ejercicios3IE.pdf

(Ej 5 modificado) Relación 3 resuelta

3º Inferencia Estadística

Grado en Matemáticas

Facultad de Ciencias
Universidad de Granada

Relación de ejercicios 3:[1] Suficiencia y completitud.

Ejercicio 1.

Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable $X \rightsquigarrow \{B(k,p); p \in (0,1)\}$ y sea $T(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$. Probar,

a) usando la definición.
b) aplicando el teorema de factorización.

que T es suficiente para p.

SOLUCIÓN

$X \rightsquigarrow B(k,p) \Rightarrow P[X = x] = \binom{k}{x} p^x (1-p)^{k-x}$

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de X.

a) Consideramos el estadístico $T = \sum_{i=1}^{n} X_i \rightsquigarrow B(nk, p)$.

Para ver que es suficiente, demostramos que la probabilidad de cualquier valor de la muestra condicionada a cualquier valor del estadístico es independiente del parámetro p.

$$P_p[X_1 = x_1, \ldots, X_n = x_n \,/\, T(X_1, \ldots, X_n) = X_1 + \ldots + X_n = t] = \frac{P_p[X_1 = x_1, \ldots, X_n = x_n, T = t]}{P_p[T = t]}$$

lo que nos da dos posibles resultados:

i) 0 si $x_1 + \cdots + x_n \neq t$ pues el numerador se anula por ser una intersección de sucesos que dan lugar a un suceso imposible y, por tanto, de probabilidad 0. En este caso, la probabilidad (siempre 0) es independiente de p.

ii) Desarrollamos el caso $x_1 + \cdots + x_n = t$:

$$\frac{P_p[X_1 = x_1, \ldots, X_n = x_n, T = t]}{P_p[T = t]} \underbrace{=}_{\text{suceso seguro}} \frac{P_p[X_1 = x_1, \ldots, X_n = x_n]}{P_p[T = t]}$$

$$= \frac{P_p[X_1 = x_1] \cdots P_p[X_n = x_n]}{P_p[T = t]} = \frac{\binom{k}{x_1} p^{x_1} (1-p)^{k-x_1} \cdots \binom{k}{x_n} p^{x_n} (1-p)^{k-x_n}}{\binom{nk}{t} p^t (1-p)^{nk-t}}$$

$$\underbrace{=}_{t=\sum x_i} \frac{\prod_{i=1}^{n} \binom{k}{x_i} p^{\sum x_i} (1-p)^{nk-\sum x_i}}{\binom{nk}{\sum x_i} p^{\sum x_i} (1-p)^{nk-\sum x_i}} = \frac{\prod_{i=1}^{n} \binom{k}{x_i}}{\binom{nk}{\sum x_i}}$$

En este caso, tenemos que la probabilidad es independiente de p.

[1] Si encuentras algún error en el documento, escribe un correo electrónico a *melferizq@gmail.com*.

Por tanto, T es un estadístico suficiente.

b) En la función masa de probabilidad de la muestra tenemos que encontrar una descomposición en una función h que solo dependa de la muestra y otra función $g_p(T)$ que dependa de la muestra a través del estadístico. En este caso, como hemos visto en el apartado anterior, y como vimos en la relación 1:

$$f_p^n(x_1, \ldots, x_n) = \prod_{i=1}^{n} \binom{k}{x_i} p^{\sum_{i=1}^n x_i}(1-p)^{nk-\sum_{i=1}^n x_i}$$

Por tanto, tomando $h(x_1, \ldots, x_n) = \prod_{i=1}^{n} \binom{k}{x_i}$; $g_pT(x_1, \ldots, x_n) = p^{\sum_{i=1}^n x_i}(1-p)^{nk-\sum_{i=1}^n x_i}$, demostramos que el estadístico $T = \sum_{i=1}^{n}$ es suficiente.

Ejercicio 2.

Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable $X \rightsquigarrow \{P(\lambda); \lambda \in \mathbb{R}^+\}$ y sea $T(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$. Probar,

a) usando la definición.

b) aplicando el teorema de factorización.

que T es suficiente para λ.

SOLUCIÓN

$X \rightsquigarrow P(\lambda) \Rightarrow P[X = x] = e^{-\lambda}\dfrac{\lambda^x}{x!}$.

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de X.

a) Consideramos el estadístico $T = \sum_{i=1}^{n} X_i \rightsquigarrow P(n\lambda)$.

