Capítulo 4

Inferencia Estadística: Introducción.

4.1. Preliminares.

Nos hemos manejado con la Teoría de la Probabilidad y nos disponemos a afrontar problemas de Inferencia Estadística. Para ello, además del manejo de dicha teoría, es de capital importancia establecer un modo de captar información en relación con el fenómeno que estamos estudiando. Esta información será nuestra base de datos para poder inferir, proyectar resultados o emitir juicios en relación al fenómeno que es objeto de estudio. La herramienta capaz de manejar la información en términos probabilísticos es la *muestra*.

4.1.1. Noción de Muestra.

Concepto de muestra: Supongamos que estamos estudiando cierto fenómeno F cuyo espacio de muestras es Ω, i.e. $\mathsf{F} \longleftrightarrow \Omega$. Se entiende por muestra cualquier conjunto de Ω representativo de F.
Una muestra representa en mayor o menor medida al fenómeno si a partir de ésta se infiere alguna información relevante de F, veamos un ejemplo:

EJEMPLO 4.1 *Supongamos que una moneda está cargada de la siguiente manera,*

$$P(C) = 0{,}85, \quad P(X) = 0{,}15.$$

Lanzamos la moneda 10 veces obteniendo la muestra

$$\left\{C, C, C, X, \overset{7)}{\ldots}, X\right\}$$

y nos planteamos en base a dicha muestra si efectivamente $P(C) = 0{,}85$*. Con toda seguridad y a no ser que se disponga de otra información complementaria, diremos que la muestra no es representativa de la suposición inicial o que por contra se admiten dudas sobre dicha hipótesis.*

Ejemplo 4.2 *Supongamos que estamos estudiando los beneficios netos en millones de pesetas durante 1.997, de las empresas ubicadas en CLM. Después de elegir a 10 de las más representativas y preguntar por sus beneficios, la muestra de tamaño 10 resultante es*

$$\{0, 0{,}5, 10, 5, -1, 1, \ldots, -3, 4\}.$$

Si esta muestra es representativa podremos inferir, con cierta precisión, por ejemplo, sobre el beneficio medio en toda la región.

Existen varios tipos de muestras:

(a) Muestreo probabilístico: una muestra es probabilística si puede medirse la probabilidad de obtenerse cada una de las posibles muestras.

(b) Muestreo opinático: un muestreo opinático es aquél que depende únicamente de la opinión del sujeto que lo realiza sobre la representatividad de éste.

(c) Muestreo errático: es aquél que se realiza sin tener en cuenta nada a priori.

(d) Existe un cuarto tipo que estudiamos con detalle y que en todo lo que sigue será el elegido para nuestro trabajo:

Muestreo aleatorio simple, se caracteriza por lo siguiente:
(i) La muestra es elegida al azar siendo la elegida con igual probabilidad que cualquier otra.
(ii) Cada elemento de la población tiene la misma probabilidad que los demás de ser elegido para conformar la muestra. Además tal elección se hace de forma que el hecho de que hayamos seleccionado un elemento no influye para nada sobre la elección de los restantes, es decir, hay independencia entre

los elementos de la muestra.

Paso fundamental en la construcción de una muestra aleatoria simple (m.a.s.): sea Ω una población en la que se quiere estudiar cierto fenómeno y en relación con éste se define una v.a. ξ, de esta forma todos los posibles valores de la ξ se identifican con todas las posibles concreciones del fenómeno. Volvamos al Ejemplo 1 de la moneda: sea la m.a.s. de tamaño 10,

$$\{\omega_1, \omega_2, ..., \omega_{10}\}$$

correspondiente a 10 lanzamientos de la moneda. En lugar de escribir esta sucesión de caras y cruces, y con el objeto de saber si efectivamente la hipótesis $P(C) = 0{,}85$ es verdadera o no, consideraremos la v.a.

$$\xi = \begin{cases} 1 & \text{si sale cara} \\ 0 & \text{si sale cruz} \end{cases},$$

escribiremos

$$\{\xi(\omega_1), \xi(\omega_2), ..., \xi(\omega_{10})\} = \{\xi_1, \xi_2, ..., \xi_{10}\}$$

sucesión de unos y ceros que resulta más manejable que la primera.

