Apuntes de Estadísitica

por
José Antonio Belinchón

Última actualización Diciembre 2007

Índice general

Prólogo

La idea fundamental de esta notas confecionadas a modo de resumen (personal) es la de tener a mano un recordatorio de por donde iban los tiros. Sólo se demuestran los teoremas fundamentales y se acompoña el texto con una serie de ejercios más o menos trabajados. En modo alguno pretenden sustituir (porque es implosible) los manuales clásicos o las notas de clase de un profesor. Es decir, estas notas estan confeccionadas a modo de refrito entre las notas de clase y de distintos libros clásicos como los siguientes:

1. Julián de la Horra. Estadística Apliaca. Diaz de Santos (2003)
2. M.A. Gómez Villegas Inferencia EstadísticaDiaz de Santos (2005)
3. Novo, V. Problemas de Cálculo de Probabilidades y Estadística UNED (1993).
4. Daniel Peña. Fundamentos de Estadística. Alianza Editorial (2001)
5. R. Vélez Ibarrola et al. Principios de Inferencia Estadística UNED (1993)

todo ello aderezado (como he indicado antes) con una serie de ejemplos (ejercicios donde se aplica de forma inmediata los conceptos teóricos expuestos) desarrollados (eso espero) al final de cada capitulillo (todos ellos muy sencillos).

Agradezco al Prof. Julián de la Horra el haberme pasado las secciones 3.4 y 4.4 de estas notas, éstas son enteramente suyas.

ADVERTENCIA: No están concluidas y es muy posible que hayan sobrevivido numerosas erratas. Toda observación en este sentido es bien recibida.

Capítulo 1

Muestreo aleatorio.

1.1. Muestreo aleatorio simple.

Es el objetivo fundamental de la inferencia estadística. Obtener conclusiones razonadas sobre una característica X de una población a partir de los resultados de una muestra. Dicha característica X será una variable aleatoria (v.a.) (discreta o contínua) y por lo tanto estará descrita por su función de masa o de densidad, escrita de forma genérica f.

Observación 1.1.1 *f no es completamente conocida.*

Definición 1.1.1 *Una muestra aleatoria (m.a.) de una característica X cuya distribución es $f(x)$ i.e. $X \sim f(x)$, es un vector aleatorio $(X_1, ..., X_n)$ tal que*

1. *La distribución marginal de cada X_i viene dada por la misma distribución que la característica i.e. $(X_i)_{i=1}^n \sim f(x)$.*
2. *$(X_i)_{i=1}^n$ son independientes.*

El significado intuitivo es el siguiente.

a Las observaciones son representaciones de la población que estoy estudiando.

b *La muestra es con reemplazamiento.*

1. por simplicidad matemática (el caso no independiente es mucho más complejo)
2. la muestra es con reemplazamiento, significa que la muestra (su tamaño) es pequeño en comparación con la población.

Fundamentalmente por lo tanto la función de masa o de densidad de una muestra aleatoria vendrá dada por

$$f(x_1, ..., x_n) \underset{indpen.}{=} \prod_{i=1} f(x_i). \tag{1.1}$$

Ejemplo 1.1.1 *Supongamos que X es del tipo, estatura, etc... entonces es razonable pensar que*

$$X \sim N(\mu, \sigma^2),$$

donde falta por conocer μ y σ^2.

Un problema de inferencia paramétrica es aquel en el que queremos estudiar una característica $X \sim f_\theta(x)$, donde θ es el parámetro a encontrar, $\theta \in \Theta$ (espacio paramétrico). Nuestro propósito será encontrar conclusiones sobre θ por lo tanto la eq. (1.1) se reescribe como

$$f_\theta(x_1, ..., x_n) \underset{indpen.}{=} \prod_{i=1} f_\theta(x_i). \tag{1.2}$$

Definición 1.1.2 *Un estadístico T es una función medible de la muestra $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. Formalmente T es un variable o vector aleatorio. T viene a ser el resumen de la muestra.*

Ejemplo 1.1.2 *Veamos algunos ejemplos de estadísticos:*

1. *Media muestral*
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \tag{1.3}$$

2. *Varianza muestral*
$$var(X) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right), \tag{1.4}$$

3. *Cuasi-varianza muestral*
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right), \tag{1.5}$$

4. *Estaísticos de orden*
$$\left(X_{(1)}, ..., X_{(n)}\right), \tag{1.6}$$
donde
$$\begin{aligned} X_{(1)} &= \min\left(X_1, ..., X_n\right), \\ &\ \ ... \\ X_{(n)} &= \max\left(X_1, ..., X_n\right). \end{aligned} \tag{1.7}$$

Proposición 1.1.1 *Propiedades de los estadísticos. Sea $(X_1, ..., X_n)$ una m.a. tal que $E(X) = \mu$ y $var(X) = \sigma^2$, entonces:*

1. $E\left[\overline{X}\right] = \mu$,
   se observa que
   $$E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\left[X_i\right] \overset{indp.}{=} \frac{1}{n} nE\left[X_i\right] = \mu$$
2. $var\left[\overline{X}\right] = \frac{\sigma^2}{n}$,
   Vemos que
   $$var\left[\overline{X}\right] = var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} var\left(\sum_{i=1}^{n} X_i\right) \overset{indp.}{=} \frac{1}{n^2} nvar(X_i) = \frac{\sigma^2}{n}$$
3. $E\left[S^2\right] = \sigma^2$.

Sea $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, ó $\sum_{i=1}^{n} X_i$. En general tenemos, a partir de la proposición anterior, que $E\left[\overline{X}\right] = \mu$, $var\left[\overline{X}\right] = \frac{\sigma^2}{n}$, etc... pero si estamos en la distribución de $\overline{X}$ ó en la de $\sum_{i=1}^{n} X_i$, ¿qué podemos hacer?. Estudiar su función característica. Sea $(X_1, ..., X_n)$ una m.a. de X con función característica $\varphi_X(t)$, entonces encontramos que

$$\varphi_{\sum X_i}(t) \underset{indpen.}{=} \prod \varphi_{X_i}(t) \underset{i.d.}{=} \varphi_X^n(t), \tag{1.8}$$

mientras que

$$\varphi_{\overline{X}}(t) = \varphi_{\frac{1}{n}\sum_{i=1}^{n} X_i}(t) = \varphi_{\sum_{i=1}^{n} X_i}(\frac{t}{n}) = \varphi_X^n(\frac{t}{n}), \tag{1.9}$$

basándonos en las propiedades de las funciones características.

Ejemplo 1.1.3 *Sea $(X_1, ..., X_n)$ una m.a. de $X \sim N(\mu, \sigma^2)$, con función característica $\varphi_X(t)$, ¿qué podemos decir sobre la distribución de la media muestral, $\overline{X}$?. Vemos que*

$$\varphi_X(t) = e^{it\mu - \frac{1}{2}t^2\sigma^2}$$

entonces teniendo en cuenta las fórmulas anteriores encontramos que:

$$\varphi_{\overline{X}}(t) = \varphi_X^n(\frac{t}{n}) = e^{in\frac{t}{n}\mu - \frac{n}{2}\left(\frac{t}{n}\right)^2\sigma^2} = e^{it\mu - \frac{1}{2}\frac{t}{n}^2\sigma^2},$$

donde observamos que $E\left[\overline{X}\right] = \mu$, y $var\left[\overline{X}\right] = \frac{\sigma^2}{n}$.

Volmemos otra vez sobre los **estadísticos de orden.**

Teorema 1.1.1 *Consideramos una muestra aleatoria* $(X_1, ..., X_n)$ *, de* X *continua con una función de densidad* f *y una función de distribución* F*, entonces:*

$$f_{X_{(j)}} = \frac{n!}{(j-1)!\,(n-j)!}\,[F(x)]^{j-1}\,f(x)\,[1-F(x)]^{n-j}\,. \tag{1.10}$$

Interpretamos esta fórmula de la siguiente manera;

$$[F(x)]^{j-1}\,f(x)\,[1-F(x)]^{n-j}: \quad \text{probabilidad de cada posibilidad}$$

$$\frac{n!}{(j-1)!\,(n-j)!}: \quad \text{número de posibilidades.}$$

y la podemos emplear de la siguiente forma.

Teorema 1.1.2 *Sea* $(X_1, ..., X_n)$ *, de* X *continua con una función de densidad* f *y una función de distribución* F*, entonces:*

$$f_{\left(X_{(i)},X_{(j)}\right)}(u,v) = \frac{n!}{(i-1)!\,(j-1-i)!\,(n-j)!}\,[F(u)]^{i-1}\,f(u)\,[F(v)-F(u)]^{j-1-i}\,f(v)\,[1-F(v)]^{n-j}\,. \tag{1.11}$$

Teorema 1.1.3 *Sea* $(X_1, ..., X_n)$ *, de* X *continua con una función de densidad* f *y una función de distribución* F*, entonces:*

$$f_{\left(X_{(1)},.....,X_{(n)}\right)} = n!\prod_{i=1}^{n} f(x_i)\,. \tag{1.12}$$

1.2. Inferencia paramétrica.

Sea $(X_1, ..., X_n)$ una m.a. de $X \sim f_\theta(X)$. Empezaremos repasando algunas de las propiedades fundamentales sobre las principales distribuciones que se emplearán más adelante y que se derivan de la normal.

Definición 1.2.1 χ_n^2. *Distribución* ***Chi-cuadrado*** *con n-grados de libertad. Consideramos* $(X_1, ..., X_n)$ *v.a.i.i.d. que siguen una* $N(0,1)$*. Definimos*

$$\chi_n^2 = \sum_{i=1}^{n} X_i^2. \tag{1.13}$$

Definición 1.2.2 ***t-Student****, con n-grados de libertad,* t_n*.*

Considermos $X \in N(0,1)$ *e* $Y \in \chi_n^2$ *de tal forma que* (X,Y) *son independientes, definimos*

$$t_n = \frac{X}{\sqrt{\frac{Y}{n}}} = \frac{N(0,1)}{\sqrt{\chi_n^2/n}} \tag{1.14}$$

El cálculo de la función de densidad es fácil ya que al tratarse de v.a.i. entonces $f_{XY} = f_X f_Y$

Definición 1.2.3 ***F Fisher-Snedecor***. *$F_{m,n}$ con m y n grados de libertad. Sea $X \in \chi^2_m$, e $Y \in \chi^2_n$ v.a.i. entonces definimos*

$$F_{m,n} = \frac{\frac{X}{m}}{\frac{Y}{n}} = \frac{Xn}{Ym} = \frac{\chi^2_m/m}{\chi^2_n/n} \tag{1.15}$$

Teorema 1.2.1 ***Fisher***. *Sea $(X_1, ..., X_n)$ una m.a. de $X \sim N(\mu, \sigma^2)$, entonces:*

1. $\bar{X}_n \in N(\mu, \frac{\sigma^2}{n})$,
2. $\bar{X}_n$ *y* S^2_n *son independientes,*
3. $$\frac{(n-1)\,S^2_n}{\sigma^2} = \frac{\sum (X_i - \mu)^2}{\sigma^2} \in \chi^2_{n-1}, \tag{1.16}$$

El segundo apartado del teorema es el resultado realmente importante.

1.3. Estadísticos suficientes.

Empezaremos con una idea intuitiva. Como siempre consideramos $(X_1, ..., X_n)$ una m.a. de $X \sim f_\theta(X)$. La muestra aleatoria $(X_1, ..., X_n)$ quedará reducida a un determinado estadístico T, que pasaremos a denominar estadístico suficiente. ¿Cuánta información perdemos al resumir la m.a. en T?. Llegamos a la conclusión de que si T es suficiente entoncesno hay pérdida de información.

Definición 1.3.1 *Estadístico suficiente. Sea $(X_1, ..., X_n)$ una m.a. de $X \sim f_\theta(X)$, Un estadístico T es suficiente para θ si $(X_1, ..., X_n \quad | \quad T = t)$ no depende de θ.*

Esta definición no ayuda demasiado a encontrar un estadístico además el cálculo de las distribuciones condicionadas puede resultar algo pesado, por eso consideramos el siguiente teorema

Teorema 1.3.1 ***de Factorización.*** *Consideramos $(X_1, ..., X_n)$ una m.a. de $X \sim f_\theta(X)$, entonces T será estadístico suficiente para θ sii $f_\theta\,(x_1, ..., x_n)$ puede factorizarse de la siguiente forma:*

$$f_\theta\,(x_1, ..., x_n) = g\,(t, \theta) \cdot h\,(x_1, ..., x_n)\,, \tag{1.17}$$

donde $t = T\,(x_1, ..., x_n)$.

Ejemplo 1.3.1 *Sea*

$$f_\theta = \frac{(\log\theta)\,\theta^x}{\theta - 1} I_{(0,1)}(x), \qquad \theta > 1,$$

entonces un estadístico sufiente puede ser

$$\prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \left(\frac{(\log\theta)\,\theta^{x_i}}{\theta - 1} I_{(0,1)}(x_i) \right) = \left(\frac{(\log\theta)}{\theta - 1} \right)^n \theta^{\sum x_i} \prod_{i=1}^n I_{(0,1)}(x_i)$$

por lo tanto

$$T = \sum x_i$$

será un estadístico suficiente en virtud del teorema de factorización.

Ejemplo 1.3.2 *Si*

$$f_\theta = \frac{2x}{\theta^2} I_{(0,\theta)}(x), \qquad \theta > 0,$$

entonces un estadístico sufiente puede ser

$$T = \left(\prod_{i=1}^n x_i, X_{(n)} \right)$$

ya que en virtud del teorema de facotrización tenemos que:

$$\begin{aligned} \prod_{i=1}^n f_\theta(x_i) &= \prod_{i=1}^n \left(\frac{2x}{\theta^2} I_{(0,\theta)}(x_i) \right) = \left(\frac{2}{\theta^2} \right)^n \prod_{i=1}^n (x_i) \prod_{i=1}^n I_{(0,\theta)}(x_i) = \\ &= \left(\frac{2}{\theta^2} \right)^n \prod_{i=1}^n (x_i)\, I_{\left(X_{(n)},\infty\right)}(\theta), \end{aligned}$$

por lo tanto

$$T = X_{(n)}$$

será un estadístico suficiente y $T = \left(\prod_{i=1}^n x_i, X_{(n)} \right)$ *también.*

1.4. Ejercicios.

Ejercicio 1.4.1 *Sea* $(X_1, ..., X_n)$ *una muestra sin reemplazamiento de una población finita*

$$\{x_1, ..., x_N\},$$

Probar que X_1 *y* X_2 *no son independientes, pero tienen la misma distribución.*

(b) *Sea* $(X_1, ..., X_{10})$ *una muestra sin reemplazamiento de la población finita:*

$$\{1, 2, ..., 1000\}$$

Calcular la probabilidad de que todas las observaciones sean mayores que 200*, primero de manera exacta y después de manera aproximada admitiendo independencia.*

Solución. Con respecto al primer apartado, vemos que las $(X_i)_{i=1}^n$ son v.a.i.d. pero no independientes. Lo usual es trabajar con v.a.i.i.d. Por lo tanto vemos que

$$P(X_1 = x) = \frac{1}{N},$$

$$P(X_2 = x) \underset{prob.total}{=} \sum P(X_1 = y)\, P(X_2 = x \mid X_1 = y)$$

y por lo tanto

$$\begin{aligned} P(X_2 = x) &= P(X_1 = x)\, P(X_2 = x \mid X_1 = x) + \sum_{y \neq x} P(X_1 = y)\, P(X_2 = x \mid X_1 = y) = \\ &= \frac{1}{N} \cdot 0 + \sum \frac{1}{N} \cdot \frac{1}{N-1} = (N-1) \frac{1}{N} \cdot \frac{1}{N-1} = \frac{1}{N} \end{aligned}$$

de esta forma vemos que siguen igual distribución.

Para ver la independencia i.e.

$$¿P(X_1 = x, X_2 = y) = P(X_1 = x) \cdot P(X_2 = y)?$$

así que

$$P(X_1 = x, X_2 = x) = 0,$$

sin embargo

$$P(X_1 = x) \cdot P_2(X_2 = x) = \frac{1}{N} \cdot \frac{1}{N}$$

por lo que no son independientes. ¿Es grave esta pérdida?,¿es importante?.

Sea $(X_1, ..., X_{10})$ una muestra sin reemplazamiento de la población finita:

$$\{1, 2, ..., 1000\}$$

queremos calcular

$$P(\text{todas las observaciones sean } > 200)$$

para ello definimos la v.a.

$$Y := \text{ número de observaciones mayores que 200 entre las 10.}$$

Cálculo exacto: al no ser independientes los sucesos entonces las $n-$Bernoullis no forman una Binomial sino una hipergeométrica

$$P(Y = 10) = \frac{\binom{800}{10}\binom{200}{0}}{\binom{1000}{10}} \simeq 0{,}10616,$$

mientras que el cálculo aproximado lo haremos considerando que los sucesos son independientes y por lo tanto siguen un Binomial $Bin(n, p)$, donde la probabilidad de éxito vendrá dada por

$$p = P\left(\acute{E}xito\right) = \frac{800}{1000} = 0{,}8$$

de esta forma llegamos al siguiente resultado:

$$P(Y=10) = Bin(10, 0{,}8) = \begin{pmatrix} 10 \\ 10 \end{pmatrix} (0{,}8)^n (0{,}2)^0 = (0{,}8)^n \simeq 0{,}107,$$

por lo que la moraleja del ejercio es que no hay mucha pérdida de información al considerar la independencia. ■

Ejercicio 1.4.2 *Supongamos que* $(X_1, ..., X_{n+1})$ *es una muestra aleatoria de una población* X *con* $E[X] = \mu$, *y* $V(X) = \sigma^2$, *sea*

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Hallar la esperanza y la varianza del estadístico

$$T = \sqrt{\frac{n}{n+1}} \left(X_{n+1} - \bar{X}_n \right).$$

Solución. Vemos que

$$\begin{aligned} E[T] &= E\left(\sqrt{\frac{n}{n+1}} \left(X_{n+1} - \bar{X}_n \right) \right) = 0 \\ &= \sqrt{\frac{n}{n+1}} E\left(\left(X_{n+1} - \bar{X}_n \right) \right) = \sqrt{\frac{n}{n+1}} \left(E\left(X_{n+1} \right) - E\left(\bar{X}_n \right) \right) = 0, \end{aligned}$$

ya que $E(X_{n+1}) = E\left(\bar{X}_n\right) = \mu$, mientras que

$$\begin{aligned} var[T] &= var\left(\sqrt{\frac{n}{n+1}} \left(X_{n+1} - \bar{X}_n \right) \right) = \frac{n}{n+1} var\left(X_{n+1} - \bar{X}_n \right) \\ &= \frac{n}{n+1} \left(var\left(X_{n+1} \right) + var\left(\bar{X}_n \right) - 2cov\left(X_{n+1}, \bar{X}_n \right) \right) \end{aligned}$$

ya que no sé de antemano si se tratade dos v.a. independientes, pero vemos que

$$\begin{aligned} (X_1, ..., X_{n+1}) &\longrightarrow \bar{X}_n, \\ X_{n+1} &\longrightarrow X_{n+1}, \end{aligned}$$

son independientes por lo $cov\left(X_{n+1}, \bar{X}_n\right) = 0$, quedando la anterior expresión reducida a:

$$var[T] = \frac{n}{n+1} \left(var\left(X_{n+1} \right) + var\left(\bar{X}_n \right) \right),$$

ahora es fácil comprobar que

$$\begin{aligned} var\left(X_{n+1} \right) &= \sigma^2, \\ var\left(\bar{X}_n \right) &= \frac{\sigma^2}{n} \end{aligned}$$

quedando por lo tanto

$$var[T] = \frac{n}{n+1} \left(\sigma^2 + \frac{\sigma^2}{n} \right) = \sigma^2,$$

tal y como queríamos hacer ver. ■

Ejercicio 1.4.3 *Sea $(X_1, ..., X_n)$ una muestra aleatoria de una $N(0,1)$. Probar:*

1. $X_1^2 \sim Gamma\left(a = \frac{1}{2}; p = \frac{1}{2}\right)$, *esta distribución es la* χ_1^2.
2. $\sum_{i=1}^n X_i^2 \sim Gamma\left(a = \frac{1}{2}; p = \frac{n}{2}\right)$, *que es la* χ_n^2.

Solución. Con respecto al primer apartado vemos que:

$$X_1^2 \sim Gamma\left(a = \frac{1}{2}; p = \frac{1}{2}\right)$$

donde

$$X_1 \in N(0,1) \qquad \Longrightarrow \qquad f_{X_1} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x_1^2}{2}\right)$$

definimos la nueva v.a. $Y = X_1^2$, viendose así que $X_1 = \pm\sqrt{Y}$. Teniendo en cuenta el jacobiano de esta transformación la función de distribución de la nueva v.a. será:

$$f_{Y=X_1^2} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y}{2}\right) \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y}{2}\right) \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} \exp\left(-\frac{y}{2}\right)$$

donde

$$|J| = \frac{1}{2\sqrt{y}}$$

por lo que

$$f_Y = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y}{2}\right) y^{-1/2},$$

y si tenemos en cuenta que

$$f_{\gamma(a,p)} = \frac{a^p}{\Gamma(p)} \exp\left(-ax\right) x^{p-1},$$

llegamos a la igualdad deseada puesto que $\left(a = \frac{1}{2}, p = \frac{1}{2}\right)$, $\Gamma(p) = \sqrt{\pi}$, y $a^p = \frac{1}{\sqrt{2}}$.

Con respecto al segundo apartado vemos que

$$\sum_{i=1}^n X_i^2 \sim Gamma\left(a = \frac{1}{2}, p = \frac{n}{2}\right)$$

operamos de forma parecida. Teniendo en cuenta las funciones características etc... vemos que al ser independientes

$$\sum_{i=1}^n X_i^2 = \prod \varphi_{X_i^2}$$

(entiéndase la expresión) de esta forma vemos que

$$\varphi_{\sum_{i=1}^n X_i^2} = \prod \varphi_{X_i^2} = \prod \left(1 - \frac{it}{a}\right)^{-p} = \left(1 - \frac{it}{1/2}\right)^{-\frac{n}{2}} \sim Gamma\left(a = \frac{1}{2}, p = \frac{n}{2}\right)$$

donde la función característica de la función Gamma es:

$$\varphi_{\gamma(a,p)} = \left(1 - \frac{it}{a}\right)^{-p}$$

tal y como queríamos hacer ver. ■

Ejercicio 1.4.4 *Sea $(X_1, ..., X_n)$ una muestra aleatoria de una población $U(0,1)$. Hallar la esperanza y la varianza de $X_{(j)}$.*

Solución. En primer lugar recordemos que

$$f_{X_{(j)}} = \frac{n!}{(j-1)!\,(n-j)!} \left[F(x)\right]^{j-1} f\left(x\right) \left[1 - F\left(x\right)\right]^{n-j}.$$

y que las propiedades de la función Beta son: función de densidad.

$$f_X(t) = \frac{1}{B\left(p,q\right)} x^{p-1} \left(1-x\right)^{q-1} I_{(0,1)}(x),$$

y su esperanza y varianza son:

$$E(X) = \frac{p}{p+q}, \qquad var(X) = \frac{pq}{\left(p+q\right)^2 \left(p+q+1\right)}$$

De esta forma vemos que necesitamos conocer la $F(x)$ y la $f\left(x\right)$ de la distribución uniforme siendo éstas:

$$f = 1, \qquad F = \int_{-\infty}^{x} 1dx = x, \qquad \forall x \in \left(0,1\right),$$

por lo que

$$\begin{aligned} f_{X_{(j)}} &= \frac{n!}{(j-1)!\,(n-j)!} \left[F(x)\right]^{j-1} f\left(x\right) \left[1 - F\left(x\right)\right]^{n-j} = \\ &= \frac{n!}{(j-1)!\,(n-j)!} \left[x\right]^{j-1} 1 \left[1 - x\right]^{n-j} \sim Beta\left(p,q\right) \end{aligned}$$

obteniéndose de forma inmediata la analogía con la función beta

$$\frac{1}{B\left(p,q\right)} = \frac{n!}{(j-1)!\,(n-j)!} = \frac{\Gamma\left(p+q\right)}{\Gamma\left(p\right)\Gamma\left(q\right)}$$

con $j = p$, $q = n - j + 1$. De esta forma

$$\begin{aligned} E\left[X_{(j)}\right] &= \int x f_{X_{(j)}} = \frac{p}{p+q} = \frac{j}{n+1}, \\ var\left[X_{(j)}\right] &= \frac{pq}{\left(p+q\right)^2 \left(p+q+1\right)} = \frac{j\left(n-j+1\right)}{\left(n+1\right)^2 \left(n+2\right)}, \end{aligned}$$

etc.... falta desarrollar todos estos cálculos pesadísimos. ■

Ejercicio 1.4.5 *Si $\bar{X}$ es la media de una muestra aleatoria $(X_1, ..., X_n)$ de una $N(\mu, \sigma^2 = 100)$, hallar n para que*

$$P\left(\mu - 5 < \bar{X} < \mu + 5\right) = 0{,}0954.$$

Solución. Vemos que

$$\bar{X} \in N\left(\mu, \frac{100}{n}\right) = N\left(\mu, \frac{10}{\sqrt{n}}\right),$$

por lo que

$$P\left(\mu - 5 < \bar{X} < \mu + 5\right) = P\left(\frac{\mu - 5 - \mu}{\frac{10}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{10}{\sqrt{n}}} < \frac{\mu + 5 - \mu}{\frac{10}{\sqrt{n}}}\right),$$

donde

$$Z = \frac{\bar{X} - \mu}{\sigma} = \frac{\bar{X} - \mu}{10/\sqrt{n}} \in N\left(0, 1\right),$$

así que la anterior expresión se reduce a:

$$P\left(-\frac{5\sqrt{n}}{10} < Z < \frac{5\sqrt{n}}{10}\right),$$

y por simetría se obtiene

$$P\left(Z > \frac{\sqrt{n}}{2}\right) = \frac{1 - 0{,}0954}{2} = 0{,}0230,$$

por lo que

$$\frac{\sqrt{n}}{2} \sim \underset{mirar\ tablas}{2...} \qquad \Longrightarrow \qquad n \simeq 16,$$

tal y como queríamos calcular. ■

Ejercicio 1.4.6 *Sean* $(X_1, ..., X_{25})$ *e* $(Y_1, ..., Y_{25})$ *dos muestras aleatorias de dos distribuciones independientes* $N(\mu = 0, \sigma^2 = 16)$ *y* $N(\mu = 1, \sigma^2 = 9)$ *respectivamente. Calcular* $P\left(\bar{X} > \bar{Y}\right)$.