Para ver que es suficiente, demostramos que la probabilidad de cualquier valor de la muestra condicionada a cualquier valor del estadístico es independiente del parámetro λ.

$$P_\lambda[X_1 = x_1, \ldots, X_n = x_n \,/\, T(X_1, \ldots, X_n) = X_1 + \ldots + X_n = t] = \frac{P_\lambda[X_1 = x_1, \ldots, X_n = x_n, T = t]}{P_\lambda[T = t]}$$

lo que nos da dos posibles resultados:

i) 0 si $x_1 + \cdots + x_n \neq t$ pues el numerador se anula por ser una intersección de sucesos que dan lugar a un suceso imposible y, por tanto, de probabilidad 0. En este caso, la probabilidad (siempre 0) es independiente de λ.

ii) Desarrollamos el caso $x_1 + \cdots + x_n = t$:

$$\frac{P_\lambda[X_1 = x_1, \ldots, X_n = x_n, T = t]}{P_\lambda[T = t]} \underbrace{=}_{\text{suceso seguro}} \frac{P_\lambda[X_1 = x_1, \ldots X_n = x_n]}{P_\lambda[T = t]}$$

$$= \frac{P_\lambda[X_1 = x_1] \cdots P_\lambda[X_n = x_n]}{P_\lambda[T = t]} = \frac{e^{-\lambda}\frac{\lambda^{x_1}}{x_1!} \cdots e^{-\lambda}\frac{\lambda^{x_n}}{x_n!}}{e^{-n\lambda}\frac{(n\lambda)^t}{t!}} = \frac{e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^{x_i} x_i!}}{e^{-n\lambda}\frac{n\lambda^{\sum_{i=1}^n x_i}}{(\sum_{i=1}^n x_i)!}}$$

$$= \frac{\frac{1}{\prod_{i=1}^n x_i!}}{\frac{n}{(\sum_{i=1}^n x_i)!}} = \frac{\left(\sum_{i=1}^n x_i\right)!}{n \prod_{i=1}^n x_i!}$$

En este caso, tenemos que la probabilidad es independiente de λ.

Por tanto, el estadístico T es suficiente.

b) Como hemos visto en el apartado anterior, y como vimos en la relación 1,

$$f_\lambda^n(x_1, \ldots, x_n) = e^{-n\lambda}\frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^{x_i} x_i!}$$

Así, elegimos $h(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n x_i!}$; $g_\lambda T(x_1, \ldots, x_n) = e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i}$, demostrando así que el estadístico $T = \sum_{i=1}^n X_i$ es un estadístico suficiente.

Ejercicio 3.

Sea (X_1, X_2, X_3) una muestra aleatoria simple de una variable $X \rightsquigarrow \{B(1,p); p \in (0,1)\}$. Probar que el estadístico $X_1 + 2X_2 + 3X_3$ no es suficiente.

SOLUCIÓN

Tenemos que $X \rightsquigarrow B(1,p)$, luego $P[X = x] = p^x(1-p)^{1-x}$, con $x = 0, 1$. Sea X_1, X_2, X_3 una muestra aleatoria simple de X.

Vamos a considerar el valor del estadístico $T = t = 3$, que se podrá conseguir si las variables aleatorias X_i toman los siguientes valores:

- $X_1 = 1, X_2 = 1, X_3 = 0$
- $X_1 = 0, X_2 = 0, X_3 = 1$

Sea $X_1 = 0, X_2 = 0, X_3 = 1$, vamos a demostrar que la probabilidad de que las variables de la muestra tomen esos valores condicionado a que el estadístico T valga 3 no es independiente del parámetro p.

$$P_p[X_1=0,X_2=0,X_3=1 \,/\, T=3] = \frac{P_p[X_1=0,X_2=0,X_3=1,T=3]}{P_p[T=3]}$$

$$= \frac{P_p[X_1=0,X_2=0,X_3=1]}{P_p[T=3]} = \frac{P_p[X_1=0]P_p[X_2=0]P_p[X_3=1]}{P_p[X_1=1,X_2=1,X_3=0]+P_p[X_1=0,X_2=0,X_3=1]}$$

$$= \frac{(1-p)\cdot(1-p)\cdot p}{p\cdot p\cdot(1-p)+(1-p)\cdot(1-p)\cdot p} = \frac{(1-p)^{\cancel{2}}\cdot \cancel{p}}{p^{\cancel{2}}\cdot \cancel{(1-p)}+p\cdot(1-p)^{\cancel{2}}} = \frac{1-p}{p}$$

Esta probabilidad depende de p, luego T no es un estadístico suficiente.

