesgados de p^2 ¿cuál de ellos es el UMVU?

3. Contestar verdadero o falso, en esta misma hoja.

F - Un estadístico suficiente solo existe en los modelos exponenciales en general, que incluyen entre otras, a la distribución normal, binomial y exponencial.
V - Si el error cuadrático medio de un estimador existe y tiendo a cero a medida que el tamaño muestral tiende a infinito entonces es un estimador consistente.
F - Si el sesgo de un estimador para un parámetro es cero y, además, es un estimador consistente, entonces es el estimador UMVU de dicho parámetro.
F - En el caso escalar, el riesgo cuadrático es la diferencia entre el valor medio que toma el estimador menos el parámetro al cuadrado.
F - En el caso escalar, el estimador UMVU, cuando existe, es un estimador eficiente (su varianza alcanza la cota de Cramér-Rao).
V - No siempre existen estimadores insesgados para un determinado parámetro.
V - En un modelo de la família exponencial general, con una parametrización conveniente, siempre existe un estimador que alcanza la cota de Cramér-Rao (escalar o su equivalente multidimensional) de forma insesgada.
F - Dato un estimador completo, a partir de un estimador insesgado, el Teorema de Rao-Blackwell nos permite obtener otro estimador insesgado y de varianza mínima.

NOTA: Cada apartado de los dos problemas vale 1 punto. Las respuestas deben ser debidamente justificadas, exceptuando las preguntas tipo test finales. El test vale 2 puntos de la nota final del examen. La nota del test, sobre 2 puntos, es igual a 0,25 multiplicado por el número de aciertos menos fallos.

1.- Sea X una v.a. absolutamente continua con función de densidad de probabilidad:

$$f(x,\alpha) = e^{-(x-\alpha)}\, \mathbb{1}_{[0,\infty)}(x-\alpha)$$

$\alpha > 0$ $\quad X_1, \ldots, X_n$ iid X

* Halle el MLE:

$$L_x(\alpha) = \prod_{i=1}^{n} f(x_i,\alpha) = \prod_{i=1}^{n} \left\{ e^{-(x_i-\alpha)}\, \mathbb{1}_{[0,\infty)}(x_i-\alpha) \right\} =$$

$$= e^{-\sum_{i=1}^{n}(x_i-\alpha)}\, \mathbb{1}_{[0,\infty)}(x_{(1)}-\alpha)$$

$$= e^{-n(\bar{x}_n-\alpha)}\, \mathbb{1}_{[0,\infty)}(x_{(1)}-\alpha)$$

donde $x_{(1)} = \min\{x_1, \ldots, x_n\}$ $\quad \bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$\mathbb{1}_{[0,\infty)}(x_{(1)}-\alpha) = \begin{cases} 1 & \alpha \leq x_{(1)} \\ 0 & \alpha > x_{(1)} \end{cases}$$

por otra parte

$e^{-n(\bar{x}_n-\alpha)} = e^{-n\bar{x}_n} e^{n\alpha}$ es, claramente, monótona creciente en α. Por tanto

$$L_x(\alpha) \leq e^{-n\bar{x}_n} e^{n x_{(1)}}$$

y el MLE es $\boxed{\alpha^* = X_{(1)}}$, el mínimo de las v.a. muestrales

* Por Neyman-Fisher:

$$\left[f(x_1,\dots,x_n,\alpha) = \underbrace{e^{-n\bar{x}_n}}_{h(x_1 \dots x_n)} \; \underbrace{e^{n\alpha}\,\mathbb{1}_{[0,\infty)}(x_{(1)}-\alpha)}_{g(x_{(1)},\alpha)} \right]$$

la conjunta factoriza de tal manera que permite concluir que

- $X_{(1)}$ es un estadístico suficiente

* Hemos de estudiar, en primer lugar, la distribución de $X_{(1)}$.

$$F_{X_{(1)}}(u) = P(X_{(1)} \leq u) = 1 - P(X_{(1)} > u) = 1 - P([X_1 > u] \cap \dots \cap [X_n > u])$$
$$= 1 - (P(X > u))^n = 1 - (1 - F_X(u))^n$$

pero
$$F_X(u) = \begin{cases} 0 & u < \alpha \\ 1 - e^{-(u-\alpha)} & u \geq \alpha \end{cases}$$

por tanto:
$$F_{X_{(1)}}(u) = \begin{cases} 1 - (1-0)^n & u < \alpha \\ 1 - (1 - (1 - e^{-(u-\alpha)}))^n & u \geq \alpha \end{cases}$$

igual a:
$$F_{X_{(1)}}(u) = \begin{cases} 0 & u < \alpha \\ 1 - e^{-n(u-\alpha)} & u \geq \alpha \end{cases}$$

y su densidad será:
$$f_{X_{(1)}}(u) = n\,e^{-n(u-\alpha)} \qquad u > \alpha$$
$$= 0 \qquad \text{en caso contrario}$$

Observad que $X_{(1)}$ sigue una exponencial "trasladada" su origen en α. Así $X_{(1)}-\alpha$ seguiría una exponencial de parámetro n.