Solución. Vemos que

$$\begin{array}{rcl} (X_1, ..., X_{25}) & \longrightarrow & \bar{X} \sim N(\mu = 0, \sigma^2 = 16), \\ (Y_1, ..., Y_{25}) & \longrightarrow & \bar{Y} \sim N(\mu = 1, \sigma^2 = 9), \end{array}$$

y queremos calcular $P\left(\bar{X} > \bar{Y}\right)$. Lo que hacemos en este caso es lo siguiente:

$$P\left(\bar{X} > \bar{Y}\right) = P\left(\bar{X} - \bar{Y} > 0\right) = P\left(Z > 1\right) = 0{,}1587$$

donde la nueva v.a. $\left(\bar{X} - \bar{Y}\right) \sim N(\mu = -1, \sigma^2 = 1)$ y la normalizamos a $Z \in N\left(0, 1\right)$, obteniendo el resultado requerido. Ver las propiedades de la distribución normal en el último capítulo. ■

<u>ESTADÍSTICOS SUFICIENTES</u>

Ejercicio 1.4.7 *Encontrar un estadístico suficiente en cada uno de los siguientes casos basado en una muestra aleatoria de tamaño* n *de:*

1. $X \sim Beta(\alpha, \beta)$.

2. $X \sim Gamma(p, a)$.

3.
$$f_\theta(x) = \begin{cases} e^{-x+\theta} & x \in (\theta, \infty), \\ 0 & resto, \end{cases}$$

4.
$$f_{\mu,\sigma} = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(\frac{1}{2\sigma^2}\log(x-\mu)^2\right),$$

con $x > 0$.

Solución. En todos los casos se trata de una aplicación sistemática del teorema de factorización.

1. $X \sim Beta(\alpha, \beta)$

$$\begin{aligned} \theta &= (\alpha, \beta) \\ f_\theta &= \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \end{aligned}$$

por lo que

$$f_\theta(x_1, ..., x_n) = \prod f_\theta(x_i) = \left(\frac{1}{B(\alpha,\beta)}\right)^n \prod x_i^{\alpha-1}(1-x_i)^{\beta-1},$$

de esta forma

$$f_\theta(x_1, ..., x_n) = K\left(\prod x_i^{\alpha-1}\right)\left(\prod (1-x_i)^{\beta-1}\right) = g(t,\theta)h(x_1, ..., x_n)$$

donde elejimos $h(x_1, ..., x_n) = 1$. Definimos por lo tanto el estadístico

$$T = \left(\prod x_i^{\alpha-1}, \prod (1-x_i)^{\beta-1}\right)$$

es el estadístico suficiente buscado para (α, β). No obstante debemos resaltar que no se trata del único estadístico suficiente, la muestra en sí misma es un estadístico suficiente o

$$T = \left(\sum \log x_i, \sum \log(1-x_i)\right)$$

también es suficiciente ya que

$$\prod x_i = e^{\log\left(\prod x_i\right)} = e^{\sum \log x_i},$$

luego reescribiendo adecuadamente todas estas transformaciones biyectivas obtenemos lo mismo, la moraleja está en que dado cualquier estadístico suficiente (e.s.), entonces cualquier transformación biyectiva nos da otro e.s..

2. $X\sim Gamma(p,a)$, donde
$$f_{\gamma(a,p)}=\frac{a^{p}}{\Gamma(p)}\exp\left(-ax\right)x^{p-1}$$
siguiendo el anterior esquema vemos que
$$f_{\theta}\left(x_{1},....,x_{n}\right)=\prod f_{\theta}\left(x_{i}\right)=\left(\frac{a^{p}}{\Gamma(p)}\right)^{n}\left(\prod\exp\left(-ax_{i}\right)\prod x_{i}^{p-1}\right)$$
de esta forma llegamos a la conclusión de que
$$T=\left(\sum x_{i},\prod x_{i}\right)$$
es un estadístico suficiente para $\theta=(p,a)$.

Observación 1.4.1 *En realidad se han omitido los detalles más sangrientos, ya que hasta que se llega al estadístico T, la verdad, hay que hacer unas cuantas cuentas:*
$$\begin{aligned}\prod\exp\left(-ax_{i}\right) &= \exp\left(-na\sum x_{i}\right),\\ \prod x_{i}^{p-1} &= \underset{manipular}{....}.\approx f\left(n,p\right)\prod x_{i}\end{aligned}$$

3. $X\sim e^{-x+\theta}$, con $x>\theta$. Intentamos la misma argumentación i.e.
$$f_{\theta}\left(x_{1},....,x_{n}\right)=\prod f_{\theta}\left(x_{i}\right)=\left[e^{-x_{1}+\theta}......e^{-x_{n}+\theta}\right]=e^{n\theta}e^{-\sum x_{i}}$$
llegando a la conclusión de que
$$f_{\theta}\left(x_{1},....,x_{n}\right)=g(t,\theta)h(x_{i})$$
por lo que podríamos tomar
$$g(t,\theta)=e^{n\theta},\qquad h(x_{i})=e^{-\sum x_{i}}$$
de esta forma
$$T=n\theta$$
será el estadístico suficiente. NO. Hemos operado mal intencionadamente. Aquí hay que tener muy en cuenta que $x>\theta$, por lo que empezamos reescribiendo la f_{θ} como
$$f_{\theta}=e^{-x+\theta}I_{(\theta,\infty)}(x)$$
y volvemos a rehacer el ejercicio con esta nueva consideración
$$f_{\theta}\left(x_{1},....,x_{n}\right)=\prod f_{\theta}\left(x_{i}\right)=\prod e^{-x_{i}+\theta}I_{(\theta,\infty)}(x_{i})=e^{n\theta}\prod I_{(\theta,\infty)}(x_{i})e^{-\sum x_{i}}$$
donde
$$g(t,\theta)=e^{n\theta}I_{(\theta,\infty)}(X_{(1)}),\qquad h(x_{i})=e^{-\sum x_{i}}$$
observar que $\prod I_{(\theta,\infty)}(x_{i})=I_{(\theta,\infty)}(X_{(1)})$, donde $X_{(1)}=\min\left(x_{1},....,x_{n}\right)$ por lo que podemos definir el e.s. T como
$$T=\left(X_{(1)}\right),$$
sorprendente, ¿no?.

4. En este caso
$$f_{\mu,\sigma}=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left(\frac{1}{2\sigma^2}\log\left(x-\mu\right)^2\right)$$
i.e. el modelo log-normal. En este caso el estadístico sufiente será
$$T=\left(\sum\log x_i,\sum\log x_i^2\right).$$
Para llegar a esta conclusión podemos considerar $(X_1,....,X_n)$ m.a. de X una $N(\mu,\sigma^2)$, por lo que
$$T=\left(\sum x_i,\sum x_i^2\right)$$
será un e.s. para este caso, aplicando el teorema de Fisher, entonces $T=\left(\bar{X},S^2\right)$ es un e.s. (falta probarlo).

■

Ejercicio 1.4.8 *Sea $(X_1,....,X_n)$ una muestra aleatoria de $X\sim U(0,\theta)$, $(\theta>0)$. Estudiar la suficiencia de $T=X_{(n)}$*

Solución. Vemos que
$$f_U=\frac{1}{\theta},\qquad\Longrightarrow\qquad f_\theta\left(x_1,....,x_n\right)=\frac{1}{\theta}I_{(0,\infty)}(x_i)I_{(-\infty,\theta)}(x_i)$$
aquí es donde está toda la miga del problema, por lo que
$$f_\theta=\left(\frac{1}{\theta}\right)^n\prod I_{(0,\infty)}(x_i)\prod I_{(-\infty,\theta)}(x_i)$$
y al igual que antes observamos que
$$\begin{aligned}\prod I_{(0,\infty)}(x_i)&=I_{(0,\infty)}(X_{(1)}),\\ \prod I_{(-\infty,\theta)}(x_i)&=I_{(-\infty,\theta)}(X_{(n)}),\end{aligned}$$
de esta forma llegamos a la conclusión de que
$$f_\theta\left(x_1,....,x_n\right)=g(t,\theta)h(x_i)$$
por lo que podríamos tomar
$$g(t,\theta)=\left(\frac{1}{\theta}\right)^nI_{(-\infty,\theta)}(X_{(n)}),\qquad h(x_i)=I_{(0,\infty)}(X_{(1)})$$
obteniendo así T
$$T=X_{(n)}$$
será el estadístico suficiente. Lo principal del problema está en ver
$$f_\theta\left(x_1,....,x_n\right)=\frac{1}{\theta}I_{(0,\infty)}(x_i)I_{(-\infty,\theta)}(x_i),$$
lo cual no siempre es fácil. ■

Capítulo 2

Estimación puntual

2.1. Introducción

Sea $(X_i)_{i=1}^n$ una muestra aleatoria (m.a.) donde $X \sim f_\theta(x)$, $\theta \in \Theta$, el parámetro θ es lo que queremos estimar. El objetivo es por lo tanto intentar decir algo sensato sobre θ a partir de los datos que tenemos i.e. de la m.a. $(X_i)_{i=1}^n$. Para ello podemos seguir tácticas:

1. Estimación puntual (el objeto de este capítulo).
2. Los intervalos de confianza.
3. Contraste de hipótesis.

Los objetivos de la estimación puntual son los siguientes: Estimar el parámetro θ (ó $\tau(\theta)$) mediante un único valor (un punto) a apartir de la muestra $(X_i)_{i=1}^n$.

Definición 2.1.1 *Sea* $(X_i)_{i=1}^n$ *m.a. donde* $X \sim f_\theta(x)$, $\theta \in \Theta$, *un estimador puntual para* θ *es simplemente un estadístico* $T = T(x_1, ..., x_n)$ *cuyo objetivo es estimar lo mejor posible* θ *(ó* $\tau(\theta)$*).*

Los métodos de construcción son los siguientes:

1. Método de los momentos,
2. Método de máxima verosimilitud,
3. Métodos bayesianos.

Tal y como veremos el metodo de máxima verosimilitud es el método por excelencia.

2.2. Método de los momentos

Definición 2.2.1 *Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f_\theta(x)$, $\theta \in \Theta$, un estimador puntual para θ por el método de los momentos (MM) y denotado por $\tilde{\theta}$, es el valor que se obtiene al resolver el siguiente sistema de ecuaciones:*

$$\begin{aligned} E_\theta\left[X\right] &= \frac{1}{n}\sum X_i, \\ &\vdots \\ E_\theta\left[X^k\right] &= \frac{1}{n}\sum X_i^k, \end{aligned} \tag{2.1}$$

Ejemplo 2.2.1 *Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f_\theta(x)$, $\theta \in \Theta$, donde $X \sim$ Bernoulli(p), lo que queremos estimar es p así que*

$$E_p\left[X\right] = \frac{1}{n}\sum X_i,$$

por lo tanto

$$\tilde{p} = \frac{1}{n}\sum x_i.$$

Ver la sección de ejercicios para aclarar este método con ejemplos más complicados.

2.3. Método de máxima verosimilitud

Definición 2.3.1 *Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f_\theta(x)$, $\theta \in \Theta$. Definimos para cada muestra fija $(X_i)_{i=1}^n$ la función de verosimilitud $L(\theta) =$ probabilidad o densidad que los diferentes valores de θ dan a $(X_i)_{i=1}^n = f_\theta(x_i)_{i=1}^n = \prod_{i=1}^n f_\theta(x_i)$. Por lo tanto*

$$L(\theta) = \prod_{i=1}^n f_\theta(x_i).$$

El estimador de máxima verosimilitud $\hat{\theta}$ es el valor del parámetro θ que maximiza $L(\theta)$.

Observación 2.3.1 *En la mayoría de los casos para la obtención de $\hat{\theta}$ se maximiza la función $\log L(\theta)$ en vez de $L(\theta)$*

Ejemplo 2.3.1 *Se trata del mismo ejemplo que el del MM i.e. sea* $(X_i)_{i=1}^n$ *m.a. donde* $X \sim f_\theta(x)$, $\theta \in \Theta$, *donde* $X \sim Bernoulli(p)$, *i.e.*

$$X \sim Bernoulli(p) = p^x (1-p)^{1-x},$$

lo que queremos estimar es p *así que*

$$L(p) = P_p(x_1, .., x_n) \overset{indp.}{=} \prod_{i=1}^n P_p(x_i) = p^{\sum x_i} (1-p)^{n-\sum x_i}$$

tomamos logaritmos

$$\log(L(p)) = \sum x_i \log(p) + \left(n - \sum x_i\right) \log(1-p)$$

y pasamos a calcular el máximo i.e.

$$\begin{aligned} \partial_p \log(L(p)) &= 0 \\ \sum x_i \frac{1}{p} - \left(n - \sum x_i\right) \frac{1}{1-p} &= 0 \end{aligned}$$

despejamos p, *encontrándose que*

$$\hat{p} = \frac{1}{n} \sum x_i.$$

Vemos que en tos ejemplos hemos obtenido que $\hat{p} = \tilde{p}$, *pero esto no siempre pasa.*

2.4. Métodos bayesianos.

2.4.1. Introducción a los métodos bayesianos.

Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f_\theta(x)$, $\theta \in \Theta$. En primer lugar advertir que en este caso cambia la notación y se emplea

$$f_\theta(x) = f(x \mid \theta), \tag{2.2}$$

considerando a θ como una v.a.

1. Disponemos de información muestral

$$f_\theta(x_1, ..., x_n) = f(x_1, ..., x_n \mid \theta) = \prod_{i=1}^n f(x_i \mid \theta). \tag{2.3}$$

2. Nuevo. Disponemos de información previa sobre el parámetro θ, que modelizamos mediante una densidad a priori (no proviene de la muestra) $\pi(\theta)$.

3. Obtenemos la densidad conjunta

$$f(x_i \mid \theta)\, \pi(\theta). \tag{2.4}$$

4. Obtenemos la densidad o función de masa de θ condicionada por $(X_i)_{i=1}^n$, densidad a posteriori,
$$\pi(\theta \mid x_1, ..., x_n) = \frac{f(x_1, ..., x_n \mid \theta)\,\pi(\theta)}{\int_\theta f(x_1, ..., x_n \mid \theta)\,d\theta} \propto f(x_i \mid \theta)\,\pi(\theta). \tag{2.5}$$
esta expresión no es más que la expresión de Bayes para el cáculo de la probabilidad condicionada.

Los métodos bayesianos obtienen todas sus conclusiones sobre θ a partir de $\pi(\theta \mid x_i)$ (densidad condicionada) que recibe el nombre de densidad a posteriori.

El problema en la práctica está en la elección de $\pi(\theta)$. A menudo se utilizan familias conjungadas de densidades a priori facilitándose así los cálculos.

Definición 2.4.1 ***Familia conjungada.*** *Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f(x \mid \theta)$, $\theta \in \Theta$. Una familia $\mathcal{P} = \{\pi(\theta)\}$ de densidades a priori es conjungada para muestras de $X \sim f(x \mid \theta)$ cuando $\pi(\theta) \in \mathcal{P}$, entonces*
$$\pi(\theta \mid x_1, ..., x_n) \propto f(x_i \mid \theta)\,\pi(\theta) \in \mathcal{P}. \tag{2.6}$$

La idea en la práctica es que cualquier técnica estadísitica basada en métodos bayesianos obtendrá conclusiones a partir de $\pi(\theta \mid x_i)$.

2.4.2. Estimadores puntuales bayesianos.

Un estimador puntual bayesiano será una medida de centralización de la densidad a posteriori. Los más habituales son:

1. Media a posteriori
$$\int_\Theta \theta\pi(\theta \mid x_1, ..., x_n)\,d\theta, \tag{2.7}$$
2. Mediana a posteriori
$$\int_{-\infty}^{M} \theta\pi(\theta \mid x_1, ..., x_n)\,d\theta = \frac{1}{2}. \tag{2.8}$$

Ejemplo 2.4.1 *Estimación de $\theta = p = P(E)$, probabilidad de éxito. Disponemos de una m.a. $(X_i)_{i=1}^n$ donde*
$$X = \begin{cases} 1 & P(E) = \theta \\ 0 & P(F) = 1-\theta \end{cases}$$
por lo que $X \sim Bernoulli(\theta)$, tal que $\theta \in (0,1)$, así que sabemos que $X \sim f(x \mid \theta) = \theta^x (1-\theta)^{1-x}$. Consideramos la siguiente densidad a priori
$$\pi(\theta) \sim Beta(p = a, q = b)$$

por lo tanto la densidad a posteriori será

$$\begin{aligned} f\left(x_1,...,x_n \mid \theta\right) &= \prod_{i=1}^{n} f(x_i \mid \theta) = \theta^{\sum x_i}\left(1-\theta\right)^{n-\sum x_i} \\ \pi\left(\theta\right) &= \frac{\theta^{a-1}\left(1-\theta\right)^{b-1}}{B(a,b)} \end{aligned}$$

entonces

$$\begin{aligned} \pi\left(\theta \mid x_1,...,x_n\right) &\propto f\left(x_i \mid \theta\right)\pi\left(\theta\right) \\ &\propto \left(\theta^{\sum x_i}\left(1-\theta\right)^{n-\sum x_i}\right)\left(\frac{\theta^{a-1}\left(1-\theta\right)^{b-1}}{B(a,b)}\right) \\ &\propto \theta^{a+\sum x_i-1}\left(1-\theta\right)^{n+b-\sum x_i-1} \sim Beta\left(p,q\right) \end{aligned}$$

así llegamos a la conclusión de que

$$p = a + \sum x_i, \qquad q = n + b - \sum x_i.$$

Entonces la media a posteriori será

$$\int_{\Theta} \theta\pi\left(\theta \mid x_1,...,x_n\right) d\theta = \int_0^1 \theta Beta\left(p,q\right) d\theta$$

pero para hecharse esta cuenta uno debe considerar las constantes que hemos ido despreciando. Lo mejor en estos casos es fijarse en los resultados ya conocidos, por lo que

$$E\left[Beta\left(p,q\right)\right] = \frac{p}{p+q} = \frac{a+\sum x_i}{a+b+n}.$$

Vemos que la densidad a priori era una Beta y que la densidad a posteriori ha salido otra Beta i.e. hemos tomado una familia conjugada

$$\mathcal{P} = \left\{Beta\left(p,q\right),\ p,q > 0\right\}$$

si $p = q = 1 \Longrightarrow Beta\left(1,1\right) \sim U\left(\left[0,1\right]\right),$ *la uniforme sería el caso en el que no tenemos ni idea sobre* θ. *Si* $P(E) = \theta = \frac{1}{2}$ *(o se aproxima mucho a este valor) entonces tomamos* p,q *tales que* $1/2$ *sea el máximo para la función* $Beta\left(p,q\right)$.

2.5. Propiedades deseables de los estimadores.

2.5.1. Error cuadrático medio. Estimadores insesgados.

Definición 2.5.1 *El **error cuadrático medio** de un estimador* T *para estimar* $\tau\left(\theta\right)$ *es:*

$$ECM = E_{\theta}\left[\left(T-\tau\left(\theta\right)\right)^2\right] = \int f_{\theta}(x_i)\left(T-\tau\left(\theta\right)\right)^2 dx^n. \tag{2.9}$$

Si desarrollamos el formulón anterior entonces vemos que

$$ECM = E_{\theta}\left[(T - \tau(\theta))^2 \right] = var(T) + (E_{\theta}[T] - \tau(\theta))^2, \tag{2.10}$$

donde se define el sesgo como

$$sesgo := E_{\theta}[T] - \tau(\theta). \tag{2.11}$$

Definición 2.5.2 *Un* ***estimaodor*** *es* ***insesgado*** *o centrado para estimar una función del parámetro si $E_{\theta}[T] = \tau(\theta)$, i.e. el* ***sesgo,*** *$(E_{\theta}[T] - \tau(\theta) = 0)$ es cero i.e.*

$$ECM = E_{\theta}\left[(T - \tau(\theta))^2 \right] = var(T) + (E_{\theta}[T] - \tau(\theta))^2 = var(T). \tag{2.12}$$

Ejemplo 2.5.1 *Sea $(X_i)_{i=1}^{n}$ m.a. donde $X \sim N(\mu, \sigma^2)$, donde $\theta = (\mu, \sigma^2)$, entonces:*

1. *$T_1 = \bar{X}$, es insesgado para estimar $\mu = \tau(\theta)$, ya que*
$$E\left[\bar{X}\right] = \mu = \tau(\theta)$$
2. *$T_2 = S^2$, es insesgado para estimar $\sigma^2 = \tau(\theta)$, ya que*
$$E\left[S^2\right] = \sigma^2 = \tau(\theta)$$
recordar que $S^2 = \frac{1}{n-1}\sum \left(x_i - \bar{X}\right)^2$
3. *$T_3 = var(X)$, es sesgado para estimar $\sigma^2 = \tau(\theta)$, ya que*
$$E[var(X)] \neq \sigma^2 = \tau(\theta)$$
por lo tanto se trataría de un estimador sesgado.

Teorema 2.5.1 ***Cota de Frechet-Cramer-Rao****. Sea $(X_i)_{i=1}^{n}$ m.a. donde $X \sim f_{\theta}(x)$, $\theta \in \Theta$. para cualquier estimador T insesgado para $\tau(\theta)$ se verifica:*

$$var(T) \geq \frac{\left(\frac{d\tau(\theta)}{d\theta}\right)^2}{E_{\theta}\left[\left(\frac{d}{d\theta}\log f_{\theta}(x_1, ..., x_n)\right)^2\right]} = \frac{\left(\frac{d\tau(\theta)}{d\theta}\right)^2}{nE_{\theta}\left[\left(\frac{d}{d\theta}\log f_{\theta}(x)\right)^2\right]}. \tag{2.13}$$

Lema 2.5.1

$$E_{\theta}\left[\left(\frac{d}{d\theta}\log f_{\theta}(x)\right)^2\right] = -nE_{\theta}\left[\frac{d^2}{d\theta^2}\log f_{\theta}(x)\right] \tag{2.14}$$

por lo tanto

$$var(T) \geq \frac{\left(\frac{d\tau(\theta)}{d\theta}\right)^2}{-nE_{\theta}\left[\frac{d^2}{d\theta^2}\log f_{\theta}(x)\right]} \tag{2.15}$$

2.5.2. Estimadores eficientes.

Definición 2.5.3 *Un **estimador** T es **eficiente** para estimar un parámetro $\tau(\theta)$ si:*

1. *Es insesgado,*
2. *Su varianza alcanza la cota de FCR*

En consecuencia si es eficiente tiene mínima varianza. Al denominador de la cota de FCR (2.15) recibe el nombre de información de Fisher.