Observación. Otra forma de hacer este ejercicio es hallar la distribución del estadístico y comprobar que la probabilidad de cualquier valor de la muestra condicionada a cualquier valor del estadístico depende de p. En este caso, esto es posible hacerlo pues la variable aleatoria de la que partimos es una Bernoulli, cuya suma nos da una binomial. Sin embargo, en general no será posible seguir este procedimiento.

Ejercicio 4.

Aplicando el teorema de factorización, y basándose en una muestra de tamaño arbitrario, encontrar un estadístico suficiente para cada una de las siguientes familias de distribuciones (en las familias biparamétricas, suponer los casos de solo un parámetro desconocido y de los dos desconocidos).

a) $X \rightsquigarrow \{U(-\theta/2,\theta/2);\theta>0\}$
b) $X \rightsquigarrow \{\Gamma(p,a);p>0,a>0\}$
c) $X \rightsquigarrow \{\beta(p,q);p>0,q>0\}$
d) $X \rightsquigarrow \{P_{N_1,N_2};N_1,N_2\in\mathbb{N},N_1\le N_2\};P_{N_1,N_2}\{X=X\}=\frac{1}{N_2-N_1+1},x=N_1,\ldots,N_2.$

SOLUCIÓN

Solo en el segundo apartado vamos a considerar los casos de un solo parámetro desconocido y de los dos desconocidos. En el apartado **c** vamos a suponer que desconocemos los dos parámetros (en el caso en el que solo sea desconocido uno de ellos, el procedimiento será el mismo que en el apartado **b**).

a) $X \rightsquigarrow U(-\theta/2,\theta/2),\theta>0 \Rightarrow f_\theta(x)=\dfrac{1}{\theta},\theta>0,\dfrac{-\theta}{2}<x<\dfrac{\theta}{2}$ (luego no es exponencial uniparamétrica).

$$f_\theta^n(x_1,\ldots,x_n)=\frac{1}{\theta^n},\ \frac{-\theta}{2}<x_i<\frac{\theta}{2}$$

Por tanto, estamos exigiendo $X_{(1)}>\dfrac{-\theta}{2},X_{(n)}<\dfrac{\theta}{2}$. Así:

$$f_\theta^n(x_1,\ldots,x_n)=\frac{1}{\theta^n}I_{[X_{(1)}>-\theta/2]}I_{[X_{(n)}<\theta/2]}$$

Por el teorema de factorización, tomando $h(x_1,\ldots,x_n)=1$; $g_\theta(T(X_1,\ldots,x_n)) = \dfrac{1}{\theta^n}I_{[X_{(1)}>-\theta/2]}I_{[X_{(n)}<\theta/2]}$ llegamos a que $T(X_{(1)},X_{(n)})$ es un estadístico suficiente.

b) $X \rightsquigarrow \Gamma(p,a); p>0, a>0$.

$$f_{a,p}(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} \Rightarrow f^n_{a,p}(x_1,\dots,x_n) = \frac{a^{np}}{(\Gamma(p))^n} \left(\prod_{i=1}^n x_i\right)^{p-1} \exp\left(-a\sum_{i=1}^n x_i\right)$$

i) Consideremos a conocido y p desconocido. Entonces, por el Teorema de Factorización, $T = \prod_{i=1}^n X_i$ será un estadístico suficiente.

ii) Consideremos p conocido y a desconocido. Entonces, $T = \sum_{i=1}^n X_i$ será un estadístico suficiente.

iii) Si ahora consideramos desconocidos tanto a como p, entonces $T = (\prod_{i=1}^n X_i, \sum_{i=1}^n X_i)$ será un estadístico suficiente.

c) $X \rightsquigarrow \beta(p,q); p>0, q>0$.

$$f_{p,q}(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1}$$
$$\Rightarrow f^n_{p,q}(x_1,\dots,x_n) = \left(\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\right)^n \left(\prod_{i=1}^n x_i\right)^{p-1} \left(\prod_{i=1}^n (1-x_i)\right)^{q-1}$$

Por el Teorema de Factorización, $T = (\prod_{i=1}^n X_i, \prod_{i=1}^n (1-X_i))$ es un estadístico suficiente.

d) $X \rightsquigarrow P_{N_1,N_2}; N_1, N_2 \in \mathbb{N}, N_1 \le N_2$.

$$P_{N_1,N_2}[X=x] = \frac{1}{N_2-N_1+1}, x = N_1,\dots,N_2$$
$$\Rightarrow P^n_{N_1,N_2}(x_1,\dots,x_n) = \frac{1}{(N_2-N_1+1)^n} I_{[X_{(1)}>N_1]} I_{[X_{(2)}<N_2]}$$

Por el Teorema de Factorización, el estadístico $T = (X_{(1)}, X_{(n)})$ es un estadístico suficiente.