En lo que sigue, en lugar de considerar muestras del tipo $\{\omega_1, \omega_2, ..., \omega_n\}$ tendremos en cuenta $\{\xi_1 = \xi(\omega_1), \xi_2 = \xi(\omega_2), ..., \xi_n = \xi(\omega_n)\}$. Puesto que el elemento de la muestra ω_i puede ser cualquiera, son elegidos de manera aleatoria e independiente, el resultado $\xi_i = \xi(\omega_i)$, que representa al resultado de la i-ésima etapa, puede ser cualquier valor de los que toma la v.a. ξ. Hasta que no se concretan los resultados ω_i no sabremos cual es la *realización* (numérica) $\xi_1, \xi_2, ..., \xi_n$ de la muestra, sucesión de n números destinados a inferir resultados sobre F.

Así pues, dada cualquier $\{\omega_1, \omega_2, ..., \omega_n\}$ escribiremos $\{\xi_1, \xi_2, ..., \xi_n\}$, cualquier resultado posible de la realización de la muestra, con ξ_i la variable aleatoria que representa al resultado obtenido en la i-ésima componente de la muestra; además las v.a. $\xi_1, \xi_2, ..., \xi_n$ son independientes.

Definición 4.1 *Una* muestra aleatoria simple *de tamaño n, en relación con cierto fenómeno F y descrito con la ayuda de la v.a. ξ, es un conjunto de n variables aleatorias independientes $\xi_1, \xi_2, ..., \xi_n$ tal que*

$$F_{\xi_i}(x) = F_\xi(x), \quad i = 1, ..., n.$$

Por tanto una m.a.s. de tamaño n es un conjunto de n v.a. independientes e idénticamente distribuidas $\xi_1, \xi_2, ..., \xi_n$ y una realización de tal muestra es un conjunto de n números reales $x_1, x_2, ..., x_n$ que representan a los valores que toma ξ sobre los elementos ω_i seleccionados.

4.1.2. Los métodos de la Inferencia.

Dado un fenómeno aleatorio sobre el cual queremos estudiar ciertos aspectos, construimos una v.a. ξ y una m.a.s. $\xi_1, \xi_2, ..., \xi_n$. Nos propondremos por lo general determinar la naturaleza aleatoria del fenómeno y por tanto, inferir cuál es la ley de probabilidades que modela a la v.a. ξ. Se dan varios casos:

(a) Cuando no sabemos nada sobre la función de distribución F_ξ.

(b) Cuando se sabe que F_ξ pertenece a cierta familia o clase de distribuciones pero se desconoce un parámetro, que en adelante denotaremos por θ, del cual depende la distribución de probabilidad. Por ejemplo, se sabe que $\xi \sim B(1, \theta)$ pero θ no se conoce, sólo se sabe que $\theta \in \Theta$ (a Θ se le llama espacio paramétrico). Nuestro objetivo es determinar de alguna manera este parámetro θ.

El tipo de inferencia estadística que practicaremos en los capítulos 4, 5 y 6 es la que se ha descrito en (b), recibe el nombre de *Inferencia Paramétrica.*

Los métodos que proponemos para obtener información sobre θ son los siguientes:

(i) **Estimación puntual**: consiste en obtener un valor numérico que aproxime al valor real del parámetro θ. Los más interesantes son el *método de los momentos* y el de la *máxima verosimilitud* (Capítulo 5).

(ii) **Estimación por intervalo** (intervalos de confianza): consiste en la construcción de un intervalo $I = [\theta_1^*, \theta_2^*]$ de forma que la probabilidad de que θ pertenezca a I sea grande y a su vez la longitud de dicho intervalo sea pequeña (Capítulo 5).