Definición 2.5.4 ***Información de Fisher**.*

$$IF = E_{\theta}\left[\left(\frac{d}{d\theta}\log f_{\theta}(x_i)\right)^2\right] \tag{2.16}$$

<u>Obtención de un estimador eficiente.</u>

Si T es eficiente su varianza alcanza la cota de FCR por lo que

$$\rho^2\left(T, \frac{d}{d\theta}\log f_{\theta}(x_i)\right) = 1,$$

el coeficiente de correlación vale uno, de este modo "casi seguro"

$$y - \bar{y} = \frac{cov(x,y)}{var(x)}(x - \bar{x})$$

obtenemos la recta de regresión que traducimos en este caso como

$$T - E_{\theta}[T] = \frac{cov(T, \frac{d}{d\theta}\log f_{\theta}(x_i))}{var(\frac{d}{d\theta}\log f_{\theta}(x_i))}\left(\frac{d}{d\theta}\log f_{\theta}(x_i) - E_{\theta}\left[\frac{d}{d\theta}\log f_{\theta}(x_i)\right]\right)$$

de donde sabemos que $E_{\theta}\left[\frac{d}{d\theta}\log f_{\theta}(x_i)\right] = 0$, desarrollando el formulón obtenemos el siguiente resultado:

$$T = \tau(\theta) + \frac{\tau'(\theta)}{-nE_{\theta}\left[\frac{d^2}{d\theta^2}\log f_{\theta}(x)\right]}(\log f_{\theta}(x_i))', \tag{2.17}$$

i.e.

$$T = \tau + \frac{\tau'}{IF}(\log f_{\theta})',$$

observándose que

$$\begin{aligned} E_{\theta}[T] &= \tau(\theta), \\ cov(T, \frac{d}{d\theta}\log f_{\theta}(x_i)) &= \frac{d\tau(\theta)}{d\theta} = \tau'(\theta), \\ var(\frac{d}{d\theta}\log f_{\theta}(x_i)) &= -nE_{\theta}\left[\frac{d^2}{d\theta^2}\log f_{\theta}(x)\right], \end{aligned}$$

así que si T es eficiente entonces debe tener la pinta de (2.17). *Si T no depende de θ, entonces será eficiente.*

2.5.3. Estimadores consistentes.

Sea $(X_i)_{i=1}^n$ m.a. donde $X \sim f_\theta(x), \ \ \theta \in \Theta$. Queremos estimar $\tau(\theta)$, Buscamos estimadores que se aproximen todo lo posible a lo que queremos estimar a medida que el número de observaciones n crece i.e. $T(X_i) \underset{L}{\simeq} \tau(\theta)$ i.e. hay convergencia en distribución (en Ley). Buscamos estimadores que funcionen bien asintóticamente.

Definición 2.5.5 *Decimos que un estimador T_n es consistente para estimar $\tau(\theta)$ si el estimador converge en probabilidad a $\tau(\theta)$ i.e.*

$$T_n \xrightarrow{P} \tau(\theta). \tag{2.18}$$

Veamos las distintas formas de abordar el estudio de la consistencia.

1. Supongamos que la función de distribución F_{T_n} es sencilla de obtener. Estudiamos entonces si
$$F_{T_n} \xrightarrow{L} F(y) = \begin{cases} 0 & y < \tau(\theta) \\ 1 & y \geq \tau(\theta) \end{cases} \tag{2.19}$$
entonces $T_n \xrightarrow{P} \tau(\theta)$ y por lo tanto se trata de un estimador consistente para $\tau(\theta)$.

   Esta táctica suele ser útil cuando estamos trabajando con máximos y mínimos

2. Supongamos que $\varphi_{T_n}(t)$ son fáciles de obtener y estudiamos la convergencia
$$\lim_{n\to\infty} \varphi_{T_n}(t) = e^{it\tau(\theta)} \tag{2.20}$$
entonces $T_n \xrightarrow{P} \tau(\theta)$ y por lo tanto se trata de un estimador consistente para $\tau(\theta)$.

   Esta táctica suele ser útil cuando estamos trabajando con sumas y medias muestrales.

Teorema 2.5.2 *Sea T_n un estimador tal que*

1. $\lim_{n\to\infty} var_\theta(T_n) = 0$,

2. $\lim_{n\to\infty} E_\theta[T_n] = \tau(\theta)$,

entonces T_n es consistente para estimar $\tau(\theta)$.

Teorema 2.5.3 *Bajo ciertas condiciones de regularidad el estimador de máxima verosimilitud es consistente.*

2.6. Ejercicios

Ejercicio 2.6.1 *Dada una m.a. de tamaño n de una v.a. X calcular el estimador de máxima verosimilitud y el del los momentos en los siguentes casos:*

1. $X \sim Poisson(\lambda)$,
2. $X \sim Exponencial(\lambda)$,
3. $X \sim N(\mu, \sigma^2)$, σ *conocido,*
4. $X \sim N(\mu, \sigma^2)$, μ *conocido,*
5. $X \sim N(\mu, \sigma^2)$.

Solución. Siguiendo el mismo esquema que el propuesto en los ejemplos de teoría vemos que:

1. $X \sim Poisson(\lambda)$, Veremos los dos métodos

   MM Sabemos que $X \sim Poisson(\lambda), \Longrightarrow E[X] = \lambda$
   $$E[X] = \lambda = \frac{1}{n}\sum_{i=1}^{n} x_i$$
   por lo que
   $$\tilde{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$
   MMV Tomamos como función de verosimilitud
   $$L(\lambda) = \prod P(x_i) = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{(x_1!)\ldots(x_n!)}$$
   tomando logaritmos entonces:
   $$\begin{aligned} \log(L(\lambda)) &= \log\left(\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} x_i}}{(x_1!)\ldots(x_n!)}\right) = \\ &= -n\lambda + \left(\sum_{i=1}^{n} x_i\right)\log\lambda - \sum_{i=1}^{n}\log((x_i)!) \end{aligned}$$
   y por lo tanto
   $$\begin{aligned} \partial_\lambda \log(L(\lambda)) &= 0, \\ -n + \left(\sum_{i=1}^{n} x_i\right)\frac{1}{\lambda} &= 0 \end{aligned}$$
   despejando se llega con facilidad a:
   $$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
   como era de esperar.

2. $X \sim Exponencial(\lambda)$, por lo que
$$f(x) = \lambda e^{-\lambda x},$$

MM Sabemos que $X \sim \exp(\lambda), \Longrightarrow E[X] = \lambda^{-1}$
$$E[X] = \lambda^{-1} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
por lo que
$$\tilde{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}.$$

MMV Tomamos como función de verosimilitud
$$L(\lambda) = \prod f(x_i) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$
tomando logaritmos entonces:
$$\log\left(L(\lambda)\right) = \log\left(\lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}\right) = n\log\lambda - \lambda\left(\sum_{i=1}^{n} x_i\right)$$
y por lo tanto
$$\partial_\lambda \log\left(L(\lambda)\right) = \frac{n}{\lambda} + \left(\sum_{i=1}^{n} x_i\right) = 0$$
despejando se llega con facilidad a:
$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}.$$
como era de esperar.

3. $X \sim N(\mu, \sigma^2)$, σ conocido,

MM Sabemos que $X \sim N(\mu, \sigma^2), \Longrightarrow E[X] = \mu$
$$E[X] = \mu = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \Longrightarrow \qquad \tilde{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

MMV Tomamos como función de verosimilitud
$$L(\mu) = \prod f(x_i) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right),$$
tomando logaritmos entonces:
$$\begin{aligned}\log\left(L(\mu)\right) &= \log\left(\frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\right) = \\ &= -n\log\sigma - n\log\left(\sqrt{2\pi}\right) - \left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)\end{aligned}$$

y por lo tanto

$$\partial_\mu \log\left(L(\mu)\right) = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i - n\mu = 0$$

despejando se llega con facilidad a:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

como era de esperar.

4. $X \sim N(\mu, \sigma^2)$, μ conocido,

MM Sabemos que $X \sim N(\mu, \sigma^2) \Longrightarrow E\left[X\right] = \mu$

$$E\left[X\right] = \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

no sirve en este caso pues este dato es conocido, por lo que tendremos que recurrir al momento de orden 2 i.e.

$$E\left[X^2\right] = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

de donde se obtiene que (recordar la definición de varianza, $var(X) = E\left[X^2\right] - E\left[X\right]^2$)

$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

despejando

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \mu^2$$

pero este resultado nos lleva a obtener un absurdo pues puede ser negativa esta magnitud !¡ Por ejemplo tomando $\mu = 3, x_i = (1, 2, 4)$.

MMV Tomamos como función de verosimilitud

$$L(\sigma) = \prod f(x_i) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

tomando logaritmos entonces:

$$\log\left(L(\sigma)\right) = -n\log\sigma - n\log\left(\sqrt{2\pi}\right) - \left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

y por lo tanto

$$\partial_\sigma \log\left(L(\sigma)\right) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0,$$

despejando se llega con facilidad a:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 .$$

como era de esperar. Aquí se ve que no siempre $\tilde{\sigma}^2 = \hat{\sigma}^2$, y que siempre el mejor etimador será el de MV i.e. $\hat{\sigma}^2$.

5. $X \sim N(\mu, \sigma^2)$,

MM Sabemos que $X \sim N(\mu, \sigma^2)$

$$\begin{aligned} E[X] &= \mu = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \Longrightarrow \qquad \tilde{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i, \\ E\left[X^2\right] &= \sigma^2+\mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \qquad \Longrightarrow \qquad \tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 . \end{aligned}$$

MMV Tomamos como función de verosimilitud

$$L(\mu,\sigma) = \prod f(x_i) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

tomando logaritmos entonces:

$$\log\left(L(\mu,\sigma)\right) = -n\log\sigma - n\log\left(\sqrt{2\pi}\right) - \left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

y por lo tanto

$$\begin{aligned} \partial_\mu \log\left(L(\mu,\sigma)\right) &= \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i - n\mu = 0, \\ \partial_\sigma \log\left(L(\mu,\sigma)\right) &= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i-\mu)^2 = 0, \end{aligned}$$

despejando se llega con facilidad a:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 .$$

como era de esperar.

■

Ejercicio 2.6.2 *Para cada uno de los siguientes casos encontrar la familia conjugada natural y hallar la distribución a posteriori.*

1. *$(X_i)_{i=1}^n$ es una m.a. de $X \sim Poisson(\theta)$,*
2. *$(X_i)_{i=1}^n$ es una m.a. de $X \sim Gamma\,(p=1, a=\theta)$*
3. *$(X_i)_{i=1}^n$ es una m.a. de $X \sim N\left(\theta, var = \frac{1}{r}\right)$, donde r es conocido.*

Solución. Seguimos el guión expuesto en el ejemplo de la teoría

1. $(X_i)_{i=1}^n$ es una m.a. de $X \sim Poisson(\theta)$,

$$f_\theta(x) = P(x \mid \theta) = \frac{e^{-\theta}\theta^x}{x!},\ x = 0, 1$$

familia de muestras para una Pisson

$$P(x \mid \theta) = \prod P(x_i \mid \theta) \propto e^{-n\theta}\theta^{\sum x_i}$$

$P(x \mid \theta)$ es el núcleo de una densidad tipo "Gamma", entonces tomamos como posible familia conjugada

$$\mathcal{P} = \{\pi(\theta) \in Gamma\,(p, a)\}$$

recordamos que

$$f \sim \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}$$

por lo que

$$\pi(\theta) \sim \frac{a^\theta}{\Gamma(\theta)} e^{-a\theta}\theta^{p-1}$$

así que

$$\pi(\theta \mid x_1, ..., x_n) \propto P(x_i \mid \theta)\,\pi(\theta) \propto e^{-n\theta}\theta^{\sum x_i}\frac{a^\theta}{\Gamma(\theta)} e^{-a\theta}\theta^{p-1} \propto e^{-(n+a)\theta}\theta^{p+\sum x_i - 1}$$

para $\theta > 0$, por lo tanto

$$\pi(\theta \mid x_1, ..., x_n) \sim Gamma\left(p = \alpha + \sum x_i, a = \beta + n\right).$$

Además podemos calcular la media a posteriori

$$\int_\Theta \theta\pi(\theta \mid x_1, ..., x_n)\,d\theta = \frac{p}{a} = \frac{\alpha + \sum x_i}{\beta + n}$$

este sería un estimador puntual bayesiano.

2. $(X_i)_{i=1}^n$ es una m.a. de $X \sim Gamma\,(p=1, a=\theta)$

$$f_\theta(x) = f(x \mid \theta) = \frac{\theta}{\Gamma(1)} e^{-\theta x} x^{1-1} = \theta e^{-\theta x}, \qquad x > 0$$

familia de muestras para una Poisson

$$P(x \mid \theta) = \prod P(x_i \mid \theta) \propto \theta^n e^{-\theta\sum x_i},$$

$P(x \mid \theta)$ es el núcleo de una densidad tipo Gamma, entonces tomamos como posible familia conjugada

$$\mathcal{P} = \{\pi(\theta) \in Gamma(p, a)\},$$

recordamos que

$$f \sim \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1},$$

por lo que

$$\pi(\theta) \sim \frac{a^\theta}{\Gamma(\theta)} e^{-a\theta} \theta^{p-1},$$

así que

$$\pi(\theta \mid x_1, ..., x_n) \propto f(x_i \mid \theta)\, \pi(\theta) \propto \theta^n e^{-\theta \sum x_i},$$

para $\theta > 0$. por lo tanto

$$\pi(\theta \mid x_1, ..., x_n) \sim Gamma\left(p = \alpha + \sum x_i, a = \beta + n\right).$$

■

Ejercicio 2.6.3 *Sea X una observación de la densidad*

$$f(x \mid \theta) = \frac{2x}{\theta^2} I_{(0,\theta)}(x), \qquad \theta > 0,$$

supongamos que θ tiene una distribución a priori uniforme sobre $(0,1)$. Hallar la mediana de la distribución a posteriori.

Solución. Tenemos la siguiente situación:

$$f(x \mid \theta) = \frac{2x}{\theta^2} I_{(0,\theta)}(x), \qquad \theta > 0,$$

y suponemos que

$$\pi(\theta) = U[0,1] = 1, \qquad \theta \in (0,1),$$

entonces:

$$\pi(\theta \mid x_i) \propto f(x_i \mid \theta)\, \pi(\theta) \propto \frac{2x}{\theta^2} 1 = \frac{2x}{\theta^2}, \qquad \theta \in (0,1),$$

pero $x \in (0,1)$, $x < \theta$, por lo que $\theta \in (x, 1)$, así que

$$\pi(\theta \mid x_i) \propto \frac{f(x_i \mid \theta)\, \pi(\theta)}{\int_x^1 f(x_i \mid \theta)\, d\theta} = \frac{\frac{2x}{\theta^2}}{\int_x^1 \frac{2x}{\theta^2} d\theta} = \frac{x}{\theta^2 (1-x)}, \qquad \theta \in (x, 1).$$

Para calcular la media a posteriori vemos que

$$\frac{1}{2} = \int_{-\infty}^{M} \pi(\theta \mid x_i)\, d\theta = \int_{-\infty}^{M} \frac{x}{\theta^2 (1-x)} d\theta$$

por lo que
$$M=\frac{2x}{1+x},$$
tal y como queríamos calcular. ■

Ejercicio 2.6.4 *Encontrar la cota de FCR y el estimador eficiente (si existe) en los siguientes casos:*

1. *m.a. de*
$$f_{\theta}(x)=\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right),\qquad \theta>0,\qquad x>0,$$
para estimar θ
2. *m.a. de*
$$f_{\theta}(x)=\theta\left(1-\theta\right)^{x},\qquad x=0,1..,\qquad \theta\in\left(0,1\right),$$
para estimar θ
3. *m.a. de*
$$f_{\sigma}(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^{2}}{2\sigma^{2}}\right),$$
para estimar σ, lo mismo para estimar σ^2.

Solución. Recordamos que
$$FCR=\frac{\left(\frac{d\tau}{d\theta}\right)^{2}}{-nE_{\theta}\left[\frac{d^{2}\log f_{\theta}(x)}{d\theta^{2}}\right]},\qquad I.F.=-nE_{\theta}\left[\frac{d^{2}\log f_{\theta}(x)}{d\theta^{2}}\right]$$
mientras que el estimador eficiente
$$\tau+\frac{\frac{d\tau}{d\theta}}{I.Fisher}\frac{d\log f_{\theta}(x_{i})}{d\theta},\qquad T=\tau+\frac{\tau'}{IF}\left(\log f_{\theta}\right)',$$
así que:

1. En este caso
$$f_{\theta}(x)=\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right),\qquad \theta>0,\qquad x>0,$$
por lo que tomando logaritmos tenemos que
$$\log f_{\theta}(x)=\log\left(\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)\right)=-\log\theta-\frac{x}{\theta},$$
y de forma inmediata se calculan las derivadas etc...
$$\begin{aligned}\partial_{\theta}\log f_{\theta}(x)&=-\frac{1}{\theta}+\frac{x}{\theta^{2}},\\ \partial_{\theta^{2}}^{2}\log f_{\theta}(x)&=\frac{1}{\theta^{2}}-\frac{2x}{\theta^{3}},\end{aligned}$$

de esta forma vemos que

$$E_\theta\left[\frac{1}{\theta^2}-\frac{2x}{\theta^3}\right]=\frac{1}{\theta^2}-\frac{2}{\theta^3}E\left[x\right]=\frac{1}{\theta^2}-\frac{2}{\theta^3}\theta=-\frac{1}{\theta^2}$$

ya que

$$E\left[X\right]=\int_{\mathbb{R}}x\frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right)dx=\theta$$

además podemos observar que

$$E\left[X\right]=\frac{p}{a}=\theta$$

ya que $\Gamma\left(p=1,a=\frac{1}{\theta}\right)$. Por lo tanto:

$$FCR=\frac{\left(\frac{d\tau}{d\theta}\right)^2}{-nE_\theta\left[\frac{d^2\log f_\theta(x)}{d\theta^2}\right]}=\frac{1}{-n\left(-\frac{1}{\theta^2}\right)}=\frac{\theta^2}{n}.$$

Para calcular el estimador eficiente vemos que

$$f_\theta(x_i)=\prod f_\theta(x_i)=\frac{1}{\theta^n}\exp\left(-\frac{\sum x_i}{\theta}\right)$$

tomando logaritmos

$$\log\left(f_\theta(x_i)\right)=\log\left(\frac{1}{\theta^n}\exp\left(-\frac{\sum x_i}{\theta}\right)\right)=-n\log\theta-\frac{\sum x_i}{\theta}$$

y por lo tanto

$$\partial_\theta\left(\log\left(f_\theta(x_i)\right)\right)=\partial_\theta\left(-\frac{n}{\theta}--\frac{\sum x_i}{\theta}\right)=-\frac{n}{\theta}+\frac{\sum x_i}{\theta^2}$$

de esta forma llegamos a que

$$\begin{aligned} T &= \tau+\frac{\frac{d\tau}{d\theta}}{I.Fisher}\frac{d\log f_\theta(x_i)}{d\theta}= \\ &= \theta+\frac{1}{-n\left(-\frac{1}{\theta^2}\right)}\left(-\frac{n}{\theta}+\frac{\sum x_i}{\theta^2}\right)=\theta+\frac{\theta^2}{n}\left(-\frac{n}{\theta}+\frac{\sum x_i}{\theta^2}\right)=\frac{\sum x_i}{n}.\end{aligned}$$

2. De forma mécanica vemos que

$$f_\theta(x)=\theta\left(1-\theta\right)^x,\qquad x=0,1..,\qquad \theta\in(0,1),$$

es una geométrica

$$E\left[X\right]=\frac{1-\theta}{\theta},\qquad I.F.=\frac{n}{\theta^2\left(1-\theta\right)}$$

por lo que

$$FCR=\frac{\theta^2\left(1-\theta\right)}{n}$$

sin embargo

$$T=\theta-\frac{\theta^2\left(1-\theta\right)}{n}\left(\frac{d\log f_\theta(x_i)}{d\theta}\right)=2\theta-\theta^2-\theta^2\sum x_i$$

i.e. no es un estimador eficiente. FALTA completar las cuentecillas.

$$\begin{aligned} \log\left(f_{\theta}(x_i)\right) &= \log\left(\theta^n\left(1-\theta\right)^{\sum x_i}\right). \\ \frac{d\log f_{\theta}(x_i)}{d\theta} &= \frac{n}{\theta}-\frac{\sum x_i}{(1-\theta)}, \end{aligned}$$

3. En este caso

$$f_{\sigma}(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)\in N\left(0,\sigma^2\right),$$

evitando pasos intermedios vemos que

$$FCR=\frac{\left(\frac{d\tau}{d\sigma}\right)^2}{-nE_{\sigma}\left[\frac{d^2\log f_{\sigma}(x)}{d\sigma^2}\right]}=\frac{1}{-nE_{\sigma}\left[\frac{1}{\sigma^2}-\frac{3x^2}{\sigma^4}\right]}=\frac{1}{-\frac{n}{\sigma^2}+\frac{3n}{\sigma^4}E_{\sigma}\left[X^2\right]}$$

observándose que: $var(X)=E_{\sigma}\left[X^2\right]-E_{\sigma}^2\left[X\right]$

$$E_{\sigma}\left[X^2\right]=var(X)+E_{\sigma}^2\left[X\right]=var(X)=\sigma^2$$

ya que $X\in N\left(0,\sigma^2\right)$, así que

$$FCR=\frac{\sigma^2}{2n}.$$

El estimador eficiente:

$$T=\tau+\frac{\frac{d\tau}{d\theta}}{I.Fisher}\frac{d\log f_{\theta}(x_i)}{d\theta}=\sigma+\frac{\sigma^2}{2n}\left(\frac{d\log f_{\theta}(x_i)}{d\theta}\right)$$

donde

$$\left(\frac{d\log f_{\theta}(x_i)}{d\theta}\right)=\frac{d}{d\theta}\log\left(\frac{1}{\left(\sigma\sqrt{2\pi}\right)^n}\exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)\right)=-\frac{n}{\sigma}+\frac{\sum x_i^2}{\sigma^3}$$

entonces

$$T=\sigma+\frac{\sigma^2}{2n}\left(-\frac{n}{\sigma}+\frac{\sum x_i^2}{\sigma^3}\right)$$

por lo que no hay estimador eiciente para σ.

Sin embargo si repetimos estas cuentecillas para σ^2, vemos que

$$FCR=\frac{\left(\frac{d\tau}{d\sigma}\right)^2}{-nE_{\sigma}\left[\frac{d^2\log f_{\sigma}(x)}{d\sigma^2}\right]}=\frac{4\sigma^2}{-\frac{n}{\sigma^2}+\frac{3n}{\sigma^2}}=\frac{2}{n}\sigma^4.$$

mientras que el estimador eficiente verifica

$$\begin{aligned} T &= \tau+\frac{\frac{d\tau}{d\theta}}{I.Fisher}\frac{d\log f_{\theta}(x_i)}{d\theta}= \\ &= \sigma^2+\frac{2\sigma}{\frac{2n}{\sigma^2}}\left(-\frac{n}{\sigma}+\frac{\sum x_i^2}{\sigma^3}\right)=\frac{\sum x_i^2}{n} \end{aligned}$$

luego $T=\frac{\sum x_i^2}{n}$ es un estimador eficiente en este caso.

Tal y comoqueríamos hacer ver. ■

Ejercicio 2.6.5 *Sea* $(X_i)_{i=1}^n$ *m.a. de una distribución tipo Poisson* λ.

1. *Hallar la cota inferior de FCR para la varianza de los estimadores insesgados de* $e^{-\lambda}$.
2. *Determinar el valor de la constante* c *tal que el estimador* $\exp\left(-c\sum X_i\right)$ *sea un estimador insesgado de* $e^{-\lambda}$.

Solución. De forma inmediata vemos que

$$FCR = \frac{\left(\frac{d\tau}{d\lambda}\right)^2}{-nE_\lambda\left[\frac{d^2\log f_\lambda(x)}{d\lambda^2}\right]} = \frac{-e^{2\lambda}}{nE_\lambda\left[\frac{d^2\log f_\lambda(x)}{d\lambda^2}\right]} = \frac{\lambda e^{-2\lambda}}{n}$$

no olvidar que estamos estimando $e^{-\lambda}$.

$T = \exp\left(-c\sum X_i\right)$ insesgado para estimar $e^{-\lambda}$.?? Recordamos que insesgado significa:

$$E_\theta[T] - \tau(\theta) = 0 \qquad \Longleftrightarrow \qquad E_\theta[T] = e^{-\lambda}$$

por lo tanto

$$\begin{aligned} e^{-\lambda} &= E[T] = E\left[\exp\left(-c\sum X_i\right)\right] = \prod E\left[\exp(-cx_i)\right] = \left(E\left[\exp(-cx_i)\right]\right)^n \\ &= \left(\exp\left(-\lambda\left(1-e^c\right)\right)\right)^n = \exp\left(-\lambda n\left(1-e^c\right)\right) \end{aligned}$$

despejando se obtiene

$$c = -\log\left(1-\frac{1}{n}\right).$$

veamos un paso delicado de la anterior cadenas de igualdades:

$$E\left[\exp(-cx)\right] = \sum e^{-cx}P(X=x) = \sum_{x=0}^{\infty} e^{-cx}\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(e^{-c}\lambda\right)^x}{x!} = e^{-\lambda}e^{-\lambda e^{-c}}.$$

tal y como queríamos hacer ver. ■

Ejercicio 2.6.6 *Sea* $(X_i)_{i=1}^n$ *m.a. de una distribución tipo uniforme* $U(0,\theta)$, $\theta > 0$.

1. *Hallar la cota inferior de FCR para la varianza de los estimadores insesgados de* θ.
2. *Estudiar si* $X_{(n)}$ *es un estimador insesgado para* θ.
3. *Hallar el estimador de máxima verosimilitud para* θ.