Ejercicio 5.

Sea $X \rightsquigarrow \{P_N; N \in \mathbb{N}\}$, siendo P_N la distribución uniforme en los puntos $\{1,\dots,N\}$, y sea $(X_1,\dots,X_n)$ una muestra aleatoria simple de X. Probar que $\max(X_1,\dots,X_n)$ es un estadístico suficiente y completo.

SOLUCIÓN

$X \rightsquigarrow \{P_N : N \in \mathbb{N}\}, P_N(X=x) = \dfrac{1}{N}, x = 1,\dots,N \leftarrow$ F.m.p.

El conjunto de valores de X para N fijo es $\chi_N = \{1,\dots,N\}$. Así, el conjunto de valores de X es $\bigcup_{n\in\mathbb{N}} \chi_N = \bigcup_{n\in\mathbb{N}} \{1,\dots,N\} = \mathbb{N} \Rightarrow \chi^n = \mathbb{N}^n$.

Calculemos la función masa de probabilidad de una muestra aleatoria simple de tamaño n de X, con el fin de aplicar el teorema de factorización.

$$P_N[X_1 = x_1, \dots, X_n = x_n] = P_N[X_1 = x_1] \cdot \dots \cdot P_N[X_n = x_n] = P_N[X = x_1] \cdots P_N[X = x_n]$$
$$= \frac{1}{N} I_{[1,N]}(x_1) \cdots \frac{1}{N} I_{[1,N]}(x_n) = \frac{1}{N^n} I_{[1,N]}(x_1) \cdots I_{[1,N]}(x_n), (x_1, \dots, x_n) \in \mathbb{N}^n$$

Por tanto, la función masa de probabilidad de una muestra aleatoria simple de X de tamaño n es:

$$f_N^n(x_1, \dots, x_n) = P[X_1 = x_1, \dots, X_n = x_n] = \begin{cases} \frac{1}{N^n} & \text{máx}\, x_i \leq N \\ & \\ 0 & \text{máx}\, x_i > N \end{cases}, (x_1, \dots, x_n) \in \mathbb{N}^n$$

Observemos que en este caso particular no es necesario especificar $\text{mín}\, x_i \geq 1$ puesto que, por definición, X toma valores naturales.

Otra forma de expresar la f.m.p. muestral es:

$$f_N^n(x_1, \dots, x_n) = \frac{1}{N^n} I_{[X_{(n)} \leq N]}$$

Así, tomando $h(x_1, \dots, x_n) = 1$ se llega a que $T(X_1, \dots, X_n) = X_{(n)}$ es un estadístico suficiente.

Comprobemos que es completo.

Lo que tenemos que probar es que si g es una transformación tal que $\mathrm{E}_N[g(T(x_1, \dots, x_n))] = 0, \forall N \in \mathbb{N}$ entonces $P_N[g(T) = 0] = 1$. Para ello, vamos a calcular la función masa de probabilidad de $X_{(n)}$, para lo cual necesitamos previamente su función de distribución. Para $x \in \{1, \dots, N\}$,

$$P_N[X_{(n)} \leq x] = P_N[X_1 \leq x] \cdots P_N[X_n \leq x] = P_N[X \leq x] \cdot P_N[X \leq x] = \left(\frac{x}{N}\right)^n$$

$$P_N[X_{(n)} = x] = P_N[X_{(n)} \leq x] - P_N[X_{(n)} \leq x - 1] = \left(\frac{x}{N}\right)^n - \left(\frac{x-1}{N}\right)^n = \frac{x^n - (x-1)^n}{N^n}$$

Por tanto,

$$\mathrm{E}_N[g(T)] = \sum_{t=1}^{N} g(t) \frac{t^n - (t-1)^n}{N^n} = \frac{1}{N^n} \sum_{i=1}^{N} g(t)(t^n - (t-1)^n) = 0 \Leftrightarrow \sum_{t=1}^{N} g(t)(t^n - (t-1)^n) = 0$$

Vamos ahora a darle valores a N para ver que $g(t) = 0, \forall t \in \mathbb{N}$.