(iii) **Contrastación de Hipótesis**: consiste en emitir una conjetura sobre θ para más tarde, y en base a una serie de reglas y una muestra, dilucidar si tal conjetura es o no aceptable (Capítulo 6).

4.2. Estadísticos.

Los valores que se observan al realizar una m.a.s. $\xi_1, \xi_2, ..., \xi_n$ son utilizados para inferir sobre la distribución de la v.a. ξ que representa al fenómeno. Nótese que al realizar una muestra las ξ_i se convierten en números concretos, deja de ser aleatorio. Para proyectar información es fundamental el uso de estadísticos.

DEFINICIÓN 4.2 *Un* estadístico T *es cualquier función de la m.a.s.* $\xi_1, \xi_2, ..., \xi_n$ *que no depende de parámetros desconocidos:*

$$T = T(\xi_1, \xi_2, ..., \xi_n).$$

Los estadísticos son cantidades, magnitudes que se calculan una vez que se haya realizado la muestra.

DEFINICIÓN 4.3 *Los estadísticos más importantes:*

1. *El momento de orden* $k \in \mathbb{N}$ *de la m.a.s.* $\xi_1, \xi_2, ..., \xi_n$ *es el estadístico definido como*
$$M_k = \frac{\sum_{j=1}^{n} \xi_j^k}{n}.$$

2. *Se define la* media muestral *como el momento de orden* $1:$
$$\overline{\xi} = \frac{1}{n} \sum_{j=1}^{n} \xi_j.$$

3. *La* cuasivarianza muestral *se define como*
$$S^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(\xi_j - \overline{\xi}\right)^2.$$

4. *Se conoce con el nombre de* varianza muestral *al estadístico*
$$V^2 = \frac{1}{n} \sum_{j=1}^{n} \left(\xi_j - \overline{\xi}\right)^2$$

5. *Definimos la* desviación standard de la muestra *como*

$$S = \sqrt{S^2}.$$

EJEMPLO 4.3 *Supongamos que nos interesamos por la altura de los individuos pertenecientes a cierta población. Se supone que este fenómeno esta modelizado con la ayuda de la v.a.* $\xi =$ *altura del individuo, con* $\xi \sim N(\mu, \sigma)$. *Aquí el parámetro puede ser unidimensional o bidimensional, dependiendo de si se conoce a* μ *o* σ. *Se realiza una m.a.s. de tamaño* $n = 5$ *obteniéndose los resultados*

$$x_1 = 1{,}70,\ \ x_2 = 1{,}74,\ \ x_3 = 1{,}66,\ \ x_4 = 1{,}169\ \ y\ x_5 = 1{,}72.$$

entonces podemos evaluar los distintos estadísticos definidos anteriormente:

$$m_k = \sum_{i=1}^{5} x_i^k = (1{,}70)^k + ... + (1{,}72)^k$$

o

$$s^2 = \frac{1}{4}\sum_{i=1}^{5}\left(x_i^k - m_1\right)^2.$$

Veamos propiedades interesantes de los estadísticos, pero antes necesitamos introducir nuevos tipos de distribuciones de probabilidad:

4.2.1. $\mathcal{X}^2$ de Pearson y t de Student.

Sean $\xi_1, ..., \xi_n$ n v.a. independientes distribuidas todas ellas según una $N(0,1)$. Definimos

$$\eta = \sum_{i=1}^{n} \xi_i^2.$$

Toda v.a. construida como η se distribuye como una $\mathcal{X}_n^2$ de Pearson (con n grados de libertad).

