tema3teorIE.pdf

(Provisional) Apuntes tema 3

3º Inferencia Estadística

Grado en Matemáticas

Facultad de Ciencias
Universidad de Granada

Tema 3: Suficiencia y completitud. [1]

Índice

En este tema se estudian algunas propiedades que deben cumplir los estadísticos para poder simplificar los datos de una muestra aleatoria simple sin que se pierda información relevante, así como familias de distribuciones para las que es seguro que existen estadísticos con tales propiedades.

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

1. Introducción.

Aunque el objetivo de seleccionar una muestra aleatoria simple de una población es reducir el número de datos poblacionales a estudiar, es posible que incluso una muestra que sea representativa de la población no pueda ser estudiada de forma exhaustiva. Por ello, será también necesario encontrar métodos que nos permitan sacar conclusiones sobre la muestra sin que ello implique un esfuerzo inasequible.

Nos centraremos en estudiar lo que, en el tema 1, definimos como *estadístico*, que es una función construida a partir de los datos muestrales. Así, el proceso de investigación se simplifica y ya no hay que observar cada elemento de la muestra sino solo los estadísticos asociados a ella. Por tanto, aunque el uso de cualquier estadístico (a los que denotamos $T(X_1, \dots, X_n)$, o simplemente, $T(X)$) implica una reducción en los datos muestrales, también supone una cierta pérdida de la información sobre la muestra.

Por ejemplo, supongamos que se lanza una moneda al aire 100 veces, y tomamos dos muestras aleatorias simples de los resultados de tales lanzamientos. Sean $(X_1, \dots, X_5)$ e $(Y_1, \dots, Y_5)$ dichas muestras. Si llamamos 1 al suceso *salir cara* y 0 al suceso *salir cruz* y se tiene que $\bar{X} = \bar{Y} = 0{,}4$, entonces sabemos que en cada muestra deberá haber dos 1 y tres 0. Sin embargo, no tenemos forma de saber si las muestras son $(1, 1, 0, 0, 0)$ o $(0, 1, 0, 0, 1)$, luego a pesar de que es más sencillo estudiar y obtener conclusiones a partir de las medias muestrales, hemos perdido la información sobre la forma en que se distribuyen los datos de tales muestras.

Se plantea el problema de encontrar estadísticos T tales que la información que se pierde al utilizarlos sea irrelevante para los fines con los que los estamos empleando. Para ello, definimos las propiedades de suficiencia y completitud.

Demostrar que un estadístico es suficiente y completo en ocasiones será complejo. Por ello, se estudiarán familias de distribuciones para las cuales la existencia de un estadístico que sea a la vez suficiente y completo está asegurada, incluso habiendo métodos para su obtención directa.

2. Estadísticos suficientes.

Dada una variable aleatoria X con distribución en una familia de distribuciones $\{F_\theta, \theta \in \Theta\}$ dependientes de un parámetro θ desconocido. Recordemos que nuestro objetivo es inferir el valor de θ a partir de una muestra aleatoria simple de X.

Estadístico.

Definición 1. *Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de una variable aleatoria X. Un estadístico muestral asociado a X es una función $T : (\mathbb{R}^n, \mathcal{B}^n) \longrightarrow (\mathbb{R}^k, \mathcal{B}^k)$ medible, es decir, cumple $T^{-1}(B) \in \mathcal{B}^n, \forall B \in \mathcal{B}^k$, e independiente de cualquier parámetro desconocido. Lo denotamos por $T(X_1, \dots, X_n)$, o cuando se sobreentienda que es un estadístico, simplemente T o $T(X)$.*

Así, cualquier estadístico $T(X_1, \dots, X_n)$ de una variable aleatoria X resume la información contenida en una muestra aleatoria simple asociada, y las inferencias que hagamos sobre θ se harán a partir de los valores de tales estadísticos. Sin embargo, con el fin de perder la menor cantidad de información de la muestra que sea posible, se utilizarán estadísticos que cumplan ciertas propiedades. La primera de ellas será la de *suficiencia*.

Fisher (1922) dijo que *un estadístico es suficiente cuando contiene toda la información contenida en la muestra sobre el parámetro que se está considerando.* Sin embargo, esta definición es demasiado general y poco rigurosa, aunque sí que puede ser útil si la reescribimos con cuidado. En efecto, si un estadístico es suficiente al contener toda la información que puede dar la muestra sobre el parámetro desconocido eso significa que cualquier información adicional que la muestra pueda aportar no proporciona información relevante sobre θ.

Es evidente que la información más importante que nos puede dar la muestra es el propio valor del estadístico. Sin embargo, su valor cambia para cada muestra considerada, y debemos entender que un estadístico es suficiente cuando, al cambiar la muestra seleccionada, su validez para inferir el parámetro θ no cambia. Es decir, sea cual sea el valor de T que se obtenga, este sigue siendo igual de válido a la hora de hacer inferencias sobre θ.

Pero, ¿cómo puede ser un estadístico igual de válido para cualquier valor que se obtenga? Esto solo sería posible si, para cualquier muestra, el valor del estadístico fuese aproximadamente el mismo. Así, buscamos estadísticos cuyos valores no difieran demasiado dependiendo de la muestra utilizada, es decir, buscamos estadísticos cuyos posibles valores están *igualmente distribuidos*. Por ejemplo, si demostramos que una variable aleatoria toma siempre valores entre 0 y 5, no nos equivocamos al suponer que puede estar distribuida según una binomial $B(5, 0'4)$. Por tanto, si un estadístico toma siempre los mismos valores (aproximadamente), para cualquier muestra considerada, podemos suponer que su distribución es siempre la misma.

Si tal suposición se convierte en una imposición, es evidente que estaremos ante estadísticos que siempre tienen la misma distribución, independientemente de la muestra que se considere y del valor concreto del estadístico que esta pueda proporcionar. Esta es la definición de estadístico suficiente.

Estadístico suficiente.

Definición 2. *Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de una variable aleatoria $X \rightsquigarrow F \in \{F_\theta, \theta \in \Theta\}$. Un estadístico $T(X_1, \dots, X_n)$ es suficiente para la familia de distribuciones considerada (o suficiente para θ) si la distribución de la muestra condicionada a cualquier valor del estadístico, $T(X_1, \dots, X_n) = t$, es independiente de θ.*

Observación 1. *Para cualquier variable aleatoria y para cualquier muestra aleatoria simple, siempre existe un estadístico suficiente asociado. En efecto, si $T(X_1, \dots, X_n) = (X_1, \dots, X_n)$ es evidente que no se pierde información. A la propia muestra le llamaremos estadístico suficiente trivial.*

Ejemplo 1. *Sea X_1, X_2 variables aleatorias independientes con distribución de Poisson de parámetro λ. Probar que $X_1 + X_2$ es suficiente para λ, y que $X_1 + 2X_2$ no lo es.*

Solución.