Solución. En este caso vemos que

$$E[T] = E\left[X_{(n)}\right] = \int_{\mathbb{R}} t f_{X_{(n)}}(t) dt$$

para calcular $f_{X_{(n)}}(t)$ puedo mirar la función de densidad del máximo:

$$F_{X_{(n)}}(t) = \begin{cases} 0 & t < 0, \\ \frac{t^n}{\theta^n} & t \in (0,\theta), \\ 1 & t > \theta, \end{cases}$$

vemos que

$$F_{X_{(n)}}(t) = \prod P(X_i \leq t) = (P(X_i \leq t))^n = \left(\int_{-\infty}^{t} f dx\right)^n = \left(\int_0^t \frac{1}{\theta} dx\right)^n = \frac{t^n}{\theta^n}$$

así

$$f_{X_{(n)}} = \begin{cases} n\frac{t^{n-1}}{\theta^n} & t \in (0,\theta) \\ 0 & resto \end{cases}$$

de esta forma

$$E[T] = E\left[X_{(n)}\right] = \int_{\mathbb{R}} t f_{X_{(n)}}(t) dt = \int_0^\theta n\frac{t^n}{\theta^n} dt = \frac{n}{n+1}\theta$$

por lo que no es insesgado para estimar θ. Recordatorio: $E_\theta[T] - \tau(\theta) = 0$.

Hallar el estimador de máxima verosimilitud para θ. En este caso

$$f_\theta(x) = \frac{1}{\theta}, \qquad x \in (0,\theta),$$

por lo que la función

$$L(\theta) = f_\theta(x_i) = \prod f_\theta(x_i) = \frac{1}{\theta^n},$$

así que

$$\log L(\theta) = -n\log\theta, \qquad \Longrightarrow \qquad \partial_\theta \log L(\theta) = -\frac{n}{\theta}$$

de esta forma llegamos a que

$$\partial_\theta \log L(\theta) = -\frac{n}{\theta} = 0 \qquad \Longrightarrow \qquad \hat{\theta} = \infty,$$

llegando así a una cagada monumental.

Rehacemos el camino para hacer ver que en esta ocasión $f_\theta(x) = \frac{1}{\theta}$, $x \in (0,\theta)$, pero $L(\theta) = \frac{1}{\theta^n}$, si $\theta \geq X_{(n)}$, en este caso el rango de las observaciones depende del parámetro por lo que hay que ser un poco más cuidadoso a la hora de escribir las cosas. Así se llega a la conclusión de que

$$\hat{\theta} = X_{(n)}$$

tal y como queríamos hacer ver. ■

Ejercicio 2.6.7 *Sea* $(X_i)_{i=1}^n$ *m.a. de la densidad*

$$f_\theta(x) = \theta \exp(-\theta x), \qquad x \geq 0, \theta > 0.$$

Encontrar el estimador de máxima versimilitud de θ*. ¿Es consistente para* θ*?.*

Solución. En este caso $f_\theta(x)$ es una exponencial (ver primer ejercicio), además el rango no depende del parámetro. $\hat{\theta} = \frac{n}{\sum x_i}$, ¿Es consistente para θ? *.i.e.* $\hat{\theta} \xrightarrow{P} \theta$.?? Para verlo tendremos que comprobar si $E\left[\hat{\theta}\right] \overset{n\to\infty}{\longrightarrow} \theta$, y $var\left[\hat{\theta}\right] \overset{n\to\infty}{\longrightarrow} 0$.

Por lo tanto:

$$\begin{aligned} E\left[\hat{\theta}\right] &= E\left[\frac{n}{\sum x_i}\right] = E\left[\frac{n}{Y}\right] = \int_0^\infty \frac{n}{y} f(y)\, dy = \int_0^\infty \frac{n}{y} \frac{\theta^n}{\Gamma(n)} e^{-\theta y} y^{n-1} dy = \\ &= n\frac{\theta^n}{\Gamma(n)} \int_0^\infty e^{-\theta y} y^{n-2} dy = n\frac{\theta^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\theta^{n-1}} = \frac{n}{n-1}\theta \end{aligned}$$

por lo tanto

$$E\left[\hat{\theta}\right] = E\left[\frac{n}{Y}\right] = \frac{n}{n-1}\theta, \qquad E\left[\hat{\theta}\right] \overset{n\to\infty}{\longrightarrow} \theta$$

en estos cálculos hemos empleado los siguientes trucos: $Y = \sum x_i \sim Gamma(p = n, a = \theta)$. las exponenciales son casos particulares de una Gamma y la suma de Gammas es otra Gamma. Si $X \sim Gamma(p, a)$ entonces:

$$f(x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}, \qquad \int_0^\infty e^{-ax} x^{p-1} dx = \frac{\Gamma(p)}{a^p}, \qquad \Gamma(p) = (p-1)\Gamma(p-1).$$

Mientras que

$$\begin{aligned} var\left[\hat{\theta}\right] &= var\left(\frac{n}{\sum x_i}\right) = n^2 var\left(\frac{1}{Y}\right) = n^2\left(E\left[Y^{-2}\right] - E^2\left[Y^{-1}\right]\right) = \\ &= n^2\left(\int_0^\infty \frac{1}{y^2}\frac{\theta^n}{\Gamma(n)} e^{-\theta y} y^{n-1} dy - \frac{\theta^2}{(n-1)^2}\right) = n^2\left(\frac{\theta^2}{(n-1)(n-2)} - \frac{\theta^2}{(n-1)^2}\right) \end{aligned}$$

de esta forma $var\left[\hat{\theta}\right] \overset{n\to\infty}{\longrightarrow} 0$, así que $\hat{\theta}$ es consistente. ■

Ejercicio 2.6.8 *Sea* $(X_i)_{i=1}^n$ *m.a. de una distribución tipo uniforme* $U(0,\theta)$, $\theta \in (0,\infty)$*. Demostrar que* $X_{(n)}$ *es consistente para* θ*. ¿Es* $Y_n = 2\overline{X}_n$ *consistente para* θ*?*

Solución. Seguimos los mismos que en ejercicios anteriores viendo que

$$F_{X_{(n)}}(t) = \begin{cases} 0 & t < 0, \\ \frac{t^n}{\theta^n} & t \in (0,\theta), \\ 1 & t > \theta, \end{cases}$$

observamos que

$$F_{X_{(n)}}(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq \theta \end{cases}$$

comprobando que

$$X_{(n)} \xrightarrow{L} D(\theta) \Longleftrightarrow X_{(n)} \xrightarrow{P} \theta \Longleftrightarrow X_{(n)} \text{ consistente para } \theta.$$

¿Es $Y_n = 2\overline{X}_n$ consistente para θ?. Esperamos que $2\overline{X}_n$ esté cerca de θ.

$$\begin{aligned} E[Y_n] &= E\left[2\overline{X}_n\right] = 2E\left[\overline{X}_n\right] = 2E[X] = 2\frac{\theta}{2} = \theta \overset{n\to\infty}{\longrightarrow} \theta, \\ var(Y_n) &= var(2\overline{X}_n) = 4var(\overline{X}_n) = 4\frac{var(X)}{n} = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n} \overset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

por lo que es consistente para θ. Sin embargo tiene una pequeña pega pues si tomamos $X \sim U(0,10)$ supongamos que $\bar{X} = 6$, entonces $Y_n = 2\overline{X}_n = 12$, por lo que a veces sobre-estima. ■

Ejercicio 2.6.9 *Sea $(X_i)_{i=1}^n$ m.a. de una distribución discreta con función de masa*

$$P_\theta(X = x) = \theta(1-\theta)^{x-1}, \qquad x = 1, 2, ..., \theta \in (0,1)$$

1. *Estudiar si $T = \sum X_i$, es un estadístico suficiente para θ.*
2. *Hallar el estimador de θ por el método de los momentos,*
3. *Obtener el estiamdor de máxima verosimilitud de θ.*
4. *Calcular la media de la distribución a posteriori cuando la distribución a priori para el parámetro $\theta \in (0,1)$.*
5. *Calcular la cota de FCR para la varianza de los estiamdores insesgados para $\tau(\theta) = \frac{1}{\theta}$.*
6. *¿Hay estimador eficiente para estimar $\tau(\theta) = \frac{1}{\theta}$?.*

Solución. De forma telegráfica vemos que

1. Sí, aplicando el teorema de facotrización obtenemos

$$\prod \left(\theta(1-\theta)^{x_i-1}\right) = \left(\theta^n (1-\theta)^{\sum x_i - n}\right)$$

   por lo tanto $T = \sum X_i$, es un estadístico suficiente para θ

2. Sabemos que al tratarse de un geométrica

$$E[X] = \frac{1}{p} = \frac{1}{\theta}$$

   y por lo tanto

$$\tilde{\theta} = \frac{n}{\sum x_i},$$

3. Vemos que
$$L(\theta) = \left(\theta^n (1-\theta)^{\sum x_i - n}\right)$$
así que tomando logaritmos
$$\log L = n \log \theta + \left(\sum x_i - n\right) \log (1-\theta)$$
calculamos el máximo, encontrando
$$\frac{n}{\theta} - \frac{(\sum x_i - n)}{(1-\theta)} = 0$$
de esta forma llegamos a:
$$\hat{\theta} = \frac{n}{\sum x_i},$$

4.
$$\pi(\theta \mid x_i) \sim Beta\left(p = n+1, q = \sum x_i - n + 1\right)$$
La media a posteriori es:
$$E[X] = \frac{p}{p+q} = \frac{n+1}{\sum x_i + 2}$$

5.
$$T = \frac{\sum x_i}{n}$$
es eficiente para $\tau = \frac{1}{\theta}$, donde
$$E[X] = \sum x\theta (1-\theta)^{x-1} = \frac{1}{\theta}.$$

■

Capítulo 3

Intervalos de confinaza

3.1. Introducción

Disponemos de $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. Como siempre queremos estimar θ.

Definición 3.1.1 *Un estimador por intervalos de confienza para estimar θ_i es una función que a cada muestra $(X_i)_{i=1}^n$ le asigna un intervalo*

$$(L, U) = \left(L\left(x_i\right)_{i=1}^n, U\left(x_i\right)_{i=1}^n\right).$$

Definición 3.1.2 *El nivel de confianza de un intervalo (L, U) es $1-\alpha$ cuando*

$$P_\theta\left\{\theta \in (L, U)\right\} = P_\theta\left\{(x_i)_{i=1}^n : L < \theta < U\right\} = 1-\alpha.$$

Consideramos los siguientes métodos de construcción

1. Método clásico (cantidad pivotal), intervalos de confianza más frecuentes en las aplicaciones.
2. Métodos bayasianos.

3.2. Método de la cantidad pivotal.

Definición 3.2.1 *Una cantidad pivotal para estimar θ_i es una función*

$$T\left((x_i)_{i=1}^n : \theta_i\right)$$

cuya distribución no depende de θ_i.

La idea es la de obtener una cantidad pivotal T para θ_i, que sea una función continua y monótona de θ_i. Queremos hallar $\left(\varepsilon\left(\alpha_1\right), \varepsilon\left(\alpha_2\right)\right)$ tales que

$$P_\theta\left\{(x_i)_{i=1}^n : \varepsilon\left(\alpha_1\right) < T < \varepsilon\left(\alpha_2\right)\right\} = 1-\alpha.$$

donde los $(\varepsilon(\alpha_1), \varepsilon(\alpha_2))$ no dependen de θ_i por ser T una cantidad pivotal.

Despejamos θ_i de la anterior ecuación i.e.

$$\begin{aligned} \varepsilon(\alpha_1) &< T \\ T &< \varepsilon(\alpha_2) \end{aligned}$$

obteniendo así (L, U).

3.2.1. Cantidades pivotales e intervalos de confianza más frecuentes.

1. $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x) = N\left(\mu, \sigma_0^2\right)$ donde σ_0^2 es una cantidad conocida. El objetivo es estimar μ mediante IC con un nivel de confianza (NC) $1-\alpha$.

   Para ello seguiremos la siguiente táctica:

   Consideramos $\bar{X} \sim N\left(\mu, \frac{\sigma_0^2}{n}\right)$ y tipificando i.e.

   $$\frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$$

   nos preguntamos si

   $$T = \frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}}$$

   es una cantidad pivotal para estimar μ. Además, para una muestra fija ¿es T continua y monótona?. Ambas respuestas son afirmativas.

   $$P_\theta\left\{(x_i)_{i=1}^n : \varepsilon(\alpha_1) < T < \varepsilon(\alpha_2)\right\} = 1-\alpha,$$
   $$P_\theta\left\{(x_i)_{i=1}^n : \varepsilon(\alpha_1) < \frac{\bar{X}-\mu}{\sigma_0/\sqrt{n}} < \varepsilon(\alpha_2)\right\} = 1-\alpha,$$

   vemos que $T \sim N(0,1)$ entonces por simetría la última ecuación la podemos reescribir como

   $$P_\theta\left\{(x_i)_{i=1}^n : -Z_{\alpha/2} < T < Z_{\alpha/2}\right\} = 1-\alpha,$$

   así que despejamos

   $$\begin{aligned} -Z_{\alpha/2} &< T, \\ T &< Z_{\alpha/2}, \end{aligned}$$

   de donde obtenemos que

   $$\mu < \bar{X} + Z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}, \qquad \mu > \bar{X} - Z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}},$$

   por lo que llegamos a:

   $$IC_{1-\alpha} = (L, U) = \left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right).$$

   Vemos de esta forma que al aumentar n la longitud del intervalo decrece (i.e. la estimación es más precisa). De igual forma se observa que al aumentar el NC $1-\alpha$ entonces $Z_{\alpha/2}$ aumenta i.e. la longitud del intervalo aumenta y por lo tanto la estimación es menos precisa).

2. $(X_i)_{i=1}^n$ m.a. de $X \sim N\left(\mu, \sigma^2\right)$. El objetivo es estimar μ mediante IC con un nivel de confianza (NC) $1-\alpha$. Podemos intentar seguir la táctica anterior i.e. consideramos $\bar{X} \sim N\left(\mu, \frac{\sigma_0^2}{n}\right)$ y tipificando i.e.
$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$
y nos preguntamos si
$$T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$
es una cantidad pivotal para estimar μ. Además, para una muestra fija ¿es T continua y monótona?. En esta caso no, ya que T depende de un parámetro desconocido, σ, por lo tanto no puede ser cantidad pivotal.

   En este caso deberemos apelar al teorema de Fisher (recordar el primer capítulillo) viendo que
$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1), \qquad \frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2_{n-1}$$
por lo tanto
$$\frac{N(0,1)}{\sqrt{\chi^2_{n-1}/(n-1)}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$
por lo tanto si definimos
$$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$
ahora sí es cantidad pivotal para estimar μ. Por lo tanto y siguiendo el procedimiento anterior vemos que
$$P\left\{-t_{n-1;\alpha/2} < T < t_{n-1;\alpha/2}\right\} = 1-\alpha$$
donde hemos vuelto a tener en cuenta las simetrías de la distribución. Así que despejando se llega a que
$$IC_{1-\alpha}(\mu) = \left(\bar{X} \pm t_{n-1;\alpha/2}\frac{S}{\sqrt{n}}\right).$$
Siguiendo los mismos paso, podemos ahora estimar σ^2. Para ello consideramos
$$T = \frac{(n-1)\,S^2}{\sigma^2},$$
como cantidad pivotal y obtenemos por lo tanto
$$IC_{1-\alpha}\left(\sigma^2\right) = \left(\frac{(n-1)\,S^2}{\chi^2_{n-1;\alpha/2}}, \frac{(n-1)\,S^2}{\chi^2_{n-1;1-\alpha/2}}\right).$$

3. $(X_i)_{i=1}^n$ m.a. de $X \sim Bernoulli(p)$. El objetivo es estimar p mediante IC con un nivel de confianza (NC) $1-\alpha$. En este caso seguremos las siguientes consideraciones:
$$\sum_{i=1}^n X_i = \text{“\# \textit{éxitos}”} \sim Bin(n,p) \overset{T^a \text{ De Moivre}}{\sim} N\left(\mu = np, \sigma^2 = npq\right)$$

por lo tanto

$$\sum_{i=1}^{n} X_i \sim N\left(\mu = np, \sigma^2 = npq\right)$$

tipificando

$$\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np\,(1-p)}} \sim N(0,1)$$

arreglando las cuentecillas llegamos a

$$\frac{\bar{X} - p}{\sqrt{\frac{p(1-p)}{n}}} \simeq \frac{\bar{X} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = \frac{\bar{X} - p}{\sqrt{\frac{\bar{X}\left(1-\bar{X}\right)}{n}}} \sim N(0,1)$$

así que

$$T = \frac{\bar{X} - p}{\sqrt{\frac{\bar{X}\left(1-\bar{X}\right)}{n}}}.$$

observamos que para llegar a este resultado hemos utilizado resultdos previos como $\hat{p} = \bar{X}$. Por lo tanto y siguiendo los mismos pasos que los anteriores casos vemos que

$$IC_{1-\alpha}\,(p) = \left(\bar{X} \pm Z_{\alpha/2}\sqrt{\frac{\bar{X}\left(1-\bar{X}\right)}{n}}\right).$$

3.2.2. Cantidad pivotal general.

Sea $(X_i)_{i=1}^{n}$ m.a. de $X \sim F_\theta$ distribución continua. Podemos definir una cantidad pivotal genérica como sigue:

$$T = -\sum_{i=1}^{n} \log F_\theta\,(x_i)$$

es cantidad pivotal para estimar el parámetro θ, pues depende de $(X_i)_{i=1}^{n}$ y θ y se comprueba con facilidad que la función de distribución de T no depende de θ.

Definición 3.2.2 *Definimos el error en la estimación de un intervalo de confienaza como*

$$E = \frac{longitud\ del\ intervalo}{2}.$$

En muchos estudios es muy interesante saber el tamaño muestral necesario para obtener un error en la estimación inferior a E.

Ejemplo 3.2.1 *Sea $(X_i)_{i=1}^{n}$ m.a. de $X \sim N\left(\mu, \sigma_0^2\right)$ donde σ_0^2 es una cantidad conocida. ¿Cuál es el mínimo valor de n para que el error en la estimación de un intervalo con nivel de confianza $1-\alpha$ sea inferior a E?*

Sabemos que

$$IC_{1-\alpha} = \left(\bar{X} \pm Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}} \right).$$

por lo tanto

$$E = Z_{\alpha/2} \frac{\sigma_0}{\sqrt{n}}$$

por lo que de aquí despejamos n

$$Z_{\alpha/2}\sigma_0 \leq E\sqrt{n}$$

así que

$$n \geq \left(\frac{Z_{\alpha/2}\sigma_0}{E} \right)^2.$$

3.3. Intervalos de confianza bayesianos.

Sea $(X_i)_{i=1}^n$ m.a. de $X \sim f(X \mid \theta)$ donde $\theta \in \Theta$. Tenemos la información muestral (verosimilitud)

$$f(x_1, ..., x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

y de la información a priori

$$\pi(\theta)$$

así que aplicando Bayes

$$\pi(\theta \mid x_1, ..., x_n) = \frac{\pi(\theta) \prod_{i=1}^{n} f(x_i \mid \theta)}{\int_\theta \pi(\theta) \prod_{i=1}^{n} f(x_i \mid \theta)} \propto \pi(\theta) f(x_1, ..., x_n \mid \theta)$$

y todas las conclusiones las obtendremos a partir de $\pi(\theta \mid x_1, ..., x_n)$ la distribución a posteriori.

Definición 3.3.1 *Un intervalo de confianza bayesiano al nivel* $1-\alpha$ *es un intervalo* (L, U) *tal que*

$$\int_L^U \pi(\theta \mid x_1, ..., x_n)\, d\theta = 1 - \alpha.$$

Definición 3.3.2 *El intervalo de confianza bayesiano con máxima densidad a posteriori (HPD) al nivel* $1-\alpha$ *es el intervalo* (L, U) *tal que*

$$\int_L^U \pi(\theta \mid x_1, ..., x_n)\, d\theta = 1 - \alpha,$$

y $\forall \theta_1 \in (L, U)$ *y* $\forall \theta_2 \notin (L, U)$,

$$\pi(\theta_1 \mid x_1, ..., x_n) > \pi(\theta_2 \mid x_1, ..., x_n).$$

El intervalo HPD es el de mínima longitud para un nivel $1-\alpha$ fijado.

3.4. Resumen de intervalos

INTERVALOS DE CONFIANZA

NOTACIÓN:

$(X_1, \ldots, X_n)$ muestra aleatoria (m. a.) de X.

$$\bar{x} = \frac{1}{n}\sum x_i \qquad s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2.$$

3.4.1. $X \sim N(\mu, \sigma)$

Intervalos de confianza $1-\alpha$ **para** μ:

1. σ conocida:
$$I = \left(\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

2. σ desconocida
$$I = \left(\bar{x} \pm t_{n-1;\alpha/2}\frac{s}{\sqrt{n}}\right).$$

Intervalo de confianza $1-\alpha$ **para** σ^2:

$$I = \left(\frac{(n-1)s^2}{\chi^2_{n-1;\alpha/2}}\ ,\ \frac{(n-1)s^2}{\chi^2_{n-1;1-\alpha/2}}\right).$$

3.4.2. $X \sim$ Bernoulli(p) (muestras grandes).

Intervalo de confianza $1-\alpha$ **para** p:

$$I = \left(\bar{x} \pm z_{\alpha/2}\sqrt{\frac{\bar{x}(1-\bar{x})}{n}}\right).$$

3.4.3. $X \sim$ Poisson(λ) (muestras grandes)

Intervalo de confianza $1-\alpha$ **para** λ:

$$I = \left(\bar{x} \pm z_{\alpha/2}\sqrt{\bar{x}/n}\right).$$

3.4.4. Dos poblaciones Normales independientes

$X \sim N(\mu_1, \sigma_1)$; $(X_1, \ldots, X_{n_1})$ m. a. de X; se calcula $\bar{x}$ y s_1^2.

$Y \sim N(\mu_2, \sigma_2)$; $(Y_1, \ldots, Y_{n_2})$ m. a. de Y; se calcula $\bar{y}$ y s_2^2.

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}.$$

Intervalos de confianza $1-\alpha$ para $\mu_1 - \mu_2$:

1. $\sigma_1,\ \sigma_2$ conocidas:
$$I = \left(\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right).$$

2. $\sigma_1,\ \sigma_2$ desconocidas, $\sigma_1 = \sigma_2$:
$$I = \left(\bar{x} - \bar{y} \pm t_{n_1+n_2-2;\alpha/2}\ s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right).$$

3. $\sigma_1,\ \sigma_2$ desconocidas, $\sigma_1 \neq \sigma_2$:
$$I = \left(\bar{x} - \bar{y} \pm t_{f;\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right),$$
donde f = entero más próximo a $\dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$.

Intervalo de confianza $1-\alpha$ para σ_1^2/σ_2^2:

$$I = \left(\frac{s_1^2/s_2^2}{F_{n_1-1;n_2-1;\alpha/2}}\ ,\ \frac{s_1^2/s_2^2}{F_{n_1-1;n_2-1;1-\alpha/2}}\right).$$

3.4.5. Comparación de proporciones (muestras grandes e independientes)

$X \sim \text{Bernoulli}(p_1); (X_1, \ldots X_{n_1})$ m. a. de X.

$Y \sim \text{Bernoulli}(p_2); (Y_1, \ldots Y_{n_2})$ m. a. de Y.

Intervalo de confianza $1-\alpha$ para $p_1 - p_2$:

$$I = \left(\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{\bar{x}(1-\bar{x})}{n_1} + \frac{\bar{y}(1-\bar{y})}{n_2}}\right).$$

3.5. Ejercicios

Ejercicio 3.5.1 *Sea* (X_1, X_2, X_3, X_4) *una m.a. de* $X \sim N\left(\mu, \sigma^2 = 1\right)$. *Hallar el nivel de confianza del estimador por intervalos* $\left[\bar{X} - 1, \bar{X} + 1\right]$.

Solución. Queremos calcular el NC que como sabemos

$$NC = P_\mu \left\{ (X_i)_{i=1}^4 : \mu \in \left[\bar{X} - 1, \bar{X} + 1\right] \right\} = P_\mu \left\{ \bar{X} - 1 < \mu < \bar{X} + 1 \right\}$$

si tenemos en cuenta que

$$\bar{X} = \frac{\sum X_i}{n} = \mu, \qquad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(\mu, \frac{1}{4}\right), \qquad \Longrightarrow \qquad \sigma = 1/2,$$

entonces

$$P_\mu \left\{ \bar{X} - 1 < \mu < \bar{X} + 1 \right\} = P\left\{ \frac{\mu - 1 - \mu}{1/2} < \frac{\bar{X} - \mu}{1/2} < \frac{\mu + 1 - \mu}{1/2} \right\}$$

simplificando queda:

$$P\left\{-2 < Z < 2\right\}$$

donde como es habitual $Z = \frac{\bar{X} - \mu}{1/2}$. Teniendo en cuenta las simetrías de la $N(0,1)$ la anterior expresión queda reducida a:

$$2P\left(Z > 2\right) = 0{,}9544$$

donde dicho valor se ha obtenido de las tablas para la nornal tipificada. ■

Ejercicio 3.5.2 *Sea* $(X_i)_{i=1}^n$ *una m.a. de* $X \sim U(0, \theta), \theta > 0$.