Por tanto:

$$E_\alpha(X_{(1)}) = E_\alpha(X_{(1)}-\alpha+\alpha) = E_\alpha(X_{(1)}-\alpha)+\alpha = \frac{1}{n}+\alpha$$

El estimador UMVU se obtendrá corrigiendo su sesgo y aplicando Lehmann-Scheffé:

$$U(X_1 \ldots X_n) := X_{(1)} - \frac{1}{n}$$

por construcción es insesgado y al ser función de un estadístico suficiente completo es el UMVU.

* $EQM_\alpha(U) = B_\alpha(U)^2 + var_\alpha(U)$

pero $B_\alpha(U)=0$ (insesgado) por tanto basta hallar $var_\alpha(U)$

$$var_\alpha(U) = var_\alpha\left(X_{(1)}-\frac{1}{n}\right) = var_\alpha(X_{(1)}-\alpha) = \frac{1}{n^2}$$

(varianza exponencial: $Y \sim exp(\lambda)$ $var(Y)=\frac{1}{\lambda^2}$)

$$EQM_\alpha(U) = \frac{1}{n^2}$$

* La función de densidad de U será:

$$f_U(y) = f_{X_{(1)}}(u(y)) \cdot \left|\frac{du}{dy}\right|$$ siempre que $u(y) > \alpha$

donde $u(y) = y - \frac{1}{n}$ $\quad \frac{du}{dy} = 1$ por tanto

$$f_U(y) = \begin{cases} n\, e^{-n(y-\frac{1}{n}-\alpha)} & y > \alpha+\frac{1}{n} \\ 0 & y \le \alpha+\frac{1}{n} \end{cases}$$

2.- La función de densidad discreta:

$$f(x;p) = p^x (1-p) \mathbb{1}_{\{0,1\}}(x) \qquad 0 < p < 1$$

corresponde a una distribución de Bernoulli

* La distribución conjunta muestra/parámetro será proporcional a

$$h(x,p) \propto p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \mathbb{1}_{\{0,1\}}(x_i)$$

la distribución a posteriori será:

$$g(p|x_1-x_n) = C\, p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i} \qquad p \in (0,1)$$

La constante de normalización es:

$$C\int_0^1 p^{n\bar{x}_n} (1-p)^{n(1-\bar{x}_n)} dp = C\, B(n\bar{x}_n+1, n(1-\bar{x}_n)+1)$$

$$C = \frac{1}{B(n\bar{x}_n+1, n(1-\bar{x}_n)+1)} = \frac{\Gamma(n+2)}{\Gamma(n\bar{x}_n+1)\Gamma(n(1-\bar{x}_n)+1)}$$

El estimador de Bayes correspondiente a la pérdida cuadrática es la esperanza de p condicionada a $x_1 - x_n$ (posteriori)

$$E(p) = \int_0^1 p\, g(p|x_1-x_n)\, dp = \int_0^1 \frac{\Gamma(n+2)}{\Gamma(n\bar{x}_n+1)\Gamma(n(1-\bar{x}_n)+1)} p^{n\bar{x}_n+1} (1-p)^{n(1-\bar{x}_n)} dp$$

$$= \frac{\Gamma(n+2)}{\Gamma(n\bar{x}_n+1)\Gamma(n(1-\bar{x}_n)+1)} \frac{\Gamma(n\bar{x}_n+2)\Gamma(n(1-\bar{x}_n)+1)}{\Gamma(n+3)}$$

$$= \frac{n\bar{x}_n+1}{n+2} = \frac{n}{n+2}\bar{x}_n + \frac{1}{n+2}$$

Por tanto el estimador de Bayes es: $\boxed{\mathcal{U}(x_1-x_n) = \frac{n}{n+2}\bar{X}_n + \frac{1}{n+2}}$

* El MLE se obtendrá maximizando L_x :

$$L_x(p) = \prod_{i=1}^{n} \left\{ p^{x_i}(1-p)^{1-x_i} \mathbb{1}_{\{0,1\}}(x_i) \right\}$$

$$= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \underbrace{\prod_{i=1}^n \mathbb{1}_{\{0,1\}}(x_i)}_{\substack{\| \\ 1 \text{ con probabilidad } 1 \\ \forall p \in (0,1)}}$$

$$L_x(p) = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i}$$

$$\ln L_x(p) = \sum_{i=1}^n x_i \ln p + \left(n - \sum_{i=1}^n x_i\right) \ln(1-p)$$

$$\frac{\partial \ln L_x}{\partial p} = \frac{\sum_{i=1}^n x_i}{p} - \frac{n - \sum_{i=1}^n x_i}{1-p} = \frac{n}{p(1-p)} \left\{ \bar{x}_n - p \right\}$$