- Por reproductividad, si $X_1, X_2 \rightsquigarrow P(\lambda) \Rightarrow T = X_1 + X_2 \rightsquigarrow P(2\lambda)$. Probemos que T es suficiente para λ. Para ver que lo es, demostramos que la probabilidad de cualquier valor de la muestra condicionada a cualquier valor del estadístico es independiente del parámetro λ.

$$P_\lambda[X_1 = x_1, X_2 = x_2 \,/\, T(X_1, X_2) = X_1 + X_2 = t] = \frac{P_\lambda[X_1 = x_1, X_2 = x_2, T = t]}{P_\lambda[T = t]}$$

lo que nos da dos posibles resultados:

i) 0 si $x_1 + x_2 \neq t$ pues el numerador se anula por ser una intersección de sucesos que dan lugar a un suceso imposible y, por tanto, de probabilidad 0 (no es posible que $X_1 = x_1, X_2 = x_2$ y $T = X_1 + X_2 \neq x_1 + x_2$). En este caso, la probabilidad (siempre 0) es independiente de λ.

ii) Desarrollamos el caso $x_1 + x_2 = t$:

$$\frac{P_\lambda[X_1 = x_1, X_2 = x_2, T = t]}{P_\lambda[T = t]} \underbrace{=}_{\text{suceso seguro}} \frac{P_\lambda[X_1 = x_1, X_2 = x_2]}{P_\lambda[T = t]}$$

$$\underbrace{=}_{\text{ind}} \frac{P_\lambda[X_1 = x_1] \cdot P_\lambda[X_2 = x_2]}{P_\lambda[T = t]} = \frac{\cancel{e^{-\lambda}} \frac{\cancel{\lambda}^{x_1}}{x_1!} \cdot \cancel{e^{-\lambda}} \frac{\cancel{\lambda}^{x_2}}{x_2!}}{\cancel{e^{-2\lambda}} \frac{(2\cancel{\lambda})^t}{t!}} = [t = x_1 + x_2] = \frac{(x_1 + x_2)!}{2^{x_1+x_2} x_1! x_2!}$$

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

Por tanto, T es un estadístico suficiente.

- Para demostrar que un estadístico T no es suficiente basta con encontrar un valor concreto de T para el cual su distribución condicionada a que tome tal valor no es independiente del parámetro. Como en este caso $T = X_1 + 2X_2$, tomando $X_1 = 0, X_2 = 1$ se tiene $T = 2$.

$$\frac{P_\lambda[X_1 = 0, X_2 = 1 \,/\, T = 2]}{P_\lambda[T = 2]} = \frac{P_\lambda[X_1 = 0, X_2 = 1, T = 2]}{P_\lambda[T = 2]} \underbrace{=}_{\text{suceso seguro}} \frac{P_\lambda[X_1 = 0, X_2 = 1]}{P_\lambda[T = 2]}$$

$$\underbrace{=}_{\text{ind}} \frac{P_\lambda[X_1 = 0] \cdot P_\lambda[X_2 = 1]}{P_\lambda[X_1 = 2, X_2 = 0] + P_\lambda[X_1 = 0, X_2 = 1]} = \frac{e^{-\lambda} \frac{\lambda^0}{0!} \cdot e^{-\lambda} \frac{\lambda^1}{1!}}{e^{-\lambda} \frac{\lambda^2}{2!} \cdot e^{-\lambda} \frac{\lambda^0}{0!} + e^{-\lambda} \frac{\lambda^0}{0!} \cdot e^{-\lambda} \frac{\lambda^1}{1!}}$$

$$= \frac{\cancel{e^{-2\lambda}} \cdot \cancel{\lambda}}{\cancel{e^{-2\lambda}} \left(\frac{\lambda^{\cancel{2}}}{2} + \cancel{\lambda} \right)} = \frac{1}{\frac{\lambda}{2} + 1} = \frac{2}{\lambda + 2} \quad \text{dependiente del parámetro } \lambda.$$

Ejemplo 2. *Demostrar que, en n lanzamientos de una moneda, el número de caras es suficiente para el parámetro p.*

Solución. El lanzamiento de una moneda se puede modelar con la variable aleatoria $X \rightsquigarrow B(1, p)$, donde

$$X = \begin{cases} 1 & \text{si se obtiene cara} \\ 0 & \text{si se obtiene cruz} \end{cases}$$

Sea $(X_1, \ldots, X_n)$ una muestra aleatoria simple de X, demostrar que el número de caras es suficiente es equivalente a demostrar que el estadístico $T = \sum_{i=1}^n X_i \rightsquigarrow B(n, p)$ es suficiente. Para ver que lo es, demostramos que la probabilidad de cualquier valor de la muestra condicionada a cualquier valor del estadístico es independiente del parámetro p.

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

lo que nos da dos posibles resultados:

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

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

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

$$= \frac{P_p[X_1 = x_1] \cdots P_p[X_n = x_n]}{P_p[T = t]} = \frac{p^{x_1}(1-p)^{1-x_1} p^{x_2}(1-p)^{1-x_2} \cdots p^{x_n}(1-p)^{1-x_n}}{\binom{n}{t} p^t (1-p)^{n-t}}$$

$$[t = \sum x_i] = \frac{\cancel{p^t}\cancel{(1-p)^{n-t}}}{\binom{n}{t}\cancel{p^t}\cancel{(1-p)^{n-t}}} = \frac{1}{\binom{n}{t}} = \frac{1}{\binom{n}{\sum x_i}} \rightsquigarrow \text{Uniforme discreta}$$

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

Por tanto, T es un estadístico suficiente.

Es posible que haya ocasiones en las que se disponga del estadístico asociado a la muestra, pero no la muestra en sí. A partir de un estadístico suficiente, será posible reconstruirla.

Sea X una variable aleatoria con distribución $\{F_\theta, \theta \in \Theta\}$ y $(X_1, \ldots, X_n)$ una muestra aleatoria simple de X. Se realiza un experimento que consiste en observar la muestra y se obtienen unos valores concretos (observaciones) $(x_1, \ldots, x_n)$. Asociada a la variable X se tiene un estadístico $T(X_1, \ldots, X_n)$ que, al aplicarlo a los datos observados, nos da un valor concreto de T, es decir, $T(X_1, \ldots, X_n) = t$.