1. *Nivel de confianza del estimador por intervalos* $\left[aX_{(n)}, bX_{(n)}\right]$ *tal que* $1 \leq a < b$.
2. *Hallar* $P_\theta \left\{ \theta \in \left[X_{(n)} + c, X_{(n)} + d\right] \right\}$ *tal que* $1 \leq c < d$.

Solución. Para el primero de los apartados vemos que

$$[L, U] = \left[aX_{(n)}, bX_{(n)}\right], \qquad 1 \leq a < b,$$

por lo tanto el NC será

$$NC = P_\theta \left\{ (X_i)_{i=1}^n : \theta \in [L, U] \right\} = P_\theta \left\{ aX_{(n)} < \theta < bX_{(n)} \right\}$$

o lo que es lo mismo

$$P_\theta \left\{ \frac{\theta}{b} < X_{(n)} < \frac{\theta}{a} \right\}$$

recordar que $X_{(n)} = máx(X_i)$. Además si rememoramos las cuentecillas del primer capitulillo vemos que

$$f_{X_{(n)}} = n\left[F_X(t)\right]^{n-1} f_X(t)$$

así que

$$f_{X_{(n)}} = n\left[\frac{t}{\theta}\right]^{n-1}\frac{1}{\theta} = \frac{nt^{n-1}}{\theta^n}, \qquad \forall t \in (0,\theta)$$

por lo tanto

$$P_\theta\left\{\frac{\theta}{b} < X_{(n)} < \frac{\theta}{a}\right\} = \int_{\frac{\theta}{b}}^{\frac{\theta}{a}} f_{X_{(n)}}dt = \int_{\frac{\theta}{b}}^{\frac{\theta}{a}} \frac{nt^{n-1}}{\theta^n}dt = \frac{1}{a^n} - \frac{1}{b^n}$$

y como vemos en este caso no depende de θ.

En el segundo de los casos vemos que

$$[L,U] = \left[X_{(n)} + c, X_{(n)} + d\right], \qquad 1 \leq c < d,$$

así que seguimos el mismo procedimiento por lo que llegamos sin problemas a:

$$P_\theta\left\{X_{(n)} + c < \theta < X_{(n)} + d\right\} = P_\theta\left\{\theta - d < X_{(n)} < \theta - c\right\} = \left(1 - \frac{c}{\theta}\right)^n - \left(1 - \frac{d}{\theta}\right)^n$$

al depender de θ no se le llama nivel de confienza. ■

Ejercicio 3.5.3 *Sea $(X_i)_{i=1}^n$ una m.a. de X cuya densidad viene dada por*

$$f(x \mid \lambda) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & resto \end{cases}$$

probar que $T = \lambda\sum_{i=1}^n X_i$ es una cantidad pivotal y obtener los intervalos de confianza correspondintes.

Solución. Observamos que T es función sólo de la muestra y del parámetro y tenemos que ver que la distribución es función continua y monótona.

$$X \sim f_\lambda$$

definimos la v.a. $Y = \lambda X$ así que aplicando el c.v.

$$T = \lambda\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \sim Gamma\,(p = n, a = 1)$$

vemos que si

$$Y = \lambda X \qquad \Longrightarrow \qquad X = \frac{Y}{\lambda} \qquad \Longrightarrow \qquad \frac{dX}{dY} = \frac{1}{\lambda}$$

por lo tanto

$$f_Y = f_X\left|\frac{dX}{dY}\right| = \lambda e^{-\lambda\frac{y}{\lambda}}\frac{1}{\lambda} = e^{-y}, \qquad \forall y \in (0,\infty)$$

además se observa que $f_Y \sim Gamma\,(p = 1, a = 1)$. De esta forma vemos que al solo depender de $(X_i)_{i=1}^n$ se trata de una cantidad pivotal para el parámetro λ, igualmente se observa que es continua y monótona creciente en λ.

Buscamos ahora $(\varepsilon_1, \varepsilon_2)$ tales que

$$P\left\{\varepsilon_1 < T < \varepsilon_2\right\} = 1 - \alpha$$

así que teniendo en cuenta las simetrías de la función gamma etc... despejamos λ y vemos que

$$P\left\{\varepsilon_1 < \lambda \sum_{i=1}^{n} X_i < \varepsilon_2\right\} = 1 - \alpha$$

$$\begin{aligned} \varepsilon_1 < \lambda \sum_{i=1}^{n} X_i, \qquad &\Longrightarrow \qquad \frac{\varepsilon_1(\alpha)}{\sum_{i=1}^{n} X_i} < \lambda \\ \lambda \sum_{i=1}^{n} X_i < \varepsilon_2, \qquad &\Longrightarrow \qquad \lambda < \frac{\varepsilon_2(\alpha)}{\sum_{i=1}^{n} X_i} \end{aligned}$$

por lo tanto el IC será

$$IC_{1-\alpha}(\lambda) = \left(\frac{\varepsilon_1(\alpha)}{\sum_{i=1}^{n} X_i}, \frac{\varepsilon_2(\alpha)}{\sum_{i=1}^{n} X_i}\right)$$

Para obetener los $(\varepsilon_1, \varepsilon_2)$ calculamos

$$\int_0^{\varepsilon_1(\alpha)} f_T dt = \frac{\alpha}{2}, \qquad \int_{\varepsilon_2(\alpha)}^{0} f_T dt = \frac{\alpha}{2},$$

donde

$$\begin{aligned} T &= \lambda \sum_{i=1}^{n} X_i \sim Gamma\left(p = n, a = 1\right) \\ f_T(t) &= \frac{1}{\Gamma(n)} e^{-t} t^{n-1}, \qquad \forall t > 0, \end{aligned}$$

tal y como queríamos hacer ver. ■

Ejercicio 3.5.4 *Sea $(X_i)_{i=1}^n$ una m.a. de X cuya densidad viene dada por*

$$f(x \mid \theta) = \begin{cases} e^{-(x-\theta)} & x > \theta > 0 \\ 0 & resto \end{cases}$$

Hallar el intervalo bayesiano de máxima densidad a posteriori con nivel $4/5$ si la densidad a priori viene dada por

$$\pi(\theta) = \begin{cases} e^{-\theta} & \theta > 0 \\ 0 & resto \end{cases}.$$

Solución. Teniendo en cuenta resultados del capítulo anterior vemos que

$$\pi(\theta \mid X_i) \propto \pi(\theta) \prod_{i=1}^{n} f(x_i \mid \theta) = e^{-\theta} e^{-\sum x_i + n\theta} = e^{(n-1)\theta - \sum x_i} \propto e^{(n-1)\theta}, \qquad \forall \theta \in \left(0, X_{(1)}\right)$$

el término $e^{-\sum x_i}$ lo hemos suprimido al tratarse de una constante. Por lo tanto

$$\pi\left(\theta \mid x_i\right)=\frac{e^{(n-1)\theta}}{\int_0^{X_{(1)}} e^{(n-1)\theta}d\theta} \qquad \forall\theta\in\left(0,X_{(1)}\right)$$

integrando vemos que

$$\pi\left(\theta \mid x_i\right)=\frac{e^{(n-1)\theta}}{\frac{1}{n-1}\left(e^{(n-1)X_{(1)}}-1\right)} \qquad \forall\theta\in\left(0,X_{(1)}\right)$$

de esta forma el intervalo bayesiano viene dado por

$$(L,U)=\left(\theta_1,X_{(1)}\right)$$

tal que

$$\int_{\theta_1}^{X_{(1)}}\pi\left(\theta \mid x_i\right)d\theta=1-\alpha=\frac{4}{5}$$

así que

$$\int_{\theta_1}^{X_{(1)}}\frac{e^{(n-1)\theta}}{\frac{1}{n-1}\left(e^{(n-1)X_{(1)}}-1\right)}d\theta=\frac{4}{5}$$

por lo tanto integrando obtenemos

$$\frac{4}{5}=\frac{e^{(n-1)X_{(1)}}-e^{(n-1)\theta_1}}{e^{(n-1)X_{(1)}}-1}$$

y tomando logaritmos

$$\theta_1=\frac{1}{n-1}\log\left(e^{(n-1)X_{(1)}}\right)-\frac{4}{5}\left(\left(e^{(n-1)X_{(1)}}-1\right)\right)$$

tal y como queríamos hacer ver. ■

Ejercicio 3.5.5 *Sea $(X_i)_{i=1}^n$ una m.a. de $X\sim N\left(\mu,\sigma^2=1\right)$ y supongamos que la distribución a priori viene dada por $N(0,1)$. Hallar el IC bayesiano (L,U) de máxima densidad a posteriori con NC $1-\alpha$.*

Solución. Teniendo en cuenta resultados del capítulo anterior vemos que

$$\pi\left(\theta \mid x_i\right)\sim N\left(\frac{\tau\mu+nr\overline{X}}{\tau+nr},\frac{1}{\tau+nr}\right)=N\left(\frac{\sum x_i}{n+1},\frac{1}{n+1}\right)$$

ya que $\tau=r=1$ y $\mu=0$. Apelando a las simetrías de la normal entonces:

$$(L,U)=\left(\varepsilon_1=\frac{\sum x_i}{n+1}-a,\varepsilon_2=\frac{\sum x_i}{n+1}+a\right)$$

así que

$$\int_{\varepsilon_1}^{\varepsilon_2} \pi\left(\theta \mid x_i\right) d\theta = 1-\alpha$$

i.e.

$$P\left\{\frac{\sum x_i}{n+1} - a < Y < \frac{\sum x_i}{n+1} + a\right\} = 1-\alpha$$

donde $Y \sim \pi\left(\theta \mid x_i\right)$. Así que tipificando vemos que

$$P\left\{\frac{\frac{\sum x_i}{n+1} - a - \frac{\sum x_i}{n+1}}{\sqrt{\frac{1}{n+1}}} < \frac{Y - \frac{\sum x_i}{n+1}}{\sqrt{\frac{1}{n+1}}} < \frac{\frac{\sum x_i}{n+1} + a - \frac{\sum x_i}{n+1}}{\sqrt{\frac{1}{n+1}}}\right\} = 1-\alpha$$

simplificamos

$$P\left\{-a\sqrt{n+1} < Z < a\sqrt{n+1}\right\} = 1-\alpha$$

o lo que es lo mismo

$$P\left\{Z > a\sqrt{n+1}\right\} = \frac{\alpha}{2}$$

por lo tanto despejando a

$$a\sqrt{n+1} = Z_{\alpha/2} \qquad \Longrightarrow \qquad a = \frac{Z_{\alpha/2}}{\sqrt{n+1}}$$

de esta forma vemos que

$$(L,U) = \left(\frac{\sum x_i}{n+1} - \frac{Z_{\alpha/2}}{\sqrt{n+1}}, \frac{\sum x_i}{n+1} + \frac{Z_{\alpha/2}}{\sqrt{n+1}}\right)$$

tal y como queríamos hacer ver. ■

Ejercicio 3.5.6 *Sea X una observación de la densidad*

$$f_\theta(x) = \begin{cases} \theta x^{\theta-1} & x \in (0,1), \theta > 0 \\ 0 & resto \end{cases},$$

1. *Hallar la densidad de la variable aleatoria $Y = -\log X$.*

2. *Hallar el nivel de confianza del siguiente estimador por IC para θ*

$$I = \left[\frac{1}{2(-\log X)}, \frac{1}{-\log X}\right]$$

   Estudiar si $T = -\ \theta \log X$ es una cantidad pivotal para θ.

3. *Hallar la densidad a priori para θ si la distribución a priori es la unifirme en $(0,1)$ y el valor observado es $x = 1/e$. Indicar cuál sería, en este caso, la forma del intervalo de confianza bayesiano de máxima densidad a posteriori (para un NC $1-\alpha$.*

Solución. Vemos que $Y=-\log X$, consideramos el c.v.

$$y=-\log x,\qquad -y=\log x,\qquad x=e^{-y},\qquad \frac{dx}{dy}=-e^{-y}$$

por lo tanto

$$f_Y(y)=f_X(x)\left|\frac{dx}{dy}\right|=\theta x^{\theta-1}e^{-y}=\theta\left(e^{-y}\right)^{\theta-1}e^{-y}=\theta e^{-\theta y}$$

luego

$$f_Y(y)=\begin{cases}\theta e^{-\theta y} & \theta\in(0,\infty)\,,\\ 0 & resto\end{cases}.$$

Para calcular el nivel de confianza seguimos los mismos pasos como en los ejercicios anteriores y vemos que (se trata simplemente de despejar en y)

$$NC=P\left(\frac{1}{2\theta}\leq Y\leq\frac{1}{\theta}\right)=e^{-1/2}-e^{-1}.$$

De igual forma vemos que la densidad $T=-\,\theta\log X$ (utlizando el c.v. anterior) se llega a

$$f_T=\begin{cases}e^{-t} & t>0,\\ 0 & resto\end{cases}.$$

Por ultimo

$$\pi\left(\theta\mid x_i\right)\propto\theta e^{1-\theta},\ \text{si }\theta\in(0,1)$$

por lo tanto

$$\pi\left(\theta\mid x_i\right)=\begin{cases}\frac{\theta e^{1-\theta}}{e-2} & \theta\in(0,1)\,, \qquad x=1/e\\ 0 & resto\end{cases}$$

e $I=(a,1)$ ■

Ejercicio 3.5.7 *En una población se desea conocer la probabilidad de que un individuo sea alérgico al polen de las acacias. En* 100 *individuos tomados al azar se observaron* 10 *alérgicos. Hallar el IC al* 95 % *para la probabilidad pedida. ¿Cuántos individuos se deberían observar para que con probabilidad* 0,95 *el error máximo en la estimación de la proporción de alérgicos sea del* 0,01*?*

Solución. La estrategia en estos casos es siempre la misma. Consideramos la v.a. X y queremos estimar la probabilidad de ser alérgico i.e. si/no ($Bernoulli(p)$, éxito/fracaso). Entonces tenemos $(X_i)_{i=1}^n$ una m.a. de $X\sim Bernoulli(p)$ (con $n=100$ grande) y por lo tanto del resultado de teoría sabemos que

$$IC_{0{,}95}=\left(\bar{X}\pm Z_{\alpha/2}\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}}\right)=\left(\bar{X}\pm Z_{\alpha/2}\sqrt{\frac{\bar{X}\left(1-\bar{X}\right)}{n}}\right)$$

y lo único que hacemos es calcular

$$\bar{X} = \frac{\sum X_i}{n} = \frac{\text{n}^o \text{ alérgicos}}{n} = \frac{10}{100} = 0{,}10,$$
$$Z_{\alpha/2} = Z_{0{,}025} = 1,96$$

así que

$$IC_{0{,}95} = \left(0{,}10 \pm 1{,}96\sqrt{\frac{(0{,}10)\,(0{,}90)}{100}}\right) = (0{,}04, 0{,}16)$$

i.e. la proporción de alérgicos está entre un 4 % y un 16 %.

Queremos ahora estimar el valor de n para obetener un error maximo de 0,01 (1 %) a un NC del 95 %. Sabemos que

$$error = Z_{\alpha/2}\sqrt{\frac{\bar{X}\left(1-\bar{X}\right)}{n}} < 0{,}01$$

y tomamos como estimación para $\bar{X}$ la obetenida en la prueba piloto i.e. $\bar{X} = 0{,}10$, por lo tanto

$$1{,}96\sqrt{\frac{(0{,}10)\,(0{,}90)}{n}} < 0{,}01$$

y despejamos n obteniendo

$$n > 3458$$

por lo que concluimos que necesitamos del orden de 3500 pruebas para hacer un buen estudio. ■

Ejercicio 3.5.8 *Se supone que el número de erratas por página en un libro sigue una* $Poisson\,(\lambda)$*. Elegidas al azar* 95 *páginas se obtienen los siguientes resultados*

Nº erratas	0	1	2	3	4	5
Nº página	40	30	15	7	2	1

hallar intervalos de confianza al 90 % *para el número medio de erratas por página en todo el libro.*

Solución. Como en el ejercicio anterior podríamos modelizar la situación de la siguiente manera:

$$X \sim Bernoulli(p)$$

errata/correcto, pero como en este caso tenemos n páginas entonces $n-$Bernoullis se convierten en $Binomial\,(n,p)$. Pero de forma operativa si n es grande y p pequeña entonces es mejor aproximar la binomial por una $Poisson\,(\lambda = np)$ de la que sabemos que $E\,[X] = \lambda$.

Por lo tanto queremos estimar

$$IC_{0{,}90}\,(\lambda) = \left(\bar{X} \pm Z_{\alpha/2}\sqrt{\frac{\hat{\lambda}}{n}}\right) = \bar{X} \pm Z_{\alpha/2}\sqrt{\frac{\bar{X}}{n}}$$

en este caso tenemos que

$$Z_{\alpha/2} = 1{,}64,$$
$$\bar{X} = \frac{\sum X_i}{n} = \frac{(40 \cdot 0) + (30 \cdot 1) + ... + (1 \cdot 5)}{95} = 0{,}98,$$

por lo que

$$IC_{0,90}(\lambda) = (0{,}82, 1{,}16).$$

Si queremos afinar más entonces hacemos lo mismo que en el ejercicio anterior i.e.

$$error = Z_{\alpha/2}\sqrt{\frac{\bar{X}}{n}}$$

y despejamos n. ■

Ejercicio 3.5.9 *Se mide el tiempo (en segundos) de duración de un proceso químico realizado* 20 *veces en condiciones similares obteniéndose los siguientes resultados*

$$93, 90, 97, 90, 93, 91, 96, 94, 91, 91, 88, 93, 95, 91, 89, 92, 87, 88, 90, 86$$

suponiendo que la duración sigue una dostribución normal hallar los IC al 90 % *para ambos parámetros.*

Solución. Tenemos por lo tanto $(X_i)_{i=1}^{20}$ m.a. de $X \sim N(\mu, \sigma^2)$, para estimar μ a un NC del 90 % tenemos

$$IC_{90\,\%}(\mu) = \left(\bar{X} \pm t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}\right) = \left(\bar{X} \pm t_{n-1;\alpha/2}\frac{S}{\sqrt{n}}\right) = (90{,}11, 92{,}39)$$

ya que

$$\bar{X} = \frac{\sum X_i}{n} = 91{,}25$$
$$S^2 = \frac{1}{n-1}\sum\left(X_i - \bar{X}\right)^2 = \frac{1}{19}\left(\sum X_i^2 - 20\bar{X}^2\right) = 8{,}61, \qquad \Longrightarrow S = 2{,}94,$$
$$t_{n-1;\alpha/2} = t_{19;0,005} = 1,729.$$

Como en ejercicios anteriores si queremos más precisión para estimar μ entonces tendremos que hacer más ensayos, de esta forma

$$error = t_{n-1;\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \simeq Z_{\alpha/2}\frac{2{,}94}{\sqrt{n}} = Z_{0,05}\frac{2{,}94}{\sqrt{n}}$$

y despejamos n. Vemos que hemos cambiado de la t-student a la normal, esto es así ya que cuando $n \nearrow$ entonces $t_{n-1;\alpha/2} \sim Z_{\alpha/2}$

Para estimar σ^2 con un nivel de confianza del 90 %

$$IC_{90\,\%}(\sigma^2) = \left(\frac{(n-1)S^2}{\chi^2_{n-1;\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}\right) = \left(\frac{19\,(8{,}61)}{30{,}144}, \frac{19\,(8{,}61)}{10{,}117}\right) = (5{,}43, 16{,}19)$$

así

$$IC_{90\,\%}\left(\sigma\right) = (2{,}33, 4{,}02)$$

tal y como queríamos hacer ver. ■

Ejercicio 3.5.10 *En una población la altura de los individuos varones sigue una* $N\left(\mu, 7, 5\right)$. *Hallar el tamaño de la muestra para estimar* μ *con un error inferior a* $\pm 2cm$ *con un NC del* 0,90.

Solución. Como en los ejercicios ateriores la situación es la siguiente:

$$error \leq 2$$

por lo tanto si

$$IC_{0{,}90} = \left(\bar{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

entonces

$$error = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq 2 \qquad \Longleftrightarrow \qquad n \geq 38$$

tal y como queríamos hacer ver. ■

Ejercicio 3.5.11 *Se intenta estudiar la influencia de la hipertensión en los padres sobre la presión sanguínea de los hijos. Para ello se seleccionan dos grupos de niños, uno con padres de presión sanguínea normal (G1) y el otro con padres hipertensos (G2) obteniéndose las siguientes presiones sistólicas:*

$G1$	104	88	100	98	102	92	96	100	96	96
$G2$	100	102	96	106	110	110	120	112	112	90

hallar el IC para la diferencia de las medias suponiendo que las varianzas en las de los niños son iguales.

Solución. La situación es la siguiente. Disponemos de dos m.a.

$$\begin{array}{c} (X_i)_{i=1}^{n} \text{ m.a. de } X \sim N\left(\mu_1, \sigma_1^2\right) \\ (Y_i)_{i=1}^{n} \text{ m.a. de } Y \sim N\left(\mu_2, \sigma_2^2\right) \end{array}$$

donde estamos suponiendo que $\sigma_1^2 = \sigma_2^2$. Queremos estimar la diferencia de presiones medias en las poblaciones con un NC del 95 %.

¿son las poblciones independientes? podemos tener estas dos situaciones:

1.- Datos están emparejados, entonces las poblaciones NO son independeintes,

2.- Datos no emparejados, por lo tanto las poblaciones son independientes.

En este caso seobserva que las dos poblaciones son independientes y por lo tanto utilizaremos el siguiente formulón:

$$IC_{1-\alpha}\left(\mu_1-\mu_2\right)=\left(\bar{X}-\bar{Y}\pm t_{m+n-2;\alpha/2}S_p\sqrt{\frac{1}{m}+\frac{1}{n}}\right)$$

donde

$$S_p^2=\frac{(m-1)\,S_1^2+(n-1)\,S_2^2}{m+n-2}=\frac{9\cdot(22{,}4)+9\cdot(78{,}6)}{18}=50{,}51$$

es la varianza residual. De esta forma

$$IC_{1-\alpha}\left(\mu_1-\mu_2\right)=\left(97{,}2-105{,}8\pm(2{,}101)\sqrt{50{,}51}\sqrt{\frac{1}{10}+\frac{1}{10}}\right)=(-15{,}28,-1{,}92)\,.$$

Si las muestras no son independientes entonces lo que habría que hacer sería lo siguiente. Considerar

$$\left(X_i-Y_i\right)_{i=1}^n\quad\text{m.a. de } D=(X-Y)\sim N\left(\mu_1-\mu_2,\sigma^2\right)$$

y el formulón a aplicar en esta ocasión sería:

$$IC_{1-\alpha}\left(\mu_1-\mu_2\right)=\left(\left(\overline{X-Y}\right)\pm t_{n-1;\alpha/2}\frac{S_{dif}}{\sqrt{n}}\right)$$

donde $\left(\overline{X-Y}\right)$ representa la media de las diferencias. ■

Capítulo 4

Contraste de hipótesis

4.1. Introducción

Disponemos de $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. Como siempre queremos estimar θ.Queremos elegir entre dos posibilidades

$$\text{Hipótesis nula: } H_0 : \theta \in \Theta_0 \subset \Theta,$$
$$\text{Hipótesis alternativa : } H_1 : \theta \in \Theta_1 = \Theta_0^c \subset \Theta,$$

Definición 4.1.1 *Un test de hipótesis para contrastar $H_0 : \theta \in \Theta_0$, frente a $H_1 : \theta \in \Theta_1$ consiste en dividir el espacio muestral en dos regiones complementarias:*

$\mathcal{R}$: *región de rechazo de H_0 (**Región Crítica**), si $(x_1, ..., x_n) \in \mathcal{R}$, entonces rechazamos H_0.*

$\mathcal{A}$: *region de aceptación de H_0, si $(x_1, ..., x_n) \in \mathcal{A}$, entonces acpetamos H_0.*

Errores que se pueden cometer:

1. **Error de tipo I**: Rechazar la hipótesis nula, H_0, cuando $\theta \in \Theta_0$,
2. **Error de tipo II**: Aceptar H_0, cuando $\theta \in \Theta_1$,

Se suelen elegir H_0, H_1 de modo que el error de tipo I sea el más serio.

Definición 4.1.2 *La **función de potencia** de un test con región crítica $\mathcal{R}$ es la siguiente función de θ :*

$$\beta_{\mathcal{R}}(\theta) = P_\theta(\mathcal{R})$$

i.e. la probabilidad de rechazar H_0 cuando el verdadero valor del parámetro es θ.

Situación ideal.

$$P\,(\text{error de tipo I}) = 0,$$
$$P\,(\text{error de tipo II}) = 0,$$

Definición 4.1.3 *El nivel de significación o tamaño de un test $\mathcal{R}$ para contrastar $H_0 : \theta \in \Theta_0$ frente a $H_1 : \theta \in \Theta_1$ es:*

$$\alpha = \sup_{\theta \in \Theta_0} P_\theta (\mathcal{R})$$

i.e. máxima probabilidad de rechazar H_0 cuando H_0 es cierta o lo que es lo mismo, la máxima probabilidad de cometer un error de tipo I.