Si $n = 1$ entonces se trata del cuadrado de una $N(0,1)$: $\eta = \xi_1^2$; determinemos su función de distribución: para cada $x \geq 0$ se tiene

$$\begin{aligned} F_\eta(x) &= P\left\{\omega : \xi_1^2(\omega) \leq x\right\} \\ &= P\left\{-\sqrt{x} \leq \xi_1 \leq \sqrt{x}\right\} \\ &= F_{\xi_1}\left(\sqrt{x}\right) - F_{\xi_1}\left(-\sqrt{x}\right) \\ &= \int_{-\infty}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt - \int_{-\infty}^{-\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \\ &= \int_{-\sqrt{x}}^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \end{aligned}$$

El cálculo de la función de densidad es ahora sencillo:

$$\begin{aligned} f_\eta(x) &= \frac{d}{dx} F_\eta(x) \\ &= \frac{d}{dx}\left(F_{\xi_1}\left(\sqrt{x}\right) - F_{\xi_1}\left(-\sqrt{x}\right)\right) \\ &= \frac{F'(\sqrt{x})}{2\sqrt{x}} + \frac{F'(-\sqrt{x})}{2\sqrt{x}} \\ &= \frac{1}{2\sqrt{x}}\left(\frac{1}{\sqrt{2\pi}} e^{-x/2} + \frac{1}{\sqrt{2\pi}} e^{-x/2}\right) \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x}} e^{-x/2}, \end{aligned}$$

con $x \geq 0$.

Función característica:

$$\begin{aligned} \varphi_\eta(t) &= \int_0^\infty e^{itx} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{x}} e^{-x/2} dx \\ &= (1-2it)^{-1/2}. \end{aligned}$$

La esperanza y la varianza:

$$\begin{aligned} E[\eta] &= \frac{1}{i} \varphi_\eta'(0) \\ &= \frac{1}{i}\, i\,(1-2it)^{-3/2}\Big|_{t=0} = 1, \end{aligned}$$

y puesto que $E[\eta^2] = 3$,

$$V[\eta] = 2.$$

Caso general: puesto que las ξ_i se suponen independientes entonces también lo son las ξ_i^2, por consiguiente podremos utilizar los resultados de independencia del Capítulo 3. En particular tendremos que

$$\begin{aligned} \varphi_\eta(t) &= \varphi_{\sum_{i=1}^n \xi_i^2}(t) \\ &= \Pi_{i=1}^n \varphi_{\xi_i^2}(t) \\ &= (1-2it)^{-1/2} \ldots (1-2it)^{-1/2} \\ &= (1-2it)^{-n/2}. \end{aligned}$$

Los cálculos de la esperanza y la varianza son sencillos, se obtiene sin dificultad a que

$$E[\eta] = n, \quad V[\eta] = 2n.$$

Sean $\xi_0, \xi_1, ..., \xi_n$ $n+1$ v.a. independientes idénticamente distribuidas $N(0,1)$. Definimos

$$\overline{y} = \frac{1}{n} \sum_{i=1}^n \xi_i^2$$

y

$$\eta = \frac{\xi_0}{\sqrt{\overline{y}}},$$

esto es, $\eta = \dfrac{N(0,\sigma)}{\sqrt{\frac{\mathcal{X}_n^2}{n}}}$. Una v.a. se dice que tiene una distribución $\mathfrak{t}_n$ - t de Student con n grados de libertad - si está construida como η.

La función de densidad de masa asociada a una $\mathfrak{t}_n$ es

$$f_\eta(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{x^2}{n} + 1\right)^{\frac{-(n+1)}{2}}, \quad x \in \mathbf{R}.$$

Distribución Gamma.

Una v.a. ξ sigue una distribución gamma de parámetros a y p, y se escribirá $\xi \sim \Gamma(p,a)$, si su función de densidad es

$$f_\xi(x) = \begin{cases} \dfrac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases},$$

donde $\Gamma(p)=\int_0^\infty a^p e^{-ax}x^{p-1}dx$ (se demuestra mediante un cambio de variable que $\Gamma(p)=\int_0^\infty e^{-y}y^{p-1}dy$).