Si se pierden los datos de la m.a.s. pero se conoce el valor del estadístico y no se puede volver a observar la variable para obtener datos equivalentes a los anteriores, para reconstruir la muestra se sigue el siguiente proceso:

- Se considera una variable aleatoria X^* que tendrá la distribución de la variable aleatoria $X/T = t$, de la que se toma una muestra aleatoria simple $(X_1^*, \ldots, X_n^*)$ (por tanto, tal muestra tiene la misma distribución que $(X_1, \ldots, X_n)/T = t$).
- Si la distribución $(X_1, \ldots, X_n)/T = t$ es independiente del parámetro θ, por tanto es conocida, se pueden observar las variables $(X_1^*, \ldots, X_n^*)$, que es una muestra equivalente a $(X_1, \ldots, X_n)$ ya que las dos muestras dan lugar al mismo valor del estadístico:

$$T(X_1, \ldots, X_n) = T(X_1^*, \ldots, X_n^*)$$

Reconstrucción de una muestra.

Ejemplo 3. *Se supone que se ha lanzado una moneda 100 veces, y se sabe que se han obtenido 60 caras pero se han perdido los datos originales de si salió cara o cruz en cada tirada. Reconstruir la muestra.*

Solución. La variable aleatoria que modela si en una tirada se ha obtenido cara o cruz es la misma X del Ejemplo 2, luego de nuevo, el número de caras vendrá dada por el valor del estadístico

$T = \sum X_i$, con $(X_1, \ldots, X_n)$ una muestra aleatoria simple de X.

Ya demostramos que T es un estadístico suficiente, y sabemos que se distribuye según una uniforme discreta:

$$T \rightsquigarrow U\left(\binom{n}{t}\right)$$

Así, podemos seleccionar cualquier muestra en la que aparezcan sesenta 1 y cuarenta 0, puesto que esta será una de las $\binom{n}{t}$ posibles muestras que nos darán el mismo valor para el estadístico, $T = t$, y así obtener una muestra equivalente a la original.

2.1. Teorema de factorización de Neyman-Fisher.

Comprobar que un estadístico T es suficiente aplicando la definición (como en los ejemplos 1 y 2) puede ser complejo en el caso en el que no sepamos su distribución (en los casos anteriores, sí que era sencillo saberlo aplicando reproductividad). Además, si dada una familia de distribuciones nos piden encontrar un estadístico suficiente, la definición no nos da ningún método para poder hacerlo (la definición no es constructiva). Por tanto, necesitamos un criterio que nos permita tanto probar que un estadístico es suficiente como encontrar uno que lo sea.

Teorema de factorización de Neyman-Fisher.

Teorema 1. *Sea $(X_1, \ldots, X_n)$ una m.a.s. de una v.a. $X \rightsquigarrow F \in \{F_\theta, \theta \in \Theta\}$. Sea f_θ la función masa de probabilidad o función de densidad de X bajo F_θ y sea f_θ^n la f.m.p. o f.d.d. de la muestra bajo F_θ. Un estadístico $T(X_1, \ldots, X_n)$ es suficiente si y solo si,* ***para cualquier valor*** *de $\theta \in \Theta$ se cumple*

$$f_\theta^n(x_1, \ldots, x_n) = h(x_1, \ldots, x_n) g_\theta(T(x_1, \ldots, x_n)), \quad \forall (x_1, \ldots, x_n) \in \chi^n$$

donde h es independiente de θ y g_θ depende de $(x_1, \ldots, x_n)$ solo a través de $T(x_1, \ldots, x_n)$.

Demostración. La hacemos en el caso discreto (en el caso continuo basta sustituir la función masa de probabilidad por la función de densidad)

$\boxed{\Rightarrow}$ $T(X_1, \ldots, X_n)$ suficiente $\overset{¿?}{\Rightarrow} \underbrace{P[X_1 = x_1, \ldots, X_n = x_n]}_{f_\theta^n} = h(x_1, \ldots, x_n) g_\theta(T(x_1, \ldots, x_n))$.

$$\begin{aligned} P_\theta[X_1 = x_1, \ldots, X_n = x_n \;/\; T = t] &= \frac{P_\theta[X_1 = x_1, \ldots, X_n = x_n, T = t]}{P_\theta[T = t]} \\ &= \frac{P_\theta[X_1 = x_1, \ldots, X_n = x_n]}{P_\theta[T = t]} \end{aligned}$$

Por tanto, se tiene:

$$P_\theta[X_1 = x_1, \ldots, X_n = x_n] = f_\theta^n = \underbrace{P[X_1 = x_1, \ldots, X_n = x_n \;/\; T = t]}_{h(x_1, \ldots, x_n)} \underbrace{P_\theta[T = t]}_{g_\theta(T)}$$

$\boxed{\Leftarrow}$ $\underbrace{P[X_1 = x_1, \dots, X_n = x_n]}_{f_\theta^n} = h(x_1, \dots, x_n) g_\theta(T(x_1, \dots, x_n)) \ \forall \theta \overbrace{\Rightarrow}^{¿?} T$ suficiente.

$$P_\theta[X_1 = x_1, \dots, X_n = x_n \,/\, T = t] = \frac{P_\theta[X_1 = x_1, \dots, X_n = x_n, T = t]}{P_\theta[T = t]} = \begin{cases} 0 & T(x_1, \dots, x_n) \neq t \\ \underline{*} & T(x_1, \dots, x_n) = t \end{cases}$$

Desarrollamos $\underline{*}$, donde $P_\theta[X_1 = x_1, \dots, X_n = x_n, T = t] = P_\theta[X_1 = x_1, \dots, X_n = x_n]$ por ser $T = t$ un suceso seguro.

$$\underline{*} = \frac{P_\theta[(X_1, \dots, X_n) = (x_1, \dots, x_n)]}{P_\theta[T = t]} = \frac{h(x_1, \dots, x_n) g_\theta(T)}{\displaystyle\sum_{(x_1, \dots, x_n) \,/\, T(x_1, \dots, x_n) = t} P_\theta[X_1 = x_1, \dots, X_n = x_n]}$$

$$= \frac{h(x_1, \dots, x_n) g_\theta(T)}{\displaystyle\sum_{(x_1, \dots, x_n) \,/\, T(x_1, \dots, x_n) = t} h(x_1, \dots, x_n) g_\theta(T)} = \frac{h(x_1, \dots, x_n)}{\displaystyle\sum_{(x_1, \dots, x_n) \,/\, T(x_1, \dots, x_n) = t} h(x_1, \dots, x_n)}$$

que no depende de θ. Por tanto, el estadístico T es suficiente.