Se suelen utilizar test de hipótesis que tienen un nivel de significación, $0{,}01, 0{,}05, 0{,}1$.

Veremos dos métodos de construcción de un test:

1. Método de la razón de verosimilitud,
2. Método bayesiano.

4.2. Método de la razón de verosimilitud.

El estadístico utilizado en este método es:

$$\lambda (x_1, ..., x_n) = \frac{\sup_{\theta \in \Theta_0} f (x_1, ..., x_n)}{\sup_{\theta \in \Theta} f (x_1, ..., x_n)} = \frac{\sup_{\theta \in \Theta_0} L (\theta)}{\sup_{\theta \in \Theta} L (\theta)}$$

i.e. el cociente entre la máxima verosimilitud que da a la muestra y la máxima verosimilitud de la muestra. Vemos que $\lambda (x_1, ..., x_n) \in [0, 1]$ y que por lo tanto

$$\begin{aligned} &\lambda (x_1, ..., x_n) \text{ está próximo a cero, entonces rechazamos } H_0, \\ &\lambda (x_1, ..., x_n) \text{ está próximo a uno, entonces aceptamos } H_0, \end{aligned}$$

Definición 4.2.1 *El test de razón de verosimilitud para contrastar $H_0 : \theta \in \Theta_0$ frente a $H_1 : \theta \in \Theta_1$ al nivel de significación α es:*

$$\mathcal{R} = \{(x_1, ..., x_n) : \lambda (x_1, ..., x_n) \leq c\}$$

donde c se determina a partir de $\alpha = \sup_{\theta \in \Theta_0} P_\theta (\mathcal{R})$.

Proposición 4.2.1 *Relación con los estadísticos sufucientes. El test de razón de verosimilitudes es función de cualquier estadístico suficiente T*

Consideramos a continuación un ejemplo que nos permite ver la relación entre el contraste de hipótesis y los intervalos de confianza.

Considermos $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. En este caso $f_\theta(x) := N(\mu, \sigma^2)$ donde estamos interesados en estimar μ (aquí σ^2 es un parámetro perturbador). Consideramos el siguiente contraste de hipótesis

$$\begin{aligned} H_0 &: \mu = \mu_0, \\ H_1 &: \mu \neq \mu_0, \end{aligned}$$

y fijamos un nivel de significación α. Formalmente la situación es la siguiente:

$$\Theta = \left\{\theta = \left(\mu, \sigma^2\right) : \mu \in \mathbb{R}, \sigma > 0\right\}$$
$$\Theta_0 = \left\{\theta = \left(\mu, \sigma^2\right) : \mu = \mu_0, \sigma > 0\right\}$$

así que

$$\lambda\left(x_1, ..., x_n\right) = \frac{\sup\limits_{\theta\in\Theta_0} f\left(x_1, ..., x_n\right)}{\sup\limits_{\theta\in\Theta} f\left(x_1, ..., x_n\right)} = \frac{\sup\limits_{\theta\in\Theta_0} L\left(\theta\right)}{\sup\limits_{\theta\in\Theta} L\left(\theta\right)} \tag{4.1}$$

queda en este caso como sigue:

$$f_\theta(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}\left(x-\mu\right)^2\right),$$
$$L\left(\theta\right) = f_\theta\left(x_1, ..., x_n\right) = \prod_{i=1}^{n} f_\theta\left(x_i\right) = \frac{1}{\sigma^n\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i-\mu\right)^2\right)$$

el $\sup\limits_{\theta\in\Theta} L\left(\theta\right)$ se alcanza en el EMV

$$\hat{\mu} = \bar{x}, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$$

(aquí estamos suponiendo que conocemos $\mu = \mu_0$ por lo que queremos estimar $\hat{\sigma}^2$, de esta forma (4.1) queda:

$$\lambda\left(x_1, ..., x_n\right) = \frac{\frac{1}{\left(\frac{1}{n}\sum_{i=1}^n (x_i-\mu_0)^2\right)^{n/2}\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{n}{2\sum_{i=1}^n (x_i-\mu_0)^2}\sum_{i=1}^n\left(x_i-\mu_0\right)^2\right)}{\frac{1}{\left(\frac{1}{n}\sum_{i=1}^n (x_i-\bar{x})^2\right)^{n/2}\left(\sqrt{2\pi}\right)^n} \exp\left(-\frac{n}{2\sum_{i=1}^n (x_i-\bar{x})^2}\sum_{i=1}^n\left(x_i-\bar{x}\right)^2\right)}$$

simplificamos

$$\lambda\left(x_1, ..., x_n\right) = \left(\frac{\sum_{i=1}^n\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^n\left(x_i-\mu_0\right)^2}\right)^{n/2} = \left(\frac{\sum_{i=1}^n\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^n\left(x_i-\bar{x}+\bar{x}-\mu_0\right)^2}\right)^{n/2} =$$
$$= \left(\frac{\sum_{i=1}^n\left(x_i-\bar{x}\right)^2}{\sum_{i=1}^n\left(x_i-\bar{x}\right)^2 + n\left(\bar{x}-\mu_0\right)^2}\right)^{n/2} = \left(\frac{\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{\sum_{i=1}^n (x_i-\bar{x})^2}}{\frac{\sum_{i=1}^n (x_i-\bar{x})^2 + n(\bar{x}-\mu_0)^2}{\sum_{i=1}^n (x_i-\bar{x})^2}}\right)^{n/2} =$$
$$= \left(\frac{1}{1+n\frac{(\bar{x}-\mu_0)^2}{\sum_{i=1}^n (x_i-\bar{x})^2}}\right)^{n/2} = \left(\frac{1}{1+n\frac{(\bar{x}-\mu_0)^2}{(n-1)S^2}}\right)^{n/2} = \left(\frac{1}{1+\frac{1}{n-1}\left(\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right)^2}\right)^{n/2}$$

por lo tanto llegamos a que

$$\mathcal{R} = \left\{\left(x_1, ..., x_n\right) : \lambda\left(x_1, ..., x_n\right) \leq c\right\}$$

i.e.

$$\mathcal{R} = \left\{\left(x_1, ..., x_n\right) : \left(\frac{1}{1+\frac{1}{n-1}\left(\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right)^2}\right)^{n/2} \leq c\right\} = \left\{\left(x_1, ..., x_n\right) : \left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right| > c_1\right\} \tag{4.2}$$

y $\alpha = \sup_{\theta\in\Theta_0} P_\theta(\mathcal{R})$. En esta última ecuación (4.2) llegamos a ese resultado ya que

$$\lambda(x_1,...,x_n) \text{ es función decreciente de } \left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right|$$

y que

$$\alpha = \sup_{\theta\in\Theta_0} P_\theta(\mathcal{R}) = \alpha = \sup_{\theta\in\Theta_0} P_\theta\left((x_1,...,x_n) : \left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right| > c_1\right) = \sup_{\mu=\mu_0} P_\theta((x_1,...,x_n) : |t_{n-1}| > c_1)$$

ya que cuando $(X_i)_{i=1}^n$ m.a. de $X \sim N(\mu_0,\sigma^2)$ sabemos que una cantidad para estimar μ

$$\frac{\bar{x}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

de esta forma llegamos a la conclusión de que

$$c_1 = t_{n-1;\alpha/2}$$

y que por lo tanto

$$\mathcal{R} = \left\{(x_1,...,x_n) : \left|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}\right| > t_{n-1;\alpha/2}\right\} = \left\{(x_1,...,x_n) : |\bar{x}-\mu_0| > t_{n-1;\alpha/2}S/\sqrt{n}\right\}$$

será la región de rechazo del test de razón de verosimilitud. Veamos ahora la relación con el $IC_{1-\alpha}(\mu)$. La situación hasta ahora es la siguiente:

$$\mathfrak{X} = \mathcal{R}\cup\mathcal{A}, \qquad \mathcal{A} = \left\{(x_1,...,x_n) : |\bar{x}-\mu_0| < t_{n-1;\alpha/2}S/\sqrt{n}\right\},$$

esto significa que aceptamos H_0 :

$$\begin{aligned} H_0 : \mu = \mu_0 &\iff |\bar{x}-\mu_0| < t_{n-1;\alpha/2}S/\sqrt{n} \iff -t_{n-1;\alpha/2}S/\sqrt{n} < \bar{x}-\mu_0 < t_{n-1;\alpha/2}S/\sqrt{n} \\ &\iff \bar{x} - t_{n-1;\alpha/2}S/\sqrt{n} < \mu_0 < \bar{x} + t_{n-1;\alpha/2}S/\sqrt{n} \iff \mu_0 \in \left(\bar{x} \pm t_{n-1;\alpha/2}S/\sqrt{n}\right) \\ &\iff \mu_0 \in IC_{1-\alpha}(\mu). \end{aligned}$$

De forma alternativa vemos que

$$\begin{aligned} H_0 &: \mu = \mu_0, \\ H_1 &: \mu \neq \mu_0, \end{aligned}$$

y fijamos un nivel de significación α. Entonces $IC_{1-\alpha}(\mu). = \left(\bar{x} \pm t_{n-1;\alpha/2}S/\sqrt{n}\right)$, por lo tanto si

$$\begin{array}{lcl} \mu_0 \in IC_{1-\alpha}(\mu) & \Longrightarrow & \text{aceptamos } H_0, \\ \mu_0 \notin IC_{1-\alpha}(\mu) & \Longrightarrow & \text{rechazamos } H_0 \end{array}$$

de esta forma hemos visto la relación existente entre el contraste de hipótesis y los intervalos de confianza.

Si la distribución de $\lambda(x_1,...,x_n)$, bajo H_0, es desconocida o muy complicada de calcular entonces existe una solución aproximada (asintótica) para contrastar $H_0 : \theta \in \Theta_0$, frente a $H_1 : \theta \in \Theta_1$.

Teorema 4.2.1 *Sea $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$, y $\dim \Theta = k$. Queremos contrastar $H_0 : \theta \in \Theta_0$, frente a $H_1 : \theta \in \Theta_1$, donde $\dim \Theta_0 = k' < k$, entonces bajo ciertas condiciones de regularidad*

$$-2 \log \lambda (x_1, ..., x_n) \xrightarrow{d} \chi^2_{k-k'}$$

bajo H_0.

De esta forma tenemos que

$$\mathcal{R} = \left\{(x_1, ..., x_n) : -2 \log \lambda (x_1, ..., x_n) > \chi^2_{k-k';\alpha}\right\}$$

Con el ejemplo que venimos arrastrando, si $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. En este caso $f_\theta(x) := N(\mu, \sigma^2)$ donde estamos interesados en estimar μ (aquí σ^2 es un parámetro perturbador). Consideramos el siguiente contraste de hipótesis

$$H_0 : \mu = \mu_0,$$
$$H_1 : \mu \neq \mu_0,$$

y fijamos un nivel de significación α. Formalmente la situación es la siguiente:

$$\Theta = \left\{\theta = (\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma > 0\right\} \quad \Longrightarrow \quad \dim \Theta = 2$$
$$\Theta_0 = \left\{\theta = (\mu, \sigma^2) : \mu = \mu_0, \sigma > 0\right\} \quad \Longrightarrow \quad \dim \Theta_0 = 1$$

así que

$$-2 \log \lambda (x_1, ..., x_n) \xrightarrow{d} \chi^2_1.$$

4.2.1. Asignación de hipótesis en la práctica.

1. Disponemos de $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. ¿Es razonable aceptar que $\theta = \theta_0$?. Elegimos entre $\theta = \theta_0$ y $\theta \neq \theta_0$

   La situación es por lo tanto la siguiente:

   $$H_0 : \theta = \theta_0,$$
   $$H_1 : \theta \neq \theta_0,$$

   y fijamos un nivel de significación α. Los papeles, en la práctica, de hipótesis nula y alternativa siempre se asignan igual, la igualdad para H_0 y la desigualdad para H_1. Por lo tanto

   si aceptamos H_0 concluimos que es razomable aceptar $\theta = \theta_0$,
   si rechazamos H_0 concluimos que NO es razomable aceptar $\theta = \theta_0$.

2. Disponemos de $(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. ¿Podemos considerar estadísticamente probado que $\theta > \theta_0$?

   La situación es por lo tanto la siguiente:

   $$H_0 : \theta \leq \theta_0,$$
   $$H_1 : \theta > \theta_0,$$

   y fijamos un nivel de significación α. Por lo tanto

   si aceptamos H_0 concluimos que No es razomable aceptar $\theta > \theta_0$,
   si rechazamos H_0 concluimos que es razomable aceptar $\theta > \theta_0$.

4.2.2. Concepto de p-valor de los datos.

El **p-valor** o significación de los datos es la máxima probabilidad bajo H_0 de obtener un resultado más desfavorable a H_0 que el obtenido.

Veamos un ejemplillo.

Ejemplo 4.2.1 *$(X_i)_{i=1}^n$ m.a. de $X \sim f_\theta(x)$ tal que $\theta \in \Theta$. En este caso $f_\theta(x) := N\left(\mu, \sigma^2 = 1\right)$. Consideramos el siguiente contraste de hipótesis*

$$H_0 : \mu \leq 10,$$
$$H_1 : \mu > 10,$$

y fijamos un nivel de significación α. Hemos obtenido la muestra con $n = 9$ y resulta que $\bar{x} = 12$

$$p - valor = \max_{\mu \leq 10} P\left\{\text{más desfavorable a } H_0 \text{ que } \bar{x} = 12\right\} = \max_{\mu \leq 10} P\left\{\ \bar{x} = 12\right\} = P_{\mu=10}\left\{\ \bar{x} = 12\right\}$$

pero sabemos que

$$\bar{X} \sim N\left(\mu, \frac{1}{9}\right)$$

así que

$$P\left\{\frac{\bar{x} - 10}{1/3} > \frac{12 - 10}{1/3}\right\} = P\left\{Z > \frac{12 - 10}{1/3}\right\} \approx 0.$$

Interpretmos el p-valor como el apoyo de los datos a H_0

1. p-valor próximo a cero, entonces poco apoyo de los datos a H_0
2. p-valor alejado del cero, entonces apoyo suficiente de los datos a H_0

La regla práctica que se sigue para tomar decisiones a un nivel de significación α previamente fijado es la siguiente:

1. $\mathbf{p - valor} < \boldsymbol{\alpha}$, entonces poco apoyo de los datos a H_0, **rechazamos** H_0,
2. $\mathbf{p - valor} > \boldsymbol{\alpha}$, entonces apoyo suficiente de los datos a H_0, **aceptamos** H_0.

4.3. Métodos bayesianos.

Sea $(X_i)_{i=1}^n$ m.a. de $X \sim f(X \mid \theta)$ donde $\theta \in \Theta$. Queremos contrastar $H_0 : \theta \in \Theta_0$, frente a $H_1 : \theta \in \Theta_1 = \Theta_0^c$. Tenemos la información muestral (verosimilitud)

$$f(x_1, ..., x_n \mid \theta) = \prod_{i=1}^n f(x_i \mid \theta)$$

que recoje la información sobre θ que llevan los datos y también disponemos de la información a priori

$$\pi(\theta)$$

que recoje la información sobre θ antes que los datos, así que aplicando Bayes obtenemos

$$\pi(\theta \mid x_1, ..., x_n) = \frac{\pi(\theta) \prod_{i=1}^n f(x_i \mid \theta)}{\int_\theta \pi(\theta) \prod_{i=1}^n f(x_i \mid \theta)} \propto \pi(\theta) f(x_1, ..., x_n \mid \theta)$$

y todas las conclusiones las obtendremos a partir de $\pi(\theta \mid x_1, ..., x_n)$ la distribución a posteriori.

La decisión de rechazar o aceptar H_0 estará basada en;

$$\begin{aligned} P(\theta_0 \mid x_1, ..., x_n) &= \int_{\theta_0} \pi(\theta \mid x_1, ..., x_n) \, d\theta, \\ P(\theta_1 \mid x_1, ..., x_n) &= \int_{\theta_1} \pi(\theta \mid x_1, ..., x_n) \, d\theta = 1 - P(\theta_0 \mid x_1, ..., x_n), \end{aligned}$$

por ejemplo podemos rechazar $H_0 : \theta \in \Theta_0$ si

$$P(\theta_0 \mid x_1, ..., x_n) < \frac{1}{2}.$$

4.4. Resumen de contrastes de hipótesis

CONTRASTES DE HIPÓTESIS

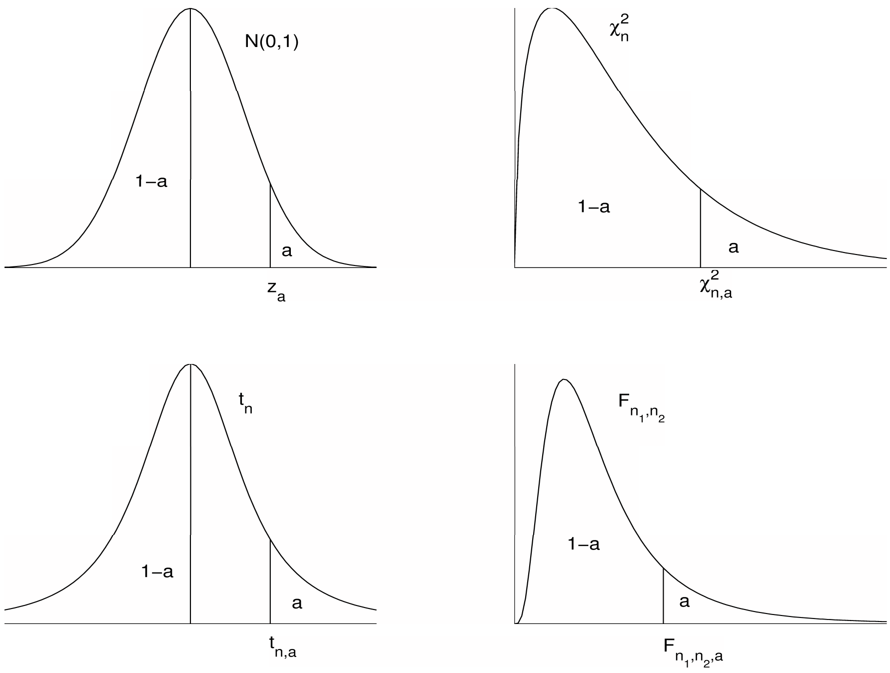
NOTACIÓN:

α= nivel de significación del contraste.

n= tamaño de la muestra.

H_0= hipótesis nula.

R= región crítica o de rechazo de H_0.



4.4.1. $X \sim N(\mu, \sigma)$

1. $H_0 : \mu = \mu_0$ (σ conocida);
$$R = \left\{ |\bar{x} - \mu_0| > z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right\}.$$

2. $H_0 : \mu = \mu_0$ (σ desconocida);
$$R = \left\{ |\bar{x} - \mu_0| > t_{n-1;\alpha/2} \frac{s}{\sqrt{n}} \right\}.$$

3. $H_0 : \mu \leq \mu_0$ (σ conocida);
$$R = \left\{ \bar{x} - \mu_0 > z_{\alpha} \frac{\sigma}{\sqrt{n}} \right\}.$$

4. $H_0 : \mu \leq \mu_0$ (σ desconocida);
$$R = \left\{ \bar{x} - \mu_0 > t_{n-1;\alpha} \frac{s}{\sqrt{n}} \right\}.$$

5. $H_0 : \mu \geq \mu_0$ (σ conocida);
$$R = \left\{ \bar{x} - \mu_0 < z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \right\}.$$

6. $H_0 : \mu \geq \mu_0$ (σ desconocida);
$$R = \left\{ \bar{x} - \mu_0 < t_{n-1;1-\alpha} \frac{s}{\sqrt{n}} \right\}.$$

7. $H_0 : \sigma = \sigma_0$;
$$R = \left\{ \frac{n-1}{\sigma_0^2} s^2 \notin \left(\chi^2_{n-1;1-\alpha/2} \ , \ \chi^2_{n-1;\alpha/2} \right) \right\}.$$

8. $H_0 : \sigma \leq \sigma_0$;
$$R = \left\{ \frac{n-1}{\sigma_0^2} s^2 > \chi^2_{n-1;\alpha} \right\}.$$

9. $H_0 : \sigma \geq \sigma_0$;
$$R = \left\{ \frac{n-1}{\sigma_0^2} s^2 < \chi^2_{n-1;1-\alpha} \right\}.$$

4.4.2. $X \sim$ Bernoulli(p) (muestras grandes)

1. $H_0 : p = p_0$;
$$R = \left\{ |\bar{x} - p_0| > z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}.$$

2. $H_0 : p \leq p_0$;
$$R = \left\{ \bar{x} - p_0 > z_{\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}.$$

3. $H_0 : p \geq p_0$;
$$R = \left\{ \bar{x} - p_0 < z_{1-\alpha} \sqrt{\frac{p_0(1-p_0)}{n}} \right\}.$$

4.4.3. $X \sim$ Poisson(λ) (muestras grandes)

1. $H_0 : \lambda = \lambda_0$;
$$R = \left\{ |\bar{x} - \lambda_0| > z_{\alpha/2} \sqrt{\lambda_0/n} \right\}.$$

2. $H_0 : \lambda \leq \lambda_0$;
$$R = \left\{ \bar{x} - \lambda_0 > z_{\alpha} \sqrt{\lambda_0/n} \right\}.$$

3. $H_0 : \lambda \geq \lambda_0$;
$$R = \left\{ \bar{x} - \lambda_0 < z_{1-\alpha} \sqrt{\lambda_0/n} \right\}.$$

4.4.4. Dos poblaciones Normales independientes

$X \sim N(\mu_1, \sigma_1)$; $(X_1, \ldots, X_{n_1})$ m. a. de X; se calcula $\bar{x}$ y s_1^2.

$Y \sim N(\mu_2, \sigma_2)$; $(Y_1, \ldots, Y_{n_2})$ m. a. de Y; se calcula $\bar{y}$ y s_2^2.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

1. $H_0 : \mu_1 = \mu_2$ (σ_1, σ_2 conocidas);
$$R = \left\{ |\bar{x} - \bar{y}| > z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}.$$

2. $H_0 : \mu_1 = \mu_2$ ($\sigma_1 = \sigma_2$);
$$R = \left\{ |\bar{x} - \bar{y}| > t_{n_1+n_2-2;\alpha/2} \; s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}.$$

3. $H_0 : \mu_1 = \mu_2$ ($\sigma_1 \neq \sigma_2$);
$$R = \left\{ |\bar{x} - \bar{y}| > t_{f;\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}.$$

4. $H_0 : \mu_1 \leq \mu_2$ (σ_1, σ_2 conocidas);
$$R = \left\{ \bar{x} - \bar{y} > z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}.$$

5. $H_0 : \mu_1 \leq \mu_2$ ($\sigma_1 = \sigma_2$);
$$R = \left\{ \bar{x} - \bar{y} > t_{n_1+n_2-2;\alpha} \; s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}.$$

6. $H_0 : \mu_1 \leq \mu_2$ ($\sigma_1 \neq \sigma_2$);
$$R = \left\{ \bar{x} - \bar{y} > t_{f;\alpha} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}.$$

7. $H_0 : \mu_1 \geq \mu_2$ (σ_1, σ_2 conocidas);
$$R = \left\{ \bar{x} - \bar{y} < z_{1-\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \right\}.$$

8. $H_0 : \mu_1 \geq \mu_2$ ($\sigma_1 = \sigma_2$);
$$R = \left\{ \bar{x} - \bar{y} < t_{n_1+n_2-2;1-\alpha} \; s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right\}.$$

9. $H_0 : \mu_1 \geq \mu_2$ ($\sigma_1 \neq \sigma_2$);
$$R = \left\{ \bar{x} - \bar{y} < t_{f;1-\alpha} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \right\}.$$

10. $H_0 : \sigma_1 = \sigma_2$;
$$R = \left\{ s_1^2/s_2^2 \notin \left(F_{n_1-1;n_2-1;1-\alpha/2} \; , \; F_{n_1-1;n_2-1;\alpha/2} \right) \right\}.$$

11. $H_0 : \sigma_1 \leq \sigma_2$;
$$R = \left\{ s_1^2/s_2^2 > F_{n_1-1;n_2-1;\alpha} \right\}.$$

12. $H_0 : \sigma_1 \geq \sigma_2$;
$$R = \left\{ s_1^2/s_2^2 < F_{n_1-1;n_2-1;1-\alpha} \right\},$$

$$\text{donde } f = \text{entero más próximo a } \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}.$$

4.4.5. Comparación de proporciones (muestras grandes e independientes)

$X \sim$ Bernoulli(p_1); $(X_1, \ldots, X_{n_1})$ m. a. de X.

$Y \sim$ Bernoulli(p_2); $(Y_1, \ldots, Y_{n_2})$ m. a. de Y.