Calculemos su función característica:

$$\begin{aligned}
\varphi_\xi(t) &= \frac{a^p}{\Gamma(p)}\int_0^\infty e^{itx}e^{-ax}x^{p-1}dx \\
&= \frac{a^p}{\Gamma(p)}\int_0^\infty e^{-(-it+a)x}x^{p-1}dx \\
&\overset{y=(-it+a)x}{=} \frac{a^p}{\Gamma(p)}\int_0^\infty e^{-y}\left(\frac{y}{-it+a}\right)^{p-1}\frac{dy}{-it+a} \\
&= \frac{a^p}{\Gamma(p)}\frac{1}{(-it+a)^p}\int_0^\infty e^{-y}(y)^{p-1}dy \\
&= \left(\frac{a-it}{a}\right)^{-p} \\
&= \left(1-\frac{it}{a}\right)^{-p}.
\end{aligned}$$

Se observa que una $\mathcal{X}_1^2$ tiene la misma función característica que una gamma Γ de parámetros $p=1/2$, $a=1/2$, y gracias al Teorema de unicidad (ver Capítulo III) tendremos que la distribución de una $\mathcal{X}_1^2$ es de tipo $\Gamma(1/2,1/2)$; de la misma manera vemos que la función característica de una $\mathcal{X}_n^2$ coincide con la de una gamma $\Gamma(n/2,1/2)$, por tanto el Teorema de unicidad garantiza que la distribución de una $\mathcal{X}_n^2$ coincide con la de una $\Gamma(n/2,1/2)$. Por consiguiente, si $\eta\sim\mathcal{X}_n^2$ entonces la densidad asociada es

$$f_\eta(x)=\begin{cases}\dfrac{(1/2)^{n/2}}{\Gamma(n/2)}e^{-x/2}x^{\frac{n}{2}-1} & \text{si } x>0\\ 0 & \text{si } x\le 0\end{cases}.$$

No obstante, tanto en el caso de una $\mathcal{X}_n^2$ como en el de una $\mathfrak{t}_n$, el cálculo de probabilidades se hace mediante tablas.

Proposición 4.1 $S^2=\frac{1}{n-1}\left[\sum_{i=1}^n\xi_i^2-n\left(\overline{\xi}\right)^2\right]$.

La prueba consiste en ejercitarse con ciertos cálculos elementales:

$$
\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{j=1}^{n}\left(\xi_j-\overline{\xi}\right)^2 \\
&= \frac{1}{n-1}\left[\sum_{j=1}^{n}\xi_j^2-2\overline{\xi}\sum_{j=1}^{n}\xi_j+\sum_{j=1}^{n}\overline{\xi}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{j=1}^{n}\xi_j^2-2n\overline{\xi}^2+n\overline{\xi}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{j=1}^{n}\xi_j^2-n\overline{\xi}^2\right].
\end{aligned}
$$

Otra operación elemental de momentos en relación con una m.a.s. $\xi_1,...,\xi_n$ es el cálculo de la varianza, se tiene

$$V\left[M_k\right]=V\left[\frac{\sum_{j=1}^{n}\xi_j^k}{n}\right]=\frac{\alpha_{2k}-\alpha_k^2}{n}$$

donde α_n es el momento de orden n de cualquier ξ_j. Nótese que dada cualquier v.a. ξ y una constante c, $V\left[c\xi\right]=c^2V\left[\xi\right].$

TEOREMA 4.2 (FISHER) *Sea una m.a.s. $\xi_1,...,\xi_n$ de distribución $N\left(\mu,\sigma\right).$ Entonces*

$$\frac{\sum_{j=1}^{n}\left(\xi_j-\overline{\xi}\right)^2}{\sigma^2}=\frac{n-1}{\sigma^2}S^2\sim\mathcal{X}_{n-1}^2$$

y además S^2 y $\overline{\xi}$ son independientes.