■

Propiedades de los estadísticos suficientes.

Corolario 1. *Si $T(X_1, \dots, X_n)$ es un estadístico suficiente para $\{F_\theta, \theta \in \Theta\}$, entonces:*

i) *$T(X_1, \dots, X_n)$ es suficiente para $\{F_\theta, \theta \in \Theta'\}$ para cualquier $\Theta' \subseteq \Theta$.*

ii) *Si $U(X_1, \dots, X_n)$ es un estadístico tal que $T = h(U)$, entonces U es suficiente para la misma familia.*

iii) *Toda transformación biunívoca de T proporciona un estadístico suficiente para la misma familia.*

Demostración.

i) Es evidente, pues se cumple $f_\theta^n(x_1, \dots, x_n) = h(x_1, \dots, x_n) g_\theta(T(x_1, \dots, x_n))$ para cualquier $\theta \in \Theta$ y $\theta' \in \Theta' \subseteq \Theta \Rightarrow \theta' \in \Theta$.

ii) Como T es suficiente, entonces $f_\theta^n(x_1, \dots, x_n) = r(x_1, \dots, x_n) g_\theta(T(x_1, \dots, x_n))$, donde r es una función independiente de θ, y si $T = h(U)$, se tiene entonces:

$$f_\theta^n(x_1, \dots, x_n) = r(x_1, \dots, x_n) g_\theta(h(U)(x_1, \dots, x_n)) = r(x_1, \dots, x_n) \underbrace{(g_\theta \circ h)}_{h^*}(U(x_1, \dots, x_n))$$

luego U es, en efecto, un estadístico suficiente.

iii) Sea m tal transformación biunívoca. Aplicando el teorema de factorización, se tiene:

$$\begin{aligned} f_\theta^n(x_1, \dots, x_n) &= h(x_1, \dots, h_n) g_\theta(T(x_1, \dots, x_n)) = h(x_1, \dots, x_n) g_\theta(m^{-1}(m(Tx_1, \dots, x_n))) \\ &= h(x_1, \dots, x_n) \underbrace{(g_\theta \circ m^{-1})}_{g^*}(m(T(x_1, \dots, x_n))) \end{aligned}$$

Por tanto, $m(T)$ es suficiente.

Observamos que la condición *ser biunívoca* es imprescindible, pues si la función no lo es, entonces es posible que no exista su función inversa.

■

Observación 2. *Podemos recordar mejor lo que dice el corolario 1 si lo reescribimos de la siguiente forma:*

i) *Si un estadístico es suficiente para una familia de distribuciones, lo es también para cualquier subfamilia suya.*

ii) *Si tengo un estadístico suficiente U que es función de otro estadístico T, entonces T también lo es.*

iii) *Para que una función de un estadístico suficiente sea otro estadístico suficiente, tal función debe ser inyectiva.*

Importante para un tipo test.

Ejemplo 4. .

- *Si un estadístico T es suficiente para el parámetro u de una familia de distribuciones Gamma, $F_1 = \{\Gamma(u, \lambda), u, \lambda \in \mathbb{R}^+\}$, entonces T es suficiente para el parámetro n de una familia de distribuciones Erlang, $F_2 = \{\varepsilon(n, \lambda), n \in \mathbb{N}, \lambda \in \mathbb{R}^+\}$, pues $F_2 \subseteq F_1$.*
- *Si $T = 3X_{(1)}^2 + 4$ es un estadístico suficiente, entonces el estadístico $X_{(1)}$ también lo es.*
- *Si $T = X_{(1)}$ es suficiente, entonces $3T^2 + 4$ no tiene por qué serlo, ya que $y = x^2$ no es una tranformación biunívoca.*

La diferencia entre estos dos últimos apartados radica en que en el primero se parte de que una transformación es suficiente, luego el estadístico que ha sido transformado también lo es, mientras que en el segundo se tiene un estadístico suficiente y se quiere comprobar si lo es también su transformación.

- *Si $T = X_{(n)}$ es un estadístico suficiente, entonces T^3 también lo es, ya que $y = x^3$ es una función inyectiva.*

Observación 3. *El teorema de factorización también es cierto si el parámetro θ es multidimensional, en cuyo caso el estadístico también es multidimensional. En particular, aunque el parámetro sea de dimensión 1, el estadístico suficiente puede tener dimensión mayor que 1. En general, se verifica:*

$$\textit{Dimensión del estadístico suficiente} \geq \textit{Dimensión del parámetro}$$

La demostración de tal hecho es evidente, puesto que si la familia de distribuciones depende de un conjunto de parámetros desconocidos, entonces en la función de densidad o función masa de probabilidad de la muestra aparecerán todos los parámetros, luego la función g dependiente de los parámetros desconocidos a través del estadístico considerado debe contenerlos a todos.

Ejemplo 5. *Sea X una variable aleatoria con distribución en $\{\mathcal{N}(\mu, \sigma^2); \mu \in \mathbb{R}, \sigma^2 > 0\}$. Estudiar la suficiencia en los casos de σ_0^2 conocida, μ_0 conocida y μ y σ^2 desconocidas.*

Solución. Si $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, entonces $f_{\mu,\sigma^2}(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{(x-\mu)^2}{2\sigma^2}\right)$.

- Supongamos σ_0^2 conocida. Entonces, estudiamos la suficiencia para el parámetro μ.

$$\begin{aligned}
f_\mu^n(x_1,\ldots,x_n) &= \frac{1}{\sqrt{2\pi\sigma_0^2}}\exp\left(-\frac{(x_1-\mu)^2}{2\sigma_0^2}\right)\cdots\frac{1}{\sqrt{2\pi\sigma_0^2}}\exp\left(-\frac{(x_n-\mu)^2}{2\sigma_0^2}\right)\\
&= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{\sum(x_i-\mu)^2}{2\sigma_0^2}\right) = \left[-\sum(x_i-\mu)^2 = -\sum x_i^2 - n\mu^2 + 2\mu\sum xi\right]\\
&= \underbrace{\frac{1}{(2\pi\sigma_0^2)^{n/2}}\exp\left(-\frac{\sum x_i^2}{2\sigma_0^2}\right)}_{h(x_1,\ldots,x_n)}\underbrace{\exp\left(\frac{-n\mu}{2\sigma_0^2}\right)\exp\left(\frac{\mu\sum x_i}{\sigma_0^2}\right)}_{g_\mu(T)}
\end{aligned}$$

Por el teorema de factorización, se tiene que $T=\sum_{i=1}^n X_i$ es un estadístico suficiente para μ.