1. $H_0 : p_1 = p_2$;
$$R = \left\{ |\bar{x} - \bar{y}| > z_{\alpha/2} \sqrt{\bar{p}(1-\bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right\}.$$

2. $H_0 : p_1 \leq p_2$;
$$R = \left\{ \bar{x} - \bar{y} > z_{\alpha} \sqrt{\bar{p}(1-\bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right\}.$$

3. $H_0 : p_1 \geq p_2$;
$$R = \left\{ \bar{x} - \bar{y} < z_{1-\alpha} \sqrt{\bar{p}(1-\bar{p}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \right\},$$

$$\text{donde } \bar{p} = \frac{\sum x_i + \sum y_i}{n_1 + n_2} = \frac{n_1\bar{x} + n_2\bar{y}}{n_1 + n_2}.$$

4.5. Ejercicios

Ejercicio 4.5.1 *Sea* (X_1, X_2) *una m.a. de*

$$f_\theta = \begin{cases} \theta x^{\theta-1} & 0 < x < 1 \\ 0 & \textit{resto} \end{cases}$$

donde $\theta \in \Theta = \{1, 2\}$. *Para contrastar* $H_0 : \theta = 1$ *frente a* $H_1 : \theta = 2$, *definimos la región de rechazo*

$$\mathcal{R} = \{(x_1, x_2) : 4x_1x_2 \geq 3\}.$$

Hallar la función de potencia test.

Solución. Queremos contrastar $H_0 : \theta = 1$ frente a $H_1 : \theta = 2$, en $\mathcal{R} = \{(x_1, x_2) : 4x_1x_2 \geq 3\}$, por lo tanto nos preguntamos si

$$\beta_{\mathcal{R}}(\theta) = P_\theta(\mathcal{R}), \qquad \theta = 1, 2.$$

Represetamos la región, $4x_1x_2 \geq 3$, por lo tanto cuando $\theta = 1$, tenemos:

$$f_{\theta=1}(x_1, x_2) = f_{\theta=1}(x_1) f_{\theta=1}(x_2) = 1 \cdot 1 = 1$$

de esta forma

$$P_{\theta=1}(\mathcal{R}) = \int\int_{\mathcal{R}} f_{\theta=1}(x_1, x_2) \, dx_1 dx_2 = \int_{x_1=3/4}^{x_1=1} \left(\int_{x_2=3/4x_1}^{x_2=1} 1 dx_2 \right) dx_1 = \frac{1}{4} + \frac{3}{4} \log\left(\frac{3}{4}\right) \approx 0{,}03$$

y hacemos lo mismo para $\theta = 2$, obreniendo así:

$$P_{\theta=1}(\mathcal{R}) = \int_{x_1=3/4}^{x_1=1} \left(\int_{x_2=3/4x_1}^{x_2=1} 4x_1x_2 dx_2 \right) dx_1 = \frac{7}{16} + \frac{9}{8} \log\left(\frac{3}{4}\right) \approx 0{,}11.$$

ya que

$$f_{\theta=2}(x_1, x_2) = f_{\theta=2}(x_1) f_{\theta=2}(x_2) = 2x_1 \cdot 2x_2 = 4x_1x_2.$$

El nivel de significación es por lo tanto

$$\sup_{\theta \in \theta_0} P_\theta(\mathcal{R}) = P_{\theta=1}(\mathcal{R}) \approx 0{,}03$$

tal y como queríamos hacer ver. ■

Ejercicio 4.5.2 *Sea $(X_i)_{i=1}^n$ una m.a. de una población $N(\mu, 1)$, donde $\mu \in \Theta = \{\mu_0, \mu_1\}$ con $(\mu_0 < \mu_1)$. Para contrastar $H_0 : \mu = \mu_0$ frente a $H_1 : \mu = \mu_1$ se considera el test que rechaza H_0 si $\bar{x} > k$. Hallar k para que el test tenga nivel de significación α.*

Solución. La situación es la siguiente:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu = \mu_1 \\ \mathcal{R} &= \{(x_i)_{i=1}^n : \bar{x} > k\} \end{aligned}$$

queremos calcular k para un nivel de significación α. Recordamos de teoría que

$$\alpha = \sup_{\mu \in \theta_0} P_\theta(\mathcal{R}) = P_{\mu_1}(\mathcal{R}) = P_{\mu_1}(\bar{x} > k) = P\left\{ \frac{\bar{x} - \mu_1}{\sqrt{\frac{1}{n}}} > \frac{k - \mu_1}{\sqrt{\frac{1}{n}}} \right\} = P\left(Z > \frac{k - \mu_1}{\sqrt{1/n}} \right)$$

recordar que

$$X \sim N\left(\mu_1, \frac{1}{n}\right)$$

por lo tanto

$$\frac{k-\mu_1}{\sqrt{1/n}} = Z_\alpha$$

obteniendo así

$$k = \mu_1 + \frac{Z_\alpha}{\sqrt{n}}$$

tal y como queríamos hacer ver. ■

Ejercicio 4.5.3 *Sea $(X_i)_{i=1}^{12}$ una m.a. de una población que sigue una $Poisson\,(\theta)$, donde $\theta \in \Theta = (1, 1/2]$. Si la región de rechazo, para contrastar $H_0 : \theta = 1/2$ frente a $H_1 : \theta < 1/2$, definimos la región de rechazo*

$$\mathcal{R} = \left\{ (x_i)_{i=1}^{12} : \sum x_1 \leq 2 \right\}.$$

Hallar la potencia del test en $\theta = 1/2$, $\theta = 1/4$, $\theta = 1/12$.

Solución. La situación es la siguiente: $(X_i)_{i=1}^{12}$ una m.a. $X \sim Poisson\,(\theta)$

$$\begin{aligned} &H_0 : \theta = 1/2, \\ &H_1 : \theta < 1/2, \\ &\mathcal{R} = \left\{ (x_i)_{i=1}^{12} : \sum_i^{12} x_1 \leq 2 \right\}. \end{aligned}$$

En $\theta = 1/2$

$$\beta_{\mathcal{R}}\,(\theta) = P_\theta\,(\mathcal{R}) = P_{\theta=1/2}\,(\mathcal{R}) = P_{\theta=1/2}\left(\sum_i^{12} x_1 \leq 2 \right) = P\left\{ Poisson\,(\lambda = 6) \leq 2 \right\} = *$$

ya que

$$\sum_i^{12} x_1 = \left(\frac{1}{2} + \overset{12-veces}{\cdots\cdots} + \frac{1}{2} \right) \leq 2 \sim Poisson\,(\lambda = 6)\,, \qquad \varphi_{X_i} = e^{\theta\left(e^{it}-1\right)}$$

y por lo tanto (mirando las tablas de la Poisson)

$$* = 0{,}0025 + 0{,}0149 + 0{,}0446 = 0{,}0620 \simeq 0{,}06$$

Si hacemos lo mismo para los otros dos casos obtenemos:

$$\begin{aligned} P_{\theta=1/4}\,(\mathcal{R}) &= ...... = 0{,}4232, \\ P_{\theta=1/12}\,(\mathcal{R}) &= ...... = 0{,}9197, \end{aligned}$$

tal y como querímos hacer ver. ■

Ejercicio 4.5.4 *En una piscifactoría se desea contrastar la hipótesis* (H_0) *de que el procentaje de adultos que miden menos de* $20cm$ *es como máximo el* $10\,\%$*. Para ello se va a tomar una muestra de* 6 *peces y rechazamos* H_0 *si encontramos más de un pez con longitud inderior a* $20cm$*.*

1. *¿Cuál es el nivel de significación de este contraste?.*
2. *Calcular la potencia del contraste si en realidad hay un* $20\,\%$ *de peces que miden menos de* $20cm$*.*

Solución. Modelizamos mediante una $Bernoulli(p)$ i.e. disponemos de una m.a. $(X_i)_{i=1}^6$ donde $X \sim Bernoulli(p)$

$$X = \begin{cases} 1 & \text{si el pez mide } < 20cm \\ 0 & \text{si el pez mide } > 20cm \end{cases}$$

así que

$$\begin{aligned} &H_0 : p \leq 0{,}10, \\ &H_1 : p > 0{,}10, \\ &\mathcal{R} = \{\text{Más de un pez con longitud inferior a 20 cm}\} = \{T > 1\}\,. \end{aligned}$$

el estadístico que estamos usando en este contraste es:

$$T = \text{ n}^o \text{ de peces, de entre los 6, que miden de 20cm} \sim Bin(n = 6, p = 0{,}10).$$

Por lo tanto, el nivel de significación de este contraste es:

$$\begin{aligned} \sup_{p \leq 0{,}10} P_p(\mathcal{R}) &= P_{p=0{,}10}(\mathcal{R}) = P_{p=0{,}10}(T > 1) = P\left(Bin(n = 6, p = 0{,}10) > 1\right) = \\ &= 1 - P\left(Bin(n = 6, p = 0{,}10) \leq 1\right) = 1 - 0{,}5314 - 0{,}3543 = 0{,}114. \end{aligned}$$

Con respecto a la segunda pregunta, i.e. Calcular la potencia del contraste si en realidad hay un $20\,\%$ de peces que miden menos de $20cm$. vemos que

$$P_{p=0{,}20}(T > 1) = P\left(Bin(n = 6, p = 0{,}20) > 1\right) = ... = 0{,}3447$$

tal y como queríamos hacer ver. ■

Ejercicio 4.5.5 *Para estudiar si una prueba de laboratorio puede resultar demasiado nociva para la salud, contrastamos la hipótesis* H_0 *de que la probabilidad de que una persona sometida a esa prueba resulte afectada sea como máximo* $0,001$*. Para ello sometemos a esa prueba a* 1000 *personas elegidas al azar y aceptamos* H_0 *si como máximo ha habido un afectado.*

1. *¿Cuál es el nivel de significación de este contraste?.*
2. *Si en realidad la prueba afecta a la salud de una persona con probabilidad* $0,003$ *¿cuál es la probabilidad deaceptar* H_0*?.*

Solución. En este ejercicio disponemos de una m.a. $(X_i)_{i=1}^{1000}$ donde $X \sim Bernoulli(p)$

$$X = \begin{cases} 1 & \text{si resulta afectada} \\ 0 & \text{si no resulta afectada} \end{cases}$$

así que

$$\begin{aligned} &H_0 : p \leq 0{,}001, \\ &H_1 : p > 0{,}001, \\ &\mathcal{R} = \{\text{Hay como máximo un afectado entre } 1000\} = \{T \leq 1\}. \end{aligned}$$

el estadístico que estamos usando en este contraste es:

$$T = \text{ n}^o \text{ de afectados entre } 1000 \sim Bin(n = 1000, p = 0{,}001),$$

al ser $n = 1000$ y $p = 0{,}001$ entonces aproximamos la binomial mediante una Poisson

$$Bin(n = 1000, p = 0{,}001) \sim Poisson\left(\lambda = np\right).$$

De esta forma encontramos que el nivel de significación de este contraste es:

$$\begin{aligned} \sup_{p \leq 0{,}001} P_p\left(\mathcal{R}\right) &= P_{p=0{,}001}\left(\mathcal{R}\right) = P_{p=0{,}001}\left(T > 1\right) = P\left(Bin(n = 1000, p = 0{,}001)\right) > 1) = \\ &= 1 - P\left(Poisson\left(\lambda = np = 1\right)\right) \leq 1) = 0{,}2642. \end{aligned}$$

Con respecto a la segunda pregunta, i.e. Si en realidad la prueba afecta a la salud de una persona con probabilidad $0,003$ ¿cuál es la probabilidad de aceptar H_0?. Vemos que

$$P\left(Poisson\left(\lambda = np = 3\right)\right) \leq 1) = 0{,}1992$$

tal y como queríamos hacer ver. ∎

Ejercicio 4.5.6 *Sea $(X_i)_{i=1}^{n}$ una m.a. de una población con una densidad*

$$f_\theta = \begin{cases} e^{-(x-\theta)} & x > \theta \in \mathbb{R} \\ 0 & resto \end{cases}$$

Hallar el test de razón de verosimilitud, de nivel α, para contrastar $H_0 : \theta \leq \theta_0$ frente a $H_1 : \theta > \theta_0$.

Solución. Tenemos

$$\begin{aligned} &H_0 : \theta \leq \theta_0, \\ &H_1 : \theta > \theta_0, \\ &\quad \text{nivel de significación } \alpha \end{aligned}$$

queremos calcular c

$$\mathcal{R} = \left\{(x_i)_{i=1}^{n} : \lambda\left(x_1, ..., x_n\right) \leq c\right\}$$

tal que

$$\alpha = \sup_{\theta \leq \theta_0} P_\theta(\mathcal{R})$$

de esta forma vemos que

$$\lambda(x_1, ..., x_n) = \frac{\sup_{\theta \leq \theta_0} f(x_1, ..., x_n)}{\sup_{\theta \in \mathbb{R}} f(x_1, ..., x_n)} = \frac{\sup_{\theta \leq \theta_0} L(\theta)}{\sup_{\theta \in \mathbb{R}} L(\theta)} \tag{4.3}$$

donde

$$L(\theta) = f_\theta(x_1, ..., x_n) = \prod_{i=1}^{n} f_\theta(x_i) = e^{n\theta} e^{-\sum x_i}, \qquad si \quad \theta \leq X_{(1)}$$

observamos que $x \geq \theta$, significa $(x_i)_{i=1}^n \geq \theta$ y por lo tanto $\theta \leq X_{(1)}$, de este modo obtenemos la siguiente función

$$L(\theta) = \begin{cases} e^{n\theta} e^{-\sum x_i} & \theta \leq X_{(1)} \\ 0 & \theta > X_{(1)} \end{cases}$$

y por lo tanto

$$\sup_{\theta \in \mathbb{R}} L(\theta) = e^{nX_{(1)} - \sum x_i},$$

$$\sup_{\theta \leq \theta_0} L(\theta) = \begin{cases} e^{nX_{(1)} - \sum x_i} & X_{(1)} \leq \theta_0 \\ e^{n\theta_0 - \sum x_i} & X_{(1)} > \theta_0 \end{cases}$$

así que la ecuación (4.3) queda

$$\begin{cases} 1 & X_{(1)} \leq \theta_0 \\ \frac{e^{n\theta_0 - \sum x_i}}{e^{nX_{(1)} - \sum x_i}} = e^{n\theta_0 - nX_{(1)}} & X_{(1)} > \theta_0 \end{cases}$$

entonces

$$\mathcal{R} = \{(x_i)_{i=1}^n : \lambda(x_1, ..., x_n) \leq c\} = \left\{(x_i)_{i=1}^n : X_{(1)} > C_0\right\} = \tag{4.4}$$

donde

$$C_0 : \alpha = \sup_{\theta \leq \theta_0} P_\theta(\mathcal{R}) = \sup_{\theta \leq \theta_0} e^{n(\theta_0 - C_0)}$$

por lo tanto

$$C_0 = \theta_0 - \frac{\log \alpha}{n}$$

ya que

$$P_\theta(\mathcal{R}) = P_\theta\left(X_{(1)} > C_0\right) = \prod_{i=1}^{n} P_\theta(X_i > C_0) = (P_\theta(X > C_0))^n =$$

$$= \left(\int_{C_0}^{\infty} f_\theta(x)\, dx\right)^n = \left(\int_{C_0}^{\infty} e^{-(x-\theta)} dx\right)^n = e^{n(\theta_0 - C_0)}$$

así que

$$\mathcal{R} = \left\{(x_i)_{i=1}^n : X_{(1)} > \theta_0 - \frac{\log \alpha}{n}\right\}$$

tal y como queríamos hacer ver. ■

Ejercicio 4.5.7 *Sea $(X_i)_{i=1}^n$ una m.a. de una población con una densidad*

$$f_\theta = \begin{cases} \theta (1-x)^{\theta-1} & 0 \leq x \leq 1, \theta > 0 \\ 0 & resto \end{cases}$$

Deseamos contrastar $H_0 : \theta = 2$ frente a $H_1 : \theta \neq 2$, al nivel de significación 0,1. Si $n = 60$ y $\prod_i (1-x_i) = 0{,}0003$. ¿Cuál sería la decisión adoptada utilizando el test de razón de verosimilitudes asintóticamente?

Solución. En este ejercicio queremos ver

$$\mathcal{R} = \mathcal{R} = \{(x_i)_{i=1}^n : \lambda(x_1,...,x_n) \leq c\} \sim \left\{(x_1,...,x_n) : -2\log\lambda(x_1,...,x_n) > \chi^2_{k-k';\alpha} = \chi^2_{1;0,1}\right\}$$

en este caso

$$\lambda(x_1,...,x_n) = \frac{\sup_{\theta\in\Theta_0} L(\theta)}{\sup_{\theta\in\not\leq} L(\theta)} = \frac{2^n\left(\prod_i (1-x_i)\right)}{\left(\hat{\theta}\right)^n \left(\prod_i (1-x_i)\right)^{\hat{\theta}-1}}$$

ya que

$$L(\theta) = f_\theta(x_1,...,x_n) = \prod_{i=1}^n f_\theta(x_i) = (\theta)^n \left(\prod_i (1-x_i)\right)^{\theta-1},$$

$$\log L(\theta) = n\log(\theta) + (\theta-1)\log\left(\prod_i (1-x_i)\right)$$

$$\frac{d\log L(\theta)}{d\theta} = \frac{n}{\theta} + \log\left(\prod_i (1-x_i)\right) = 0 \quad \Longrightarrow \quad \hat{\theta} = -\frac{n}{\log\left(\prod_i (1-x_i)\right)}$$

vemos que (en la muestra)

$$\hat{\theta} = -\frac{n}{\log\left(\prod_i (1-x_i)\right)} = \frac{(-60)}{\ln(0{,}0003)} = 7{,}3967$$

por lo tanto

$$-2\log\lambda(x_1,...,x_n) = -2\ln\left(\frac{2^{60}(0{,}0003)}{(7{,}3967)^{60}(0{,}0003)^{6{,}3967}}\right) = 69{,}39,$$

$$\chi^2_{1;0,1} = 2{,}706$$

así que rechazamos H_0, i.e. el valor $\theta = 2$ no es aceptable. ■

Ejercicio 4.5.8 *Sea $(X_i)_{i=1}^n$ una m.a. de una población X con una función de densidad*

$$f_\theta = \begin{cases} \theta e^{-\theta x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

donde $\theta > 0$. Deseamos contrastar $H_0 : \theta \geq 1$ frente a $H_1 : \theta < 1$.

1. *Si adoptamos como región de rechazo de H_0*

$$\mathcal{R} = \left\{ (x_i)_{i=1}^n : x_{(1)} \geq c \right\}.$$

determinar el valor de c, para que el test tenga tamaño α.

2. *Si tomamos como densidad a priori para θ*

$$\pi(\theta) = \begin{cases} e^{-\theta} & \theta > 0 \\ 0 & resto \end{cases}$$

calcular la probabilidad a posteriori de la región que constituye la hipótesis nula, cuando la muestra consta de una sola observación.

Solución. Tenemos

$$\begin{aligned} &H_0 : \theta \geq 1, \\ &H_1 : \theta < 1, \\ &\quad \text{nivel de significación } \alpha \end{aligned}$$

queremos calcular c

$$\mathcal{R} = \{(x_i)_{i=1}^n : \lambda(x_1, ..., x_n) \leq c\} = \left\{ (x_i)_{i=1}^n : x_{(1)} \geq c \right\}$$

tal que

$$\alpha = \sup_{\theta \geq 1} P_\theta(\mathcal{R})$$

de esta forma vemos que si

$$P_\theta(\mathcal{R}) = P_\theta\left(x_{(1)} \geq c\right) = \left(P_\theta\left(x_{(1)} \geq c\right)\right)^n = \left(\int_c^\infty \theta e^{-\theta x} dx\right)^n = e^{-n\theta c}$$

entonces

$$\alpha = \sup_{\theta \geq 1} P_\theta(\mathcal{R}) = P_{\theta=1}(\mathcal{R}) = e^{-nc}, \qquad \Longrightarrow \qquad c = -\frac{\log \alpha}{n}.$$

Con respecto al segundo de los apartados vemos que

$$\pi(\theta) = \begin{cases} e^{-\theta} & \theta > 0 \\ 0 & resto \end{cases}$$

por lo tanto

$$\pi(\theta \mid x) \propto f(x \mid \theta) \pi(\theta) = \theta e^{-\theta x} e^{-\theta} = \theta e^{-\theta(x+1)}, \qquad \theta > 0$$

por lo que

$$\pi(\theta \mid x) \sim Gamma(p=2, a=x+1) = \frac{a^p}{\Gamma(p)}\theta^{p-1}e^{-a\theta} =$$
$$= \frac{(x+1)^2}{\Gamma(2)}\theta e^{-\theta(x+1)} = (x+1)^2\,\theta e^{-\theta(x+1)}, \qquad \theta > 0$$

encontramos así que la probabilidad a posteriori de θ_0 es:

$$\int_{\theta=1}^{\infty} (x+1)^2\,\theta e^{-\theta(x+1)} d\theta = \left[-e^{-\theta(x+1)}\,(\theta + x\theta + 1)\right]_1^{\infty} = (x+2)\,e^{-(x+1)}.$$

tal y como queríamos hacer ver. ■

Ejercicio 4.5.9 *Se van a probar dos medicamentos A y B contra una enfermedad. Para esto tratamos* 100 *ratones enfermos con A y otros* 100 *con B. El número medio de horas que sobreviven con A es* $\bar{x} = 1200$ *y para B* $\bar{y} = 1400$. *Suponiendo normalidad en mabos casos se pide:*

1. *¿Se puede aceptar igualdad de varianzas si sabemos que*

$$\sum_i (x_i - \bar{x})^2 = 900000, \qquad \sum_i (y_i - \bar{y})^2 = 950000, ?.$$

   (tomar $\alpha = 0,10$*)*

2. *¿Es más efectivo el medicamento B?. Plantear el contraste adecuado para estudiar esto con un nivel de confianza del* 95 %.

Solución. Tenemos

$$X \sim N(\mu_1, \sigma_1^2), \qquad Y \sim N(\mu_2, \sigma_2^2)$$

donde además supondremos independencia.

¿Se puede aceptar igualdad de varianzas? i.e.

$$H_0 : \sigma_1 = \sigma_2,$$
$$H_1 : \sigma_1 \neq \sigma_2,$$
$$\alpha = 0,10$$

$$\mathcal{R} = \left\{ \frac{S_1^2}{S_2^2} \notin \left(F_{m-1;n-1;1-\alpha/2}, F_{m-1;n-1;\alpha/2}\right) \right\}$$

donde

$$\frac{S_1^2}{S_2^2} = \frac{\frac{1}{m-1}\sum_i (x_i - \bar{x})^2}{\frac{1}{n-1}\sum_i (y_i - \bar{y})^2} = \frac{90}{95} \sim 0{,}95,$$
$$F_{m-1;n-1;\alpha/2} = F_{99;99;0{,}05} \simeq F_{120;120;0{,}05} = 1{,}3519,$$
$$F_{m-1;n-1;1-\alpha/2} = F_{99;99;0{,}95} \simeq \frac{1}{F_{120;120;0{,}05}} = 0{,}74$$

donde estamos usando los valores más cercanos y recordamos que

$$F_{m;n;1-\alpha} = \frac{1}{F_{m,n,\alpha}}$$

de esta forma aceptamos H_0 y podemos decir que $\sigma_1 = \sigma_2$.

Con respecto a la segunda cuestión vemos que nos preguntan si $\mu_1 < \mu_2$. Así que

$$\begin{aligned} H_0 &: \mu_1 \geq \mu_2, \\ H_1 &: \mu_1 < \mu_2, \\ \alpha &= 0,05 \end{aligned}$$

con

$$\mathcal{R} = \left\{ \bar{X} - \bar{Y} < t_{m+n-2;1-\alpha} S_p \sqrt{\frac{1}{m} + \frac{1}{n}} \right\}$$

donde

$$S_p = \sqrt{\frac{(m-1)\, S_1^2 + (n-1)\, S_2^2}{m+n-1}} \sim 96{,}66,$$

$$t_{m+n-2;1-\alpha} = t_{198;0{,}95} \sim t_{200;0{,}95} = -t_{200;0{,}05} = -1{,}65$$

así que

$$\mathcal{R} = \{-200 < -22{,}60\}$$

por lo rechazamos $H_0 : \mu_1 \geq \mu_2$ y aceptamos H_1 i.e. podemos afirmar que hay suficiente evidencia muestral para asegurar que B es más efectivo. ■

Ejercicio 4.5.10 *La duración media de una m.a. de* 10 *bombillas es* $\bar{x} = 1250$ *horas, con una cuasidesviación típica muestral de* $s_X = 115$. *Se cambia el material del filamento por otro nuevo y entonces de una nueva m.a. de* 12 *bombillas se obtuvo una duración media de* $\bar{y} = 1340$ *con* $s_Y = 106$.