- $N = 1 \Rightarrow \sum_{t=1}^{1} g(t)(t^n - (t-1)^n) = g(1) \cdot 1 = 0 \Leftrightarrow g(1) = 0.$
- $N = 2 \Rightarrow \sum_{t=1}^{2} g(t)(t^n - (t-1)^n) = g(1) + g(2)(2^n - 1) \underbrace{=}_{g(1)=0} g(2)(2^n - 1) = 0 \underbrace{\Leftrightarrow}_{n>0} g(2) = 0.$

Así, por inducción, se prueba que $g(k) = 0$ pues g se anula en los $k - 1$ valores anteriores, y por tanto, $g(k + 1) = 0$. Por tanto, he demostrado que $g(t) = 0, \forall N \in \mathbb{N}$. Así que

$$\{t \in \{1 \dots, N\}\} \subseteq \{t / g(t) = 0\}$$

Y tomando ahora probabilidades, se tiene

$$1 \geq P[g(t) = 0] \geq P[t \in \{1, \dots, N\}] = 1$$

Por tanto,

$$P[g(t) = 0] = 1$$

así que T es también un estadístico completo.

Ejercicio 6.

Basándose en una muestra aleatoria simple de tamaño arbitrario, obtener un estadístico suficiente y completo para la familia de distribuciones definidas por todas las densidades de la forma

$$f_\theta(x) = e^{\theta - x}, x > \theta$$

SOLUCIÓN

Sea X una variable aleatoria con función de densidad $f_\theta(x) = e^{\theta-x}, x > \theta$.
Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de X. Se tiene que su función de densidad es:

$$f_\theta^n(x_1, \dots, x_n) = e^{\theta - x_1} \cdots e^{\theta - x_n} = e^{n\theta - \sum_{i=1}^n x_i}, x_i > \theta, \forall i$$

Por tanto, estoy exigiendo que $x_i > \theta, \forall i \in \{1, \dots, n\}$, es decir, que $X_{(1)} > \theta$. Por tanto, la función de densidad de la muestra será finalmente

$$f_\theta^n(x_1, \dots, x_n) = e^{n\theta - \sum_{i=1}^n x_i} I_{[X_{(1)} > \theta]}$$

Tomo $h(x_1, \dots, x_n) = e^{-\sum_{i=1}^n x_i}$, $g_\theta(T(x_1, \dots, x_n)) = e^{n\theta} I_{[X_{(1)} > \theta]}$, demostrando entonces que $T = X_{(1)}$ es un estadístico suficiente.
Vamos a comprobar que es completo.
Por un lado, si $f_\theta(x) = e^{\theta - x} \Rightarrow F_\theta(x) = \int_\theta^x e^{\theta - t} dt = e^\theta \int_\theta^x e^{-t} dt = -e^\theta \int_\theta^x -e^{-t} dt = -e^\theta [e^{-t}]_0^x = 1 - e^{\theta - x}$.
Por otra parte, sabemos que $F_{X(1)}(y) = 1 - (1 - F(y))^n \Rightarrow f_{X(1)}(y) = n(1 - F(y))^{n-1} f(y)$. Así que, en este caso,

$$f_{X(1)}(y) = ne^{n(\theta - y)}$$

Debo comprobar ahora que si $\mathrm{E}[g(T)] = 0 \Rightarrow P[g(t) = 0] = 1$.

$$\mathrm{E}[g(T)] = \int_\theta^{+\infty} g(t) f_\theta(t) dt = \int_\theta^{+\infty} g(t) ne^{n(\theta - t)} dt = ne^{n\theta} \int_\theta^{+\infty} g(t) e^{-nt} dt = 0$$

y esto se cumple si y solo si $\int_\theta^{+\infty} g(t) e^{-nt} dt = 0$ (pues $ne^{n\theta}$ no puede ser 0 para ningún θ).
Por el Teorema Fundamental del Cálculo sé que existe primitiva, que es $G(+\infty) - G(\theta)$, y derivando con respecto a θ se tiene que $g(\theta)e^{-n\theta} = 0, \forall \theta \Rightarrow g(\theta) = 0, \forall \theta$.
Por tanto, he demostrado que $\{t/g(t) = 0\} \supseteq \{t > \theta\}$. Tomando probabilidades, obtengo

$$1 \geq P\{g(t) = 0\} \geq P[T \geq \theta] = 1$$

Por tanto,

$$P\{g(T) = 0\} = 1$$

luego el estadístico T también es completo.