- Supongamos μ_0 es conocida. Entonces, estudiamos la suficiencia para el parámetro σ^2.

$$\begin{aligned}
f_{\sigma^2}^n(x_1,\ldots,x_n) &= \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left(\frac{-\sum(x_i-\mu_0)^2}{2\sigma^2}\right)\\
&= \frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left(\frac{-\sum x_i^2}{2\sigma^2}\right)\exp\left(\frac{-n\sigma_0^2}{2\sigma^2}\right)\exp\left(\frac{\mu_0\sum x_i}{\sigma_0^2}\right)\\
&= \underbrace{\frac{1}{(2\pi)^{n/2}}}_{h(x_1,\ldots,x_n)}\underbrace{\frac{1}{\sigma^n}\exp\left(\frac{-\sum x_i^2}{2\sigma^2}\right)\exp\left(\frac{-n\sigma_0^2}{2\sigma^2}\right)\exp\left(\frac{\mu_0\sum x_i}{\sigma_0^2}\right)}_{g_{\sigma^2}(T)}
\end{aligned}$$

Por el teorema de factorización, $T=\left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$ es un estadístico suficiente para σ^2.

Podríamos también haber decidido no desarrollar $f_{\sigma^2}^n$, ya que en ese caso escribirla como el producto de h y g_{σ^2} es trivial. En ese caso, habríamos obtenido que el estadístico $T=\sum(X_i-\mu_0)^2$ es suficiente, llegando a la misma conclusión pues T, así definido, es la transformación de los estadísticos $T_1=\sum X_i^2$ y $T_2=\sum X_i$. Observemos que, tal y como se dice en la observación 3, aunque el parámetro σ^2 es unidimensional, el estadístico suficiente para σ^2 es bidimensional.

- Basta repetir el razonamiento anterior con $\mu_0=\mu$, para obtener entonces:

$$h(x_1,\ldots,x_n)=\frac{1}{(2\pi)^{n/2}} \qquad g_{\mu,\sigma^2}(T)=\frac{1}{\sigma^n}\exp\left(\frac{-\sum x_i^2}{2\sigma^2}\right)\exp\left(\frac{-n\sigma_0^2}{2\sigma^2}\right)\exp\left(\frac{\mu\sum x_i}{\sigma_0^2}\right)$$

Así, $T=\left(\sum_{i=1}^n X_i^2, \sum_{i=1}^n X_i\right)$ es un estadístico suficiente para (μ,σ^2).

2.1.1. Suficiencia minimal.

Aunque no vayamos a estudiarlos, es conveniente destacar que, de entre todos los estadísticos suficientes para un parámetro θ de cualquier familia de distribuciones $\{F_\theta, \theta\in\Theta\}$, siempre existirá uno que sea función de todos los demás. Diremos que tal estadístico es *suficiente minimal* (al haber sido construido a partir de otros estadísticos contiene toda su información luego es, de entre todos, el que produce una menor pérdida de información relevante).

Para distribuciones discretas y continuas el estadístico minimal suficiente siempre existe, es único, y se obtiene aplicando el teorema de factorización a cocientes entre funciones masa de probabilidad o funciones de densidad puntuales en distintos puntos.

3. Estadísticos completos.

Al contrario que el concepto de suficiencia, la *completitud* es difícil de explicar de forma intuitiva. A pesar de ello, diremos que la completitud es una propiedad que descarta la existencia de funciones unidimensionales medibles no nulas con esperanza cero bajo el parámetro θ desconocido.

Familia de distribuciones completas.

Definición 3. *Sea* $\{F_\theta, \theta \in \Theta\}$ *una familia de distribuciones con función de densidad o función masa de probabilidad* $\{f_\theta, \theta \in \Theta\}$. *Diremos que dicha familia es completa si, para cualquier función medible unidimensional* g *cumpliendo* $E_\theta[g(X)] = 0, \forall\theta \in \Theta$ *se tiene que*

$$P_\theta[g(X) = 0] = 1, \qquad \forall\theta \in \Theta$$

En la Definición 3 se ha considerado una transformación g de la variable aleatoria X. Consideramos ahora una transformación de un estadístico T. Diremos que este es completo si su distribución no admite la existencia de funciones unidimensionales medibles no nulas con esperanza cero bajo el parámetro desconocido.

Estadístico completo.

Definición 4. *Un estadístico* $T(X_1, \ldots, X_n)$ *se dice que es completo para la familia de distribuciones de* X *si para cualquier función medible unidimensional* g *se tiene:*

$$E_\theta[g(T(X_1, \ldots, X_n))] = 0, \forall\theta \in \Theta \Rightarrow P_\theta[g(T(X_1, \ldots, X_n)) = 0] = 1, \forall\theta \in \Theta$$

En este caso no estudiaremos ningún teorema con el que comprobar que un estadístico es completo, sino que tendremos que hacerlo siempre aplicando la definición.

Observación 4. *El concepto de estadístico completo suele ir asociado al concepto de estadístico suficiente. De hecho, si se pide dar un estadístico suficiente y completo, lo que haremos será tomar el estadístico suficiente que proporciona el teorema de factorización, y comprobar después si es completo.*

Ejemplo 6. *Sea* $(X_1, \ldots, X_n)$ *una muestra aleatoria simple* $X \rightsquigarrow B(1, p), p \in (0, 1)$. *Encontrar un estadístico suficiente y completo asociado a la muestra.*

Solución. Si $X \rightsquigarrow B(1, p)$, entonces $P_p[X = x] = p^x(1-p)^{1-x}$. Así, la función masa de probabilidad de una m.a.s. de tamaño n de X viene dada por:

$$P_p[X_1 = x_1, \ldots, X_n = x_n] = p^{x_1}(1-p)^{1-x_1} \cdots p^{x_n}(1-p)^{1-x_n} = p^{\sum x_i}(1-p)^{n-\sum x_i}$$
$$= 1 \cdot \underbrace{(1-p)^n \left(\frac{p}{1-p}\right)^{\sum x_i}}_{g_p(T)} \Rightarrow T = \sum_{i=1}^{n} X_i \text{ es suficiente.}$$

Si $f_\theta^n = g_\theta(T)$, podemos escribir $h(x_1, \ldots, x_n) = 1$ para poder aplicar el teorema de factorización.