1. *¿puede aceptarse que las varianzas, antes y después del cambio de filamento sean iguales?*
2. *¿Ha aumentado la duración media de las bombilla?*

Solución. ¿Se puede aceptar igualdad de varianzas? i.e.

$$\begin{aligned} H_0 &: \sigma_1 = \sigma_2, \\ H_1 &: \sigma_1 \neq \sigma_2, \\ \alpha &= 0,10 \end{aligned}$$

$$\mathcal{R} = \left\{ \frac{S_1^2}{S_2^2} \notin \left(F_{m-1;n-1;1-\alpha/2}, F_{m-1;n-1;\alpha/2} \right) \right\}$$

donde

$$\begin{aligned}
\frac{S_1^2}{S_2^2} &= \frac{\frac{1}{m-1}\sum_i (x_i-\bar{x})^2}{\frac{1}{n-1}\sum_i (y_i-\bar{y})^2} = \frac{13225}{11236} = 1{,}18, \\
F_{m-1;n-1;\alpha/2} &= F_{9;11;0{,}05} \simeq 2{,}89, \\
F_{m-1;n-1;1-\alpha/2} &= F_{9;11;0{,}95} \simeq \frac{1}{2{,}89} \sim 0{,}34
\end{aligned}$$

donde estamos usando los valores más cercanos y recordamos que

$$F_{m;n;1-\alpha} = \frac{1}{F_{m,n,\alpha}}$$

de esta forma aceptamos H_0 y podemos decir que $\sigma_1 = \sigma_2$.

Con respecto a la segunda cuestión vemos que nos preguntan si $\mu_1 < \mu_2$. Así que

$$\begin{aligned}
H_0 &: \mu_1 \geq \mu_2, \\
H_1 &: \mu_1 < \mu_2, \\
\alpha &= 0,10
\end{aligned}$$

con

$$\mathcal{R} = \left\{ \bar{X} - \bar{Y} < t_{m+n-2;1-\alpha} S_p \sqrt{\frac{1}{m} + \frac{1}{n}} \right\}$$

donde

$$\begin{aligned}
S_p &= \sqrt{\frac{(m-1)S_1^2 + (n-1)S_2^2}{m+n-1}} \sim 110{,}14, \\
t_{m+n-2;1-\alpha} &= t_{20;0{,}10} = -1{,}325
\end{aligned}$$

así que

$$\mathcal{R} = \{-90 < -62{,}49\}$$

por lo rechazamos $H_0 : \mu_1 \geq \mu_2$ y aceptamos H_1 i.e. podemos afirmar que ha aumentado la duración media. ■

Ejercicio 4.5.11 *Con objeto de estudiar si las pulsaciones en los hombres (X) pueden considerarse menores que en las mujeres (Y) se tomaron muestrasde* 16 *hombre y mujeres obteniéndose los siguientes resultados (ya resumidos)*

	$\sum_{i=1}^{16} x_i$	$\sum_{i=1}^{16} x_i^2$
X	1248	97570
Y	1288	103846

¿Qué podemos decir al respecto?

Solución. Los datos no están emparejados y por lo tanto tenemos

$$X \sim N(\mu_1, \sigma_1^2), \qquad Y \sim N(\mu_2, \sigma_2^2)$$

donde además supondremos independencia. Vemos que

$$\bar{X} = \frac{1}{n}\sum x_i = \frac{1248}{16}78, \qquad \bar{Y} = 80{,}5$$

nos preguntamos si esta diferencia a favor de las mujeres es concluyente. ¿podemos afirmar que $\mu_1 < \mu_2$?. Así que

$$\begin{aligned} H_0 &: \mu_1 \geq \mu_2, \\ H_1 &: \mu_1 < \mu_2, \\ \alpha &= 0,05 \end{aligned}$$

pero consideremos una cuestión previa: ¿podemos aceptar $\sigma_1 = \sigma_2$?, i.e.

$$\begin{aligned} H_0 &: \sigma_1 = \sigma_2, \\ H_1 &: \sigma_1 \neq \sigma_2, \\ \alpha &= 0,10 \end{aligned}$$

Al ser $\sigma_1 = 3{,}7$ y $\sigma_2 = 3{,}2$, podemos concluir que son razonablemente idénticas (comprobar todas estas cuentecillas) i.e.

$$var(X) = \frac{1}{n}\left[\sum_{i=1}^{16} x_i^2 - n\bar{X}^2\right], \text{ etc..}$$

por lo tanto al considerar igualdad de varianzas (si esto no fuese así tendríamos que utilizar otro modelo), tenemos

$$\mathcal{R} = \left\{\bar{X} - \bar{Y} < t_{m+n-2;1-\alpha} S_p \sqrt{\frac{1}{m} + \frac{1}{n}}\right\} = \{-2{,}5 < -2{,}16\}$$

y por lo tanto al verificarse la condición debemos rechazar H_0 y aceptar H_1 i.e. podemos concluir que $\mu_1 < \mu_2$.

Si hubiésemos aceptado H_0 no podríamos concluir que $\mu_1 \geq \mu_2$ (no sería coherente con los datos obtenidos) y por lo tanto diríamos que los datos no son concluyentes. ■

Ejercicio 4.5.12 *Se sospecha que en cierta estación de servicio nos estafan en la gasolina. Para comprobarlo se hacen* 50 *llenados de* 30 *litros cada uno. Llamando* x_i *a cada uno de estos llenados, los resultados obtenidos son* $\sum_{i=1}^{50} x_i = 1475, \sum_{i=1}^{50} x_i^2 = 43550$. *¿Se puede considerar estadísticamente probado, al nivel* 0,01 *que nos sirven menos gasolina?.*

Solución. En este caso la situación es la siguiente:

$$\begin{aligned} H_0 &: \mu \geq 30, \\ H_1 &: \mu < 30, \\ \alpha &= 0,01 \end{aligned}$$

tenemos

$$\mathcal{R} = \left\{ \bar{X} - 30 < t_{49;0{,}99} \frac{S}{\sqrt{50}} \right\} = \{-0{,}5 < -0{,}30\}$$

donde

$$S^2 = \frac{1}{49} \left(\sum_{i=1}^{50} x_i^2 - 49\overline{X}^2 \right) = 0{,}76,$$
$$t_{49;0{,}99} = -2,403$$

por lo que debemos rechazar H_0 y aceptar H_1 i.e. podemos concluir que $\mu_1 < 30$ i.e. que NO nos estafan. ∎

Ejercicio 4.5.13 *Un fabricante de lavadoras produce un determinado modelo en dos colores A y B. De las* 1000 *primeras lavadoras vendidas* 560 *fueron de color A. ¿Proporcionan estos datos suficiente evidencia estadística (al nivel* 0,01*) para concluir que los consumidores prefieren mayoritariamente el color A?.*

Solución. Queremos estimar la proporción de personas que prefieren A frente a B, para ello modelizamos de la siguiente manera

$$X = \begin{cases} 1 & A \\ 0 & B \end{cases}$$

i.e. una Bernoulli (muestras grande $n = 1000$). ¿podemos concluir que $p > 0{,}5$? entonces haremos las siguientes hipótesis

$$H_0 : p \leq 0{,}50,$$
$$H_1 : p > 0{,}50,$$
$$\alpha = 0,01$$

tenemos

$$\mathcal{R} = \left\{ \bar{X} - p_0 > Z_\alpha \sqrt{\frac{p_0\,(1-p_0)}{n}} \right\} = \left\{ 0{,}060 > Z_\alpha \sqrt{\frac{0{,}5\,(1-0{,}5)}{1000}} \right\} = \{0{,}06 > 0{,}037\}$$

donde

$$Z_\alpha = Z_{0{,}01} = 2{,}33,$$
$$\bar{X} = \frac{1}{n} \sum_{i=1} x_i = \frac{560}{1000} = 0{,}56$$

por lo que rechazamos H_0 y aceptar H_1 i.e $p > 0{,}50$.

Antes de terminar veamos si el $p - valor$ es mayor o menor que 0,01. Sabemos por definición que el $p - valor$ es:

$$\sup P_{H_0} = \{\text{obtener un resultado más desfavorable a } H_0 \text{ que el obtenido}\}$$

i.e. apoyo de los datos a H_0?.

$$\begin{aligned} &\text{si } p - valor < \alpha \text{ entonces rechazamos } H_0, \\ &\text{si } p - valor > \alpha \text{ entonces aceptamos } H_0, \end{aligned}$$

luego un $p - valor < \alpha \iff$ rechazar H_0. Como nosotros hemos rechazado H_0 entonces el $p - valor$ debe ser menor que $\alpha = 0{,}01$ (esto nos evita el tener que calcular explícitamente el $p - valor$). ■

Ejercicio 4.5.14 *Una prueba de detección de la hepatitis vírica produce un* $2\,\%$ *de falsos positivos y un* $5\,\%$ *de falsos negativos. Se aplica esta prueba a* 800 *personas tomadas al azar.*

1. *Hallar la relación entre* $p =$ *"Probabilidad de dar positivo" y* $r =$*Probabilidad de padecer hepatitis vírica".*
2. *En* 45 *pruebas, de las* 800*, el resultado es positivo. ¿Se puede considerar estadísticamente probado que la enfermedad afecta a menos del* $8\,\%$ *de la población? (tomando* $\alpha = 0{,}01$*).*

Solución. Tenemos la siguiente situación

$$\begin{aligned} P(positivo \mid sano) &= 0{,}02 \\ P(negativo \mid enfermo) &= 0{,}05 \\ n &= 800 \end{aligned}$$

estamos interesados en calcular la proporción de personas enfermas, r.

La probabilidad de dar positivo será por lo tanto

$$\begin{aligned} p = P(positivo) &= P(sano)P(positivo \mid sano) + P(enfermo)P(positivo \mid enfermo) = \\ &= (1 - r)(0{,}02) + r(0{,}95) \end{aligned}$$

luego

$$p = 0{,}02 + 0{,}93r.$$

Con respecto al segundo apartado modelizaremos de la siguiente manera:

$$X = \begin{cases} 1 & positivo \\ 0 & negativo \end{cases}$$

i.e. mediante una $Bernoulli(p)$. ¿Cuándo podremos concluir que $r < 0{,}08$? i.e. ¿cuándo podremos concluir que $p < 0{,}02 + 0{,}93r = 0{,}094$?, entonces haremos las siguientes hipótesis

$$\begin{aligned} H_0 &: p \geq 0{,}094, \\ H_1 &: p < 0{,}094, \\ \alpha &= 0,01 \end{aligned}$$

tenemos

$$\mathcal{R} = \left\{ \bar{X} - p_0 < Z_{1-\alpha} \sqrt{\frac{p_0 \left(1 - p_0\right)}{n}} \right\}$$

por lo tanto podremos concluir que $p < 0{,}094$ cuando

$$\left\{ \bar{X} - (0{,}094) < Z_{1-\alpha} \sqrt{\frac{0{,}094 \left(1 - 0{,}094\right)}{800}} \right\} = \left\{ \bar{X} - (0{,}094) > Z_{1-\alpha} \left(1{,}031\,8 \times 10^{-2}\right) \right\}$$

donde

$$Z_{1-\alpha} = Z_{0{,}99} = -2{,}33,$$
$$\bar{X} = \frac{\#\text{ de positivos}}{800}$$

así que despejamos $\bar{X}$ y obtenemos que

$$\bar{X} < 55{,}92$$

por lo que podemos concluir que la proporción de enfermos es menor que 0,08 cuando el número de positivos sea menor que 55. ■

Capítulo 5

Distribuciones.

5.1. Distribuciones discretas

5.1.1. Bernoulli.

1. Función generatriz
$$G_X(s) = q + ps.$$
2. Función de masa
$$p_k = P(X = k) = p^k q^{1-k} = p^k (1-p)^{1-k} \qquad \text{con} \quad k = 0, 1.$$
3. Función característica y generatriz de momentos
$$\varphi_X(t) = q + pe^{it}, \qquad M_X(t) = q + pe^t.$$
4. Esperanza y varianza
$$E[X] = p, \quad var[X] = pq, \quad \sigma = \sqrt{pq}$$
5. Resumen

p_k	$\varphi_X(t)$	$M_X(t)$	$E[X]$	$var[X]$
$p^k q^{1-k}$	$q + pe^{it}$	$q + pe^t$	p	pq

5.1.2. Binomial.

Probabilidad de obtener exactamente $k-$*éxitos*.

1. Función generatriz
$$G_X(s) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} s^k = (q + ps)^n.$$

2. Función de masa
$$Bin(n,p) = \binom{n}{k} p^k q^{n-k} \qquad \text{con} \quad k = 0, 1, ...$$

3. Función característica y generatriz de momentos
$$\varphi_X(t) = \left(q + pe^{it}\right)^n, \qquad M_X(t) = \left(q + pe^{t}\right)^n.$$

4. Esperanza y varianza
$$E(X) = np, \quad var(X) = npq, \quad \sigma = \sqrt{npq}$$

5. Resumen

p_k	$\varphi_X(t)$	$M_X(t)$	$E[X]$	$var[X]$
$\binom{n}{k} p^k q^{n-k}$	$\left(q + pe^{it}\right)^n$	$\left(q + pe^{t}\right)^n$	np	npq

Teorema 5.1.1 *Suma de binomiales independientes es otra binomial.*
$$\sum X_i \in Bin\left(\sum n_i, p\right).$$

5.1.3. Geométrica.

Tiempo de espera hasta que aparece el suceso X por primera vez.

1. Función generatriz
$$G_X(s) = \sum_{k=1}^{\infty} p^k q^{k-1} s^k = \frac{ps}{1-qs}.$$

2. Función de masa
$$Geo(p) = (1-p)^{k-1} p \qquad \text{con} \quad k = 1, ...$$

3. Función característica y generatriz de momentos
$$\varphi_X(t) = \frac{pe^{it}}{1-qe^{it}}, \qquad M_X(t) = \frac{pe^{t}}{1-qe^{t}}.$$

4. Esperanza y varianza
$$E(X) = \frac{1}{p}, \quad var(X) = \frac{q}{p^2}, \quad \sigma = \sqrt{\frac{q}{p^2}}.$$

5. Resumen

p_k	$\varphi_X(t)$	$M_X(t)$	$E[X]$	$var[X]$
$(1-p)^{k-1} p$	$\frac{pe^{it}}{1-qe^{it}}$	$\frac{pe^{t}}{1-qe^{t}}$	$\frac{1}{p}$	$\frac{q}{p^2}$

5.1.4. Binomial negativa.

Si en vez de buscar la probabilidad de k éxitos en n ensayos ($n-$fijo) consideramos la variable X igual al número de fallos antes del éxito $n-$ésimo entonces encotramos la binomial negativa.

Observación 5.1.1 *Geométrica es un caso particular de una binomial negativa*

$$Geo \simeq BinN\left(1,p\right).$$

1. Función generatriz
$$G_X(s) = \left(\frac{p}{1-qs}\right)^n.$$

2. Función de masa
$$BinN(n,p) = \binom{n+k-1}{k} p^n q^k \qquad \text{con} \quad k = 0,1,....$$

3. Función característica y generatriz de momentos
$$\varphi_X(t) = \left(\frac{p}{1-qe^{it}}\right)^n, \qquad M_X(t) = \left(\frac{p}{1-qe^{t}}\right)^n.$$

4. Esperanza y varianza
$$E(X) = \frac{n}{p}, \qquad var(X) = \frac{nq}{p^2}, \qquad \sigma = \sqrt{\frac{nq}{p^2}}$$

5. Resumen

p_k	$\varphi_X(t)$	$M_X(t)$	$E[X]$	$var[X]$
$\binom{n+k-1}{k} p^n q^k$	$\left(\frac{p}{1-qe^{it}}\right)^n$	$\left(\frac{p}{1-qe^{t}}\right)^n$	$\frac{n}{p}$	$\frac{nq}{p^2}$

5.1.5. Poisson.

1. Función generatriz
$$G_X(s) = \exp(\lambda(s-1)).$$

2. Función de masa
$$Poisson(\lambda) = \exp(-\lambda)\frac{\lambda^k}{k!} \qquad \text{con} \quad k = 0,....$$

3. Función característica y generatriz de momentos
$$\varphi_X(t) = e^{\lambda}\left(e^{it}-1\right), \qquad M_X(t) = e^{\lambda\left(e^t-1\right)}.$$

4. Esperanza y varianza
$$E(X) = \lambda, \quad var(X) = \lambda, \quad \sigma = \sqrt{\lambda}$$

Teorema 5.1.2 *La suma de Poissones independientes es otra Poisson de parámetro* $\lambda = \sum \lambda_i$

Observación 5.1.2 *Binomial ≈Poisson*
$$Bin(n,p) \rightrightarrows Poiss(\lambda) \ / \ \ \lambda = np \leq 5 \ \ y \ \ p \leq 0{,}1.$$

5.1.6. Hipergeométrica

1. Función de masa
$$f_X(x) = \frac{\begin{pmatrix} k \\ x \end{pmatrix}\begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

2. Esperanza y varianza
$$E(X) = \frac{kn}{N}, \quad var(X) = \frac{N-n}{N-1} n \frac{k}{N}\left(1 - \frac{k}{N}\right),$$

5.2. Variables continuas (v.a.c.).

5.2.1. Uniforme en $[a,b]$.

1. Función de distribución.
$$F_X(t) = \begin{cases} 0 & t < a \\ t & t \in [a,b] \\ 1 & t > b \end{cases}$$

2. Función de densidad.
$$f_X(t) = \begin{cases} \frac{1}{b-a} & t \in [a,b] \\ 0 & t \notin [a,b] \end{cases}$$

3. Función característica y generatriz de momentos
$$\varphi_X(t) = \frac{e^{itb} - e^{ita}}{it\,(b-a)}, \qquad M_X(t) = \frac{e^{tb} - e^{ta}}{t\,(b-a)}.$$

4. Esperanza y varianza.
$$E(X) = \frac{1}{2}(a+b), \qquad var(X) = \frac{1}{12}\,(b-a)^2.$$

5.2.2. Gamma.

1. Función de densidad $\gamma\,(p,a)$.

$$f_X(t) = \begin{cases} \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} & x > 0 \\ 0 & resto \end{cases}$$

   donde

$$\Gamma\,(p) = \int_0^\infty x^{p-1} e^{-x} dx$$

2. Función característica y generatriz de momentos

$$\varphi_X(t) = \left(1 - \frac{it}{a}\right)^{-p}, \qquad M_X(t) = \left(1 - \frac{t}{a}\right)^{-p}.$$

3. Esperanza y varianza.

$$E(X) = \frac{p}{a}, \qquad var(X) = \frac{p}{a^2}.$$

5.2.3. Exponencial.

1. Función de distribución.

$$F_X(t) = \begin{cases} 1 - \exp(-\lambda t) & t > 0 \\ 0 & t \leq 0 \end{cases}$$

   donde $\lambda > 0$.

2. Función de densidad.

$$f_X(t) = \begin{cases} \lambda \exp(-\lambda t) & t > 0 \\ 0 & t \leq 0 \end{cases} \qquad \acute{o} \qquad f_X(t) = \lambda \exp(-\lambda t) I_{(0,\infty)}\,(x)$$

   vemos que

$$\exp\,(\lambda) = \gamma\,(p = 1, \lambda)\,.$$

3. Función característica y generatriz de momentos

$$\varphi_X(t) = \left(1 - \frac{it}{a}\right)^{-1}, \qquad M_X(t) = \left(1 - \frac{t}{a}\right)^{-1}.$$

4. Esperanza y varianza.

$$E(X) = \frac{1}{\lambda}, \qquad var(X) = \frac{1}{\lambda^2}$$

Teorema 5.2.1 *Sea X v.a. exponencial, entonces:*

$$P(X > a + t \mid X > a) = P(A > t),$$

i.e. X tiene la propiedad de no memoria.

Teorema 5.2.2 $(X_i)_{i=1}^n \in \exp(\lambda)$ *v.a.i.i.d., entonces*

$$\sum X_i \in \gamma(n, \lambda).$$

5.2.4. Beta

1. Función de densidad.
$$f_X(t) = \frac{1}{B(p,q)} x^{p-1} (1-x)^{q-1} I_{(o,1)}(x),$$
2. Esperanza y varianza.
$$E(X) = \frac{p}{p+q}, \qquad var(X) = \frac{pq}{(p+q)^2 (p+q+1)}$$

5.2.5. Normal.

1. Función de distribución.
$$F_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx \quad \forall x \in R$$
2. Función de densidad.
$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad \forall x \in R$$
3. Función característica y generatriz de momentos
$$\varphi_X(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}, \qquad M_X(t) = e^{t\mu - \frac{1}{2}\sigma^2 t^2}.$$
4. Esperanza y varianza.
$$E(X) = \mu, \qquad var(X) = \sigma^2$$

Si $X \in N(\mu, \sigma^2)$, tipificando la variable i.e.

$$Z = \frac{X - \mu}{\sigma} \Longrightarrow Z \in N(0,1)$$

Observación 5.2.1 *Aproximación Normal de una binomial. Versión más simple del TCL (ver capítulo 7)*

$$\begin{array}{rcl} Binomial & \approx & Normal. \\ Bin(n,p) & \approx & N(np, \sqrt{npq}) \end{array}$$

Si $Bin(n,p)$ tal que $np \geq 5$ y $nq \geq 5$, entonces podemos aproximar dicha binomial mediante una distribución normal.

La probabilidad binomial $P(k)$ puede aproximarse por la probabilidad normal

$$P(k) \approx P(k - 0{,}5 \leq X \leq k + 0{,}5)$$

Teorema 5.2.3 *Si $\{X_i\}_{i=1}^n$ son v.a. independientes tales que $X_i \in N(\mu_i, \sigma_i)$ $\forall i = 1, ..., n$ entonces*

$$\sum_i X_i \in N\left(\sum_i \mu_i, \sqrt{\sum_i \sigma_i}\right).$$

Teorema 5.2.4 *Si $\{X_i\}_{i=1}^n$ son v.a. independientes tales que $X_i \in N(\mu, \sigma)$ i.e. son v.a.i.i.d., entonces*

$$\frac{\sum_i X_i}{n} \in N\left(\mu, \frac{\sigma^2}{\sqrt{n}}\right).$$

5.2.6. Cauchy

1. Función de densidad.
$$f_X(t) = \frac{1}{\pi\sigma}\frac{1}{1 + \left(\frac{x-p}{\sigma}\right)^2}.$$
2. Función característica
$$\varphi_X(t) = e^{itp - \sigma|t|}.$$
3. Esperanza y varianza.
$$\nexists \qquad E(X) =, \qquad \nexists \qquad var(X) = .$$

5.2.7. χ_n^2. Chi-cuadrado

χ_n^2. Distribución **Chi-cuadrado** con n-grados de libertad. Consideramos $(X_1, ..., X_n)$ v.a.i.i.d. que siguen una $N(0,1)$. Definimos

$$\chi_n^2 = \sum_{i=1}^n X_i^2. \tag{5.1}$$

1. Función de densidad.
$$f_X(t) = \frac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)} e^{-\frac{1}{2}x} x^{\frac{n}{2}-1} I_{(o,\infty)}(x) \qquad \forall x \in R$$
viéndose que
$$\chi_n^2 = \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

2. Función característica y generatriz de momentos

$$\varphi_X(t) = (1-2it)^{-n/2}, \qquad M_X(t) = (1-2t)^{-n/2}.$$

3. Esperanza y varianza.

$$E(X) = n, \qquad var(X) = 2n.$$

5.2.8. t-Student

t-Student, con n-grados de libertad, t_n.

Definición 5.2.1 *Considermos $X \in N(0,1)$ e $Y \in \chi^2_n$ de tal forma que (X,Y) son independientes, definimos*

$$t_n = \frac{X}{\sqrt{\frac{Y}{n}}} = \frac{N(0,1)}{\sqrt{\chi^2_n/n}} \tag{5.2}$$

El cálculo de la función de densidad es fácil ya que al tratarse de v.a.i. entonces $f_{XY} = f_X f_Y$.

1. Función de densidad.

$$f_X(t) = \frac{1}{\sqrt{n}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \qquad \forall x \in \mathbb{R}$$

2. Función característica y generatriz de momentos

$$\varphi_X(t) = (1-2it)^{-n/2}, \qquad M_X(t) = (1-2t)^{-n/2}.$$

3. Esperanza y varianza.

$$E(X) = 0, \text{ si } n > 1, \qquad var(X) = \frac{n}{n-2}, \text{ si } n > 2.$$

4.

$$t_{n;1-\alpha} = -t_{n;\alpha}$$

5.2.9. F Fisher-Snedecor

F Fisher-Snedecor. $F_{m,n}$ con m y n grados de libertad. Sea $X \in \chi^2_m$, e $Y \in \chi^2_n$ v.a.i. entonces definimos

$$F_{m,n} = \frac{\frac{X}{m}}{\frac{Y}{n}} = \frac{Xn}{Ym} = \frac{\chi^2_m/m}{\chi^2_n/n}. \tag{5.3}$$

1. Función de densidad.

$$f_X(t) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m-2}{2}}\left(1+\frac{m}{n}x\right)^{-\frac{m+n}{2}}, \qquad \forall x \in \mathbb{R}$$

2. Esperanza y varianza.

$$E(X) = \frac{m}{m-2}, \quad m > 2, \quad var(X) = 2\left(\frac{n}{n-2}\right)^2 \frac{m+n-2}{m\,(n-4)}, \quad n > 4.$$

3. Propiedades

   a) Si $X \in F_{m,n} \Longrightarrow \frac{1}{X} \in F_{n,m}$

$$\begin{aligned} F_{m;n;1-\alpha} &= \frac{1}{F_{n;m;1-\alpha}} \\ F_{m;n;1-\alpha} &= \frac{1}{F_{m;n;\alpha}} \end{aligned}$$