Observación. Comprobar que un estadístico es suficiente y completo es inmediato en el caso de que tengamos familias exponenciales k-paramétricas. Así que, se nos podría haber ocurrido comprobar que esta familia, al depender de un único parámetro, puede ser exponencial uniparamétrica. Sin embargo, vemos que $\chi = \{x/f_{p,a}(x) > 0\} = (\theta, +\infty)$, es decir, depende del parámetro. Por tanto, se tiene que no es exponencial uniparamétrica (pero eso no significa que no puedan cumplirse las otras dos condiciones necesarias para serlo).

Ejercicio 7.

Comprobar que las siguientes familias de distribuciones son exponenciales uniparamétricas y, considerando una muestra aleatoria simple de una variable con distribución en dicha familia, obtener, si existe, un estadístico suficiente y completo.

a) $\{B(k_0, p); 0 < p < 1\}$
b) $\{P(\lambda); \lambda > 0\}$
c) $\{BN(k_0, p); 0 < p < 1\}$
d) $\{\exp(\lambda), \lambda > 0\}$

SOLUCIÓN

a) $X \rightsquigarrow B(k_0, p); 0 < p < 1 \Rightarrow P[X = x] = \binom{k_0}{x} p^x (1-p)^{k_0 - x}, x = 0, 1, \dots, k_0.$

i) $\Theta = (0, 1)$ intervalo de $\mathbb{R}$.

ii) $\chi = \{x / P[X = x] > 0\} = \{0, 1, 2, \dots, k_0\}$ independiente de p.

iii) Intentamos reescribir la función masa de probabilidad.

$$P[X = x] = \exp\left[\ln \binom{k_0}{x} + x \ln p + (k_0 - x)\ln(1-p)\right]$$

Así que nuestra selección será $S(x) = \ln \binom{k_0}{x}, D(p) = k_0 \ln(1-p), Q(p) = \ln(p) - \ln(1-p), T(x) = x$, donde hemos sacado factor común x para agrupar todos los sumandos que dependen de p.

Por tanto, la familia es, en efecto, exponencial uniparamétrica, y acabamos de demostrar que $T = \sum_{i=1}^n X_i$ es un estadístico suficiente. Además, como $\text{Im}(Q) = \mathbb{R}$ que contiene obviamente a abiertos de $\mathbb{R}$, entonces se tiene que T es además completo.

b) Es evidente comprobar las dos primeras condiciones para que sea familia paramétrica, y se tiene que

$$P[X = x] = e^{-\lambda}\frac{\lambda^x}{x!} = \exp[-\lambda + x \ln \lambda - \ln x!]$$

Por tanto, $S(x) = -\ln x!, D(\lambda) = -\lambda, T(x) = x, Q(\lambda) = \ln \lambda$. Así que, por el mismo razonamiento que en el apartado anterior, $T = \sum_{i=1}^n X_i$ es un estadístico suficiente y completo.

c) i) y ii) son evidentes, y

$$P[X = x] = \binom{x + k_0 - 1}{x}(1-p)^x p^{k_0} = \exp\left[\ln \binom{x + k_0 - 1}{x} + x\ln(1-p) + k_0 \ln p\right]$$

así que $S(x) = \ln \binom{x + k_0 - 1}{x}, D(p) = k_0 \ln p, T(x) = x, Q(p) = \ln(1-p)$ por lo que $T = \sum_{i=1}^n X_i$ es un estadístico suficiente y completo (por el mismo razonamiento que en a).

d) i) y ii) son evidentes, y como $f(x) = \lambda e^{-\lambda x}$, llegamos entonces a $S(x) = 0, D(\lambda) = \ln(\lambda), T(x) = x, Q(\lambda) = -\lambda$, así que, $T = \sum_{i=1}^n X_i$ es un estadístico suficiente y completo para la distribución.

Observación. Que se tenga, en todos los casos, un estadístico suficiente y completo no viene solo de que la familia sea exponencial uniparamétrica. No olvidar justificar la suficiencia y la completitud de los estadísticos con el último teorema de los apuntes de Teoría.