Comprobemos que T es completo, es decir, $\mathrm{E}_p[g(T)] = 0, \forall p \in (0,1) \overbrace{\Rightarrow}^{¿?} P_p[g(T) = 0] = 1, \forall p$. Para ello, utilizamos que $T \rightsquigarrow B(n,p)$. Así:

$$E_p[g(T)] = \underbrace{\sum_{t=0}^{n} g(t) \binom{n}{t} p^t (1-p)^{n-t}}_{(1)} = 0, \forall p \in (0,1)$$

donde, desarrollando (1) se tiene que

$$(1) = \sum_{t=0}^{n} g(t) \binom{n}{t} (1-p)^n \left(\frac{p}{1-p}\right)^t = 0 = \underbrace{(1-p)^n}_{\neq 0} \sum_{t=0}^{n} g(t) \binom{n}{t} \left(\frac{p}{1-p}\right)^t$$

y esto es posible si y solo si $\underbrace{\sum_{t=0}^{n} g(t) \binom{n}{t} \left(\frac{p}{1-p}\right)^t}_{(2)} = 0$, con (2) un polinomio en la variable $\dfrac{p}{1-p}$.

$$(2) = g(0) + g(1)n\left(\frac{p}{1-p}\right) + g(2)\binom{n}{2}\left(\frac{p}{1-p}\right)^2 + \cdots + g(n)\binom{n}{n}\left(\frac{p}{1-p}\right)^n = 0, \quad \forall p \in (0,1)$$

Así, tenemos que el polinomio dado en (2) se anula siempre, y esto es así si y solamente si todos sus coeficientes son cero. Por tanto:

$$\begin{array}{ccccc}
g(0) = 0 & g(1)\binom{n}{1} = 0 & g(2)\binom{n}{2} = 0 & \cdots & g(n)\binom{n}{n} = 0 \\
\Downarrow & \Downarrow & \Downarrow & \cdots & \Downarrow \\
g(0) = 0 & g(1) = 0 & g(2) = 0 & \cdots & g(n) = 0
\end{array}$$

Por tanto, se tiene que $\forall t \in \{0, \ldots, n\}, g(t) = 0$. Sin embargo, como es posible que $g(t)$ también se anule para otros valores de t, entonces se cumple

$$\{t / g(t) = 0\} \supseteq \{t \in \{0, \ldots, n\}\}$$

Por la monotonía de la probabilidad, se tendrá entonces

$$P_p\{g(T) = 0\} \geq P_p\{T \in \{0, \ldots, n\}\}$$

Sin embargo, T sigue una distribución binomial cuyo primer parámetro es n, luego solo puede tomar valores entre 0 y n, así que el suceso $\{T \in \{0, \ldots, n\}\}$ es seguro. Por otro lado, como cualquier suceso tiene siempre probabilidad menor o igual que 1, entonces se cumple la siguiente cadena:

$$1 \geq P_p\{g(T) = 0\} \geq P_p\{T \in \{0, \ldots, n\}\} = 1 \Rightarrow P_p\{g(T) = 0\} = 1$$

Por tanto, T es un estadístico completo.

Ejemplo 7. *Sea $(X_1, \dots, X_n)$ una muestra aleatoria simple de $X \rightsquigarrow U(0,\theta), \theta > 0$. Encontrar un estadístico suficiente y completo asociado a la muestra.*

Solución. Si $X \rightsquigarrow U(0,\theta)$, entonces $f_\theta(x) = \dfrac{1}{\theta}, 0 < x < \theta$. Por tanto, se tiene:

$$f_\theta^n = \frac{1}{\theta} \cdots \frac{1}{\theta} = \left(\frac{1}{\theta}\right)^n, 0 < x_i < \theta, \forall i = 1, \dots, n$$

Es decir, para que la función de densidad de la muestra aleatoria simple sea no nula, todos sus valores deben estar también entre 0 y θ. Por tanto, para que esto sea así, debe cumplirse que:

- El menor valor de la m.a.s. debe ser mayor que 0, puesto que esto asegura que todos los demás valores también lo son.
- El mayor valor de la m.a.s. debe ser menor que θ, ya que esto garantiza que el resto de valores también están por debajo de θ.

Por tanto, f_θ^n es no nula cuando se cumple $X_{(1)} > 0$ y $X_{(n)} < \theta$. Así, considerando la función indicadora, se llega finalmente a que la función de densidad de la muestra aleatoria simple es:

$$f_\theta^n = f_\theta(x_1, \dots, x_n) = \left(\frac{1}{\theta}\right)^n I_{[X_{(1)}>0]}(x_1, \dots, x_n) I_{[X_{(n)}<\theta]}(x_1, \dots, x_n)$$

Tomando $h(x_1, \dots, x_n) = I_{[X_{(1)}>0]}(x_1, \dots, x_n)$, $g_\theta(T) = \left(\dfrac{1}{\theta}\right)^n I_{[X_{(n)}<\theta]}(x_1, \dots, x_n)$, se llega a que $T(X_1, \dots, X_n) = X_{(n)}$ es un estadístico suficiente. Ahora, debemos calcular su función de densidad.

$$F_{X(n)}(x) = (F_X(x))^n = \left(\frac{x}{\theta}\right)^n \Rightarrow f_{X_{(n)}}(x) = n(F_X(x))^{n-1} f_\theta(x) = n\left(\frac{x}{\theta}\right)^{n-1} \cdot \frac{1}{\theta},\ 0 < x < \theta$$

Comprobemos que T es completo. Para ello, calculamos $\mathrm{E}_\theta[g(T)]$.

$$\mathrm{E}_\theta[g(T)] = \int_0^\theta g(t) f_{X_{(n)}}(t) dt = \int_0^\theta g(t) n \frac{t^{n-1}}{\theta^n} = \frac{n}{\theta^n} \int_0^\theta g(t) t^{n-1} dt$$

Como tenemos $\mathrm{E}_\theta[g(T)] = 0, \forall\theta$ para comprobar la completitud, entonces:

$$\underbrace{\frac{n}{\theta^n}}_{\neq 0} \int_0^\theta g(t) t^{n-1} dt = 0\ \forall\theta \iff \int_0^\theta g(t) t^{n-1} dt = 0\ \forall\theta$$

Por el Teorema Fundamental del Cálculo, tal integral existe, y es

$$\int_0^\theta g(t) t^{n-1} dt = G(\theta) - G(0)$$

con G una primitiva de $g(t)t^{n-1}$. Así, derivando con respecto a θ y teniendo en cuenta que $G(0)$ no depende de θ, se tiene:

$$(G(\theta) - G(0))' = G'(\theta) = g(\theta) \underbrace{\theta^{n-1}}_{\neq 0} = 0\ \forall\theta \Rightarrow g(\theta) = 0\ \forall\theta$$

Ahora, razonando como en el Ejemplo 6 se tiene:

$$\{t/g(t) = 0\} \supseteq \{t > 0\} \Rightarrow 1 \geq P_\theta\{g(T) = 0\} \geq P_\theta\{T > 0\} = 1 \Rightarrow P_\theta\{g(T) = 0\} = 1$$

Por tanto, $X_{(n)}$ es un estadístico completo.