Observar que

$$\frac{\partial \ln L_x}{\partial p} = \begin{cases} > 0 & p < \bar{x}_n \quad \ln L_x \text{ y } L_x \text{ son crecientes} \\ = 0 & p = \bar{x}_n \\ < 0 & p > \bar{x}_n \quad \ln L_x \text{ y } L_x \text{ son decrecientes} \end{cases}$$

por tanto en $p = \bar{x}_n$ hay un máximo (absoluto) y el MLE es:

$$\boxed{p^* = \bar{X}_n}$$

Si hacemos la diferencia entre el estimador de Bayes anterior, $\mathcal{U}$, resulta:

$$\mathcal{U} - p^* = \frac{n}{n+2} \bar{X}_n + \frac{1}{n+2} - \bar{X}_n = -\frac{2}{n+2} \bar{X}_n + \frac{1}{n+2}$$

Observar que ambos se comportan de forma parecida, con elevada probabilidad, cuando $n \to \infty$, con independencia del valor que tome $\bar{X}_n$.

* Para hallar la distribución asintótica calcularemos la información de Fisher:

$$I(p) = E_p\left(\left\{\frac{\partial \ln f(x;p)}{\partial p}\right\}^2\right)$$

Como

$$f(x;p) = p^x (1-p)^{1-x} \qquad x \in \{0,1\}$$

resulta:

$$\ln f(x;p) = x \ln p + (1-x) \ln(1-p)$$

$$\frac{\partial \ln f(x;p)}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p}$$

$$\left\{\frac{\partial \ln f(x;p)}{\partial p}\right\}^2 = \frac{x^2}{p^2} + \frac{(1-x)^2}{(1-p)^2} - 2\frac{x(1-x)}{p(1-p)}$$

$$I(p) = E_p\left(\left\{\frac{\partial \ln f(x;p)}{\partial p}\right\}^2\right) = E_p\left(\frac{X^2}{p^2} + \frac{(1-X)^2}{(1-p)^2} - 2\frac{X(1-X)}{p(1-p)}\right)$$

Pero $E_p(X^2) = 1 \cdot P(X^2 = 1) = p$

$$E_p((1-X)^2 = 0 \cdot P(X=0) = 1-p$$

$$E_p(X(1-X)) = 0$$

y

$$I(p) = \frac{1}{p} + \frac{1}{1-p} = \frac{1}{p(1-p)}$$

Por tanto

$$\sqrt{n}(p^* - p) \xrightarrow{\mathcal{L}} Y \sim N\left(0, \text{var} = \frac{1}{I(p)} = p(1-p)\right)$$

$$\boxed{\sqrt{n}\frac{(p^*-p)}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} Z \sim N(0,1)}$$

p^* convenientemente normalizado tiende a una $N(0,1)$.

* Para una muestra de tamaño 2, los estadísticos posibles son

$$W(x_1,x_2) = \begin{cases} W_{00} & \Longleftrightarrow \quad x_1=x_2=0 \\ W_{10} & \Longleftrightarrow \quad x_1=1 \quad x_2=0 \\ W_{01} & \Longleftrightarrow \quad x_1=0 \quad x_2=1 \\ W_{11} & \Longleftrightarrow \quad x_1=1 \quad x_2=1 \end{cases}$$

Por otra parte $P(X_1=x_1, X_2=x_2) = p^{x_1}(1-p)^{1-x_1}\, p^{x_2}(1-p)^{1-x_2} =$

$$= p^{x_1+x_2}(1-p)^{2-(x_1+x_2)}$$

Por tanto la condición de ser insesgado es:

$$W_{00}(1-p)^2 + W_{10}\,p(1-p) + W_{01}(1-p)p + W_{11}p^2 = p^2 \qquad \forall p \in [0,1]$$

que puede escribirse como:

$$p^2\{W_{00}-W_{10}-W_{01}+W_{11}-1\} + p\{-2W_{00}+W_{10}+W_{01}\} + W_{00} = 0$$

para todo $p\in(0,1)$. Por tanto: $W_{00}=0$; $-2W_{00}+W_{10}+W_{01}=0$

lo que implica que $W_{10}+W_{01}=0$ y además:

$W_{00}-W_{10}-W_{01}+W_{11}-1=0$ pero como $W_{00}=0$ y $W_{10}+W_{01}=0$

resultará:

$W_{11}-1=0$ junto con la anterior

y por tanto $W_{11}=1$

Es decir, cualquier insesgado de p^2 satisface:

$$W(0,0)=0$$
$$W(1,0)+W(0,1)=0$$
$$W(1,1)=1$$

pero como $W(1,0)\geq 0$ y $W(0,1)\geq 0 \implies W(1,0)=W(0,1)=0$

Sólo hay pues un único estimador insesgado (salvo conjuntos de probabilidad 0)

$$\left[\; W(X_1,X_2) = \begin{cases} 1 & \Longleftrightarrow \quad X_1=1 \quad X_2=1 \\ 0 & \text{en caso contrario} \end{cases} \;\right]$$

Notad que $W(X_1X_2) = W(U)$

Al ser único, automáticamente es UMVU (observar que es función de un suficiente y completo, U)