Ejercicio 8.

Estudiar si las siguientes familias de distribuciones son exponenciales biparamétricas. En caso afirmativo, considerando una muestra aleatoria simple de una variable con distribución en dicha familia, obtener, si existe, un estadístico suficiente y completo.

a) $X \rightsquigarrow \Gamma(p,a); p>0, a>0$.
b) $X \rightsquigarrow \beta(p,q); p>0, q>0$.

SOLUCIÓN

a) $X \rightsquigarrow \Gamma(p,a); p>0, a>0 \Rightarrow f_{p,a}(x) = \dfrac{a^p}{\Gamma(p)}x^{p-1}e^{-ax}, x\in\mathbb{R}$.

   i) $\Theta = (\mathbb{R}^+, \mathbb{R}^+)$ intervalo de $\mathbb{R}^2$.

   ii) $\chi = \{x / f_{p,a}(x) > 0\} = \mathbb{R}$ independiente de p, a.

   iii) Intentamos reescribir la función de densidad de la distribución como la exponencial de una serie de funciones (dependientes de la muestra o de los parámetros desconocidos) tal y como aparecen en la definición de familia exponencial k-paramétrica.

$$\begin{aligned} f_{p,a}(x) &= \exp\left[\ln\left(\frac{a^p}{\Gamma(p)}x^{p-1}e^{-ax}\right)\right] = \exp\left[\ln\left(\frac{a^p}{\Gamma(p)}\right) + \ln(x^{p-1}) + \ln(e^{-ax})\right] \\ &= \exp\left[\ln(a^p) - \ln(\Gamma(p)) + (p-1)\ln(x) - ax\right] \\ &= \exp\left[\ln(a^p) - \ln(\Gamma(p)) + p\ln(x) - \ln(x) - ax\right] \end{aligned}$$

   Por tanto, podemos tomar:

$$S(x) = -\ln(x)$$

$$D(a,p) = \ln(a^p) - \ln(\Gamma(p))$$

$$T_1(x) = \ln(x), Q_1(a,p) = p; T_2(x) = x, Q_2(a,p) = -a$$

   Por tanto, tenemos demostrado que el estadístico $T = (\sum_{i=1}^n \ln(X_i), \sum_{i=1}^n X_i)$ es suficiente (por el Teorema visto en clase). Por otro lado, $\text{Im}(Q_1) = \mathbb{R}^+, \text{Im}(Q_2) = \mathbb{R}^-$, y es evidente que $(\mathbb{R}^+, \mathbb{R}^-)$ contiene a abiertos de $\mathbb{R}^2$. Por tanto, T también es un estadístico completo.

b) $X \rightsquigarrow \beta(p,q); p>0, q>0 \Rightarrow f_{p,q}(x) = \dfrac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}x^{p-1}(1-x)^{q-1}, x\in(0,1)$.

   i) $\Theta = (\mathbb{R}^+, \mathbb{R}^+)$ intervalo de $\mathbb{R}^2$.

   ii) $\chi = \{x / f_{p,q}(x) > 0\} = (0,1)$ independiente de p, q.

iii) Intentamos reescribir la función de densidad de la distribución.

$$
\begin{aligned}
f_{p,q}(x) &= \exp\left[\ln\left(\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)}\right) + \ln(x^{p-1}) + \ln((1-x)^{q-1})\right] \\
&= \exp\left[\ln(\Gamma(p+q)) - \ln(\Gamma(p)\Gamma(q)) + \overbrace{(p-1)\ln(x)}^{=p\ln(x)-\ln(x)} + \underbrace{(q-1)\ln(1-x)}_{=q\ln(1-x)-\ln(1-x)}\right]
\end{aligned}
$$

Por tanto, podemos tomar:

$$S(x) = -\ln(x) - \ln(1-x)$$

$$D(p,q) = \ln(\Gamma(p+q)) - \ln(\Gamma(p)\Gamma(q))$$

$$T_1(x) = \ln(x), Q_1(p,q) = p; T_2(x) = \ln(1-x), Q_2(p,q) = q$$

Por el mismo razonamiento que antes, el estadístico $T = (\sum_{i=1}^{n} \ln(X_i), \sum_{i=1}^{n} \ln(1-X_i))$ es suficiente, y por contener $(\mathbb{R}^+, \mathbb{R}^+)$ abiertos en $\mathbb{R}^2$, entonces además, T es completo.