4. Suficiencia y completitud en familias exponenciales.

Estudiamos las llamadas familias de distribuciones exponenciales paramétricas, para las cuales el proceso de obtención de estadísticos suficientes y completos puede simplificarse si se dan las circunstancias adecuadas.

Sea $\mathcal{F} = \{F_\theta, \theta \in \Theta\}$ una familia de distribuciones de probabilidad paramétricas con función de densidad o función masa de probabilidad $\{f_\theta, \theta \in \Theta\}$.

4.1. Familia exponencial uniparamétrica. Familia exponencial k-paramétrica.

Familia exponencial uniparamétrica.

Definición 5. *Se dice que la familia de distribuciones $\mathcal{F}$ es exponencial uniparamétrica si se cumplen las siguientes tres condiciones:*

i) El espacio paramétrico es un intervalo real: $\Theta \subseteq \mathbb{R}$.

ii) El conjunto de valores de la variable no depende de θ:

$$\chi = \{x/f_\theta(x) > 0\} \text{ independiente de } \theta,\ \forall\theta \in \Theta$$

iii) Existen funciones real-valuadas $Q(\theta)$ y $D(\theta)$, definidas sobre Θ, y existen funciones medibles Borel T y S, también real valuadas, tales que:

$$\forall\theta \in \Theta,\ f_\theta(x) = \exp[Q(\theta)T(x) + D(\theta) + S(x)],\ x \in \chi$$

Importante para un tipo test.

Ejemplo 8. *Sea X una variable aleatoria con función de densidad de probabilidad $f_\theta(x) = e^{x-2\theta}, \theta > 0, x > \theta$. Entonces, aunque se cumplan las condiciones* i) *y* iii) *de la Definición 5, observemos que no se cumple* ii) *pues el conjunto χ está formado por valores que deben ser mayores que θ. Por tanto, χ depende de θ, luego en este caso la distribución no es exponencial uniparamétrica.*

Ejemplo 9. *Comprobar que la familia $\{P(\lambda), \lambda > 0\}$ es una familia exponencial uniparamétrica.*

Solución. Si $X \rightsquigarrow P(\lambda)$, entonces $P_\lambda[X = x] = e^{-\lambda}\dfrac{\lambda^x}{x!}, x = 0, 1, 2, \ldots$. Por tanto:

i) $\Theta = \mathbb{R}^+$ intervalo de $\mathbb{R}$, luego se cumple la primera condición.

ii) Una distribución de Poisson toma valores positivos o el 0, independientemente de cuál sea su parámetro. Por tanto, $\chi = \{x/P_\lambda[X = x] > 0\} = \{0, 1, 2, \ldots\}$ independiente de λ, luego se cumple la segunda condición.

iii) Reescribimos su función masa de probabilidad. Para ello, debemos calcular primero su exponencial y después tomar logaritmos para que no cambie:

$$P_\lambda[X = x] = e^{-\lambda}\frac{\lambda^x}{x!} = \exp\left\{\ln\left(e^{-\lambda}\frac{\lambda^x}{x!}\right)\right\} = \exp\{-\lambda + x\ln\lambda - \ln x!\}$$

y tomamos $D(\lambda) = -\lambda$, $Q(\lambda)T(x) = x\ln\lambda$, $S(x) = \ln x!$ para que se cumpla la condición.

Por tanto, es una distribución exponencial uniparamétrica.

Ejemplo 10. *Comprobar que la familia $\{B(k_0, p), p \in (0,1)\}$ es una familia exponencial uniparamétrica.*

Solución. Si $X \rightsquigarrow B(k_0, p), p \in (0,1)$, entonces $P_p[X = x] = \binom{k_0}{p} p^x (1-p)^{k_0 - x}, x = 0, 1, \ldots, k_0$.
Por tanto:

i) $\Theta = (0,1)$ intervalo de $\mathbb{R}$, luego se cumple la primera condición.

ii) Una distribución Binomial de parámetros k_0 y p puede tomar valores desde el 0 hasta k_0, parámetro que no estamos considerando desconocido. Por tanto, $\chi = \{x/P_p[X = x] > 0\} = \{0, 1, 2, \ldots, k_0\}$ independiente de p, luego se cumple la segunda condición.

iii)

$$\begin{aligned} P_p[X = x] &= \exp\left\{\ln\binom{k_0}{x} + x\ln p + (k_0 - x)\cdot\ln(1-p)\right\} \\ &= \exp\left\{\ln\binom{k_0}{x} + x(\ln p - \ln(1-p)) + k_0\ln(1-p)\right\} \end{aligned}$$

y tomamos $S(x) = \ln\binom{k_0}{x}$, $Q(p)T(x) = x(\ln p - \ln(1-p))$, $D(p) = k_0\ln(1-p)$. para que se cumpla la condición.

Por tanto, es una distribución exponencial uniparamétrica.

Familia exponencial k-paramétrica.

Definición 6. *Se dice que la familia de distribuciones $\mathcal{F}$ es exponencial k-paramétrica si se cumplen las siguientes tres condiciones:*

i) El espacio paramétrico es un intervalo real multidimensional: $\Theta \subseteq \mathbb{R}^k$.

ii) El conjunto de valores de la variable no depende de $\theta = (\theta_1, \ldots, \theta_k)$:

$$\chi = \{x/f_\theta(x) > 0\}, \forall\theta \in \Theta$$

iii) Existen funciones real-valuadas $Q_h(\theta)$ y $D_h(\theta)$ con $h = 1, \ldots, k$ definidas sobre Θ, y existen funciones medibles Borel T y S, también real valuadas, tales que:

$$\forall\theta \in \Theta,\ f_\theta(x) = \exp\left[\sum_{h=1}^{k} Q_h(\theta)T_h(x) + D(\theta) + S(x)\right],\ x \in \chi$$

Ejemplo 11.

- *Comprobar que la familia de distribuciones normales $\{\mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0\}$ es bi-paramétrica en los dos parámetros.*
- *Comprobar que es uniparamétrica en cada parámetro.*

Solución. Si $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0$, entonces $f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), x \in \mathbb{R}$.

- Caso biparamétrico.
  i) $\Theta = (\mathbb{R}, \mathbb{R}^+)$ intervalo de $\mathbb{R}^2$, luego se cumple la primera condición.
  ii) $\chi = \{x/f_{\mu,\sigma^2} > 0\} = \mathbb{R}$ independiente de μ y σ^2, luego se cumple la segunda condición.
  iii) Haciendo la exponencial de la función de densidad y después tomando logaritmos, se llega a:

$$\begin{aligned} f_{\mu,\sigma^2}(x) &= \exp\left\{-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{(x-\mu)^2}{2\sigma^2}\right\} \\ &= \exp\left\{-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{x^2}{2\sigma^2} - \frac{\mu^2}{2\sigma^2} + \frac{\mu x}{\sigma^2}\right\} \end{aligned}$$

  Para que se cumpla la condición, se toman las funciones $S(x) = -\frac{1}{2}\ln(2\pi)$, $D(\mu,\sigma^2) = -\frac{1}{2}\ln\sigma^2 - \frac{\mu^2}{2\sigma^2}$, $T_1(x) = x^2$, $Q_1(\mu,\sigma^2) = -\frac{1}{2\sigma^2}$, $T_2(x) = x$, $Q_2(\mu,\sigma^2) = \frac{\mu}{\sigma^2}$.

  Por tanto, es una distribución exponencial bi-paramétrica.

  Observemos que podríamos haber cogido una función distinta para $T_1(x)$ (podríamos haberle puesto un denominador, por ejemplo). En general, nos conviene (y se verá en el teorema siguiente) coger, para $T_h(x)$, las funciones más simples.

- Para el caso uniparamétrico, con cambios mínimos se puede comprobar a partir del punto anterior que se cumplen las condiciones *i)* y *ii)*
  - Si suponemos que μ es el parámetro desconocido, entonces para que se cumpla *iii)* escogemos las funciones:

$$S(x) = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 - \frac{x^2}{2\sigma^2}, \quad D(\mu) = \frac{-\mu^2}{2\sigma^2}, \quad Q(\mu)T(x) = \frac{\mu x}{\sigma^2}$$

  - Si suponemos que el parámetro que no conocemos es σ^2, entonces se toman las siguientes funciones:

$$S(x) = -\frac{1}{2}\ln(2\pi), \quad D(\sigma^2) = -\frac{1}{2}\ln\sigma^2, \quad Q(\sigma^2)T(x) = \frac{1}{\sigma^2}\left(-\frac{x^2+\mu^2}{2} + \mu x\right)$$

  En ambos casos, se tiene que la distribución normal es exponencial uniparamétrica.

4.2. Suficiencia y completitud.

Suficiencia y completitud en familias exponenciales.

Teorema 2. *Sea una variable aleatoria $X \rightsquigarrow \{F_\theta, \theta \in \Theta\}$ con familia de funciones asociadas $\{f_\theta, \theta \in \Theta\}$, siendo f.d.d. o f.m.p. según el caso. Si la familia de distribuciones es exponencial k-paramétrica, entonces la familia de distribuciones asociadas a una muestra aleatoria simple $(X_1, \dots, X_n)$ de X es también exponencial k-paramétrica:*

$$f_\theta^n(x_1,\dots,x_n) = \exp\left\{\sum_{h=1}^{k} Q_h(\theta)\left(\sum_{i=1}^{n} T_h(x_i)\right) + \sum_{i=1}^{n} S(x_i) + nD(\theta)\right\}, (x_1,\dots,x_n) \in \chi^n$$

Además, se tiene:

- *El estadístico* $\left(\sum_{i=1}^{n} T_1(X_i), \ldots, \sum_{i=1}^{n} T_k(X_i)\right)$ *es suficiente para* θ.
- *Si* $k \leq n$ *y el conjunto imagen de la función* $Q(\theta) = (Q_1(\theta), \ldots, Q_k(\theta))$ *contiene a un abierto de* $\mathbb{R}^k$, *el estadístico* $\left(\sum_{i=1}^{n} T_1(X_i), \ldots, \sum_{i=1}^{n} T_k(X_i)\right)$ *es también completo.*

Demostración. Demostramos solo la primera parte del teorema: si la familia de distribuciones es exponencial, también lo es la familia de distribuciones asociadas a una muestra aleatoria simple.

i) Θ es un intervalo de $\mathbb{R}^k$ por ser el producto de intervalos de $\mathbb{R}$.

ii) $\chi^n = \chi \times \cdots \times \chi$ es independiente de θ por ser el producto cartesiano de conjuntos independientes de θ.

iii)

$$\begin{aligned}
f_\theta(n) &= f_\theta(x_1) \cdots f_\theta(x_n) = \exp\{\cdots\} \cdots \exp\{\cdots\} \\
&= \exp\left\{\sum_{h=1}^{k} T_h(x_1)Q_h(\theta) + D(\theta) + S(x_1) + \cdots + \sum_{h=1}^{k} T_h(x_n)Q_h(\theta) + D(\theta) + S(x_n)\right\} \\
&= \exp\left\{nD(\theta) + \sum_{i=1}^{n} S(x_i) + \sum_{i=1}^{n}\sum_{h=1}^{k} T_k(x_i)Q_h(\theta)\right\} \\
&= \exp\left\{nD(\theta) + \sum_{i=1}^{n} S(x_i) + \sum_{h=1}^{k} Q_h(\theta)\sum_{i=1}^{n} T_h(x_i)\right\}
\end{aligned}$$

■

Ejemplo 12. *Sea* X *una variable aleatoria con distribución* $\mathcal{N}(\mu, \sigma^2)$ *con* $\mu \in \mathbb{R}$ *y* $\sigma^2 > 0$. *Hallar un estadístico suficiente y completo basado en una muestra aleatoria simple de tamaño* n.

Solución. Recordemos que, en este caso, se tenían las funciones

$$T_1(x) = x^2,\ Q_1(\mu, \sigma^2) = -\frac{1}{2\sigma^2},\ T_2(x) = x,\ Q_2(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$$

Por el Teorema 2 se tiene que $\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ es un estadístico suficiente para (μ, σ^2).

Como $\text{Im}(Q_1) = \mathbb{R}^-, \text{Im}(Q_2) = \mathbb{R}$ y evidentemente $(\mathbb{R}^-, \mathbb{R})$ contiene a un abierto de $\mathbb{R}^2$ (de hecho, él mismo lo es), entonces $\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i^2\right)$ es también completo.