Bajo las mismas hipótesis que el Teorema 2.4 es sencillo demostrar que el estadístico $\dfrac{\overline{\xi}-\mu}{\sigma/\sqrt{n}}$ sigue una distribución $N(0,1)$, i.e.

$$\frac{\overline{\xi}-\mu}{\sigma/\sqrt{n}}\sim N(0,1).$$

Para la prueba se tiene presente la siguiente observación: sabemos que si $\xi_1,...,\xi_n$ es una m.a.s. $N\left(\mu,\sigma\right)$ entonces

$$E\left[\overline{\xi}\right]=\mu$$

y

$$V\left[\overline{\xi}\right]=\frac{\sigma^2}{n}.$$

Asimismo se demuestra que $\overline{\xi}\sim N\left(\mu,\sigma/\sqrt{n}\right).$ Así es ya que si $\xi\sim N\left(\mu,\sigma\right)$ entonces $\varphi_{\xi}\left(t\right)=e^{it\mu-\frac{t^2\sigma^2}{2}}$, y gracias a la independencia de los elementos de muestra,

$$\begin{aligned}
\varphi_{\overline{\xi}}\left(t\right) &= \varphi_{\sum_{j=1}^{n}\frac{\xi_j}{n}}\left(t\right)\\
&= \Pi_{j=1}^{n}\varphi_{\frac{\xi_j}{n}}\left(t\right)\\
&= \Pi_{j=1}^{n}\varphi_{\xi_j}\left(t/n\right)\\
&= e^{it\mu/n-\frac{t^2\sigma^2}{2n^2}}...e^{it\mu/n-\frac{t^2\sigma^2}{2n^2}}\\
&= \left(e^{it\mu/n-\frac{t^2\sigma^2}{2n^2}}\right)^n=e^{it\mu-\frac{t^2\left(\sigma/\sqrt{n}\right)^2}{2}},
\end{aligned}$$

que se corresponde con la función característica de una normal de parámetros $\left(\mu,\sigma/\sqrt{n}\right)$, por tanto en virtud del Teorema de unicidad se tiene que $\overline{\xi}\sim N\left(\mu,\sigma/\sqrt{n}\right).$

Este estadístico, $\dfrac{\overline{\xi}-\mu}{\sigma/\sqrt{n}}$, es a menudo usado en inferencia, sin embargo hay ocasiones en las que la desviación típica es desconocida y como parámetro desconocido no puede aparecer en la definición del estadístico. Para solventar esta situación sustituiremos a σ por S, haremos uso de la siguiente proposición:

Teorema 4.3

$$\frac{\overline{\xi}-\mu}{S/\sqrt{n}}\sim \mathfrak{t}_{n-1}.$$

La prueba de este resultado no encierra ninguna dificultad, basta usar el Teorema 4.2 y recordar la definición de una $\mathfrak{t}$ de Student:

$$\begin{aligned}
\frac{\overline{\xi}-\mu}{S/\sqrt{n}} &= \frac{\overline{\xi}-\mu}{\sqrt{S^2/\sigma^2}}\frac{\sigma}{\sqrt{n}}\\
&= \frac{\frac{\overline{\xi}-\mu}{\sigma/\sqrt{n}}}{\frac{1}{\sqrt{n-1}}\sqrt{\frac{n-1}{\sigma^2}S^2}}\\
&= \frac{N(0,1)}{\sqrt{\mathcal{X}_{n-1}^2/n-1}}\sim \mathfrak{t}_{n-1}.
\end{aligned}$$

Dejamos como ejercicio comprobar que si $\xi_1, ..., \xi_n$ es una m.a.s. $N(\mu_1, \sigma_1)$ y $\eta_1, ..., \eta_m$ es una m.a.s. $N(\mu_2, \sigma_2)$, con $\xi_1, ..., \xi_n, \eta_1, ..., \eta_m$ independientes (lo cual implica que $\overline{\xi}$ y $\overline{\eta}$ son independientes), entonces

$$\overline{\xi} - \overline{\eta} \sim N\left(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right).$$

Y un resultado un poco más complicado de demostrar es el siguiente:

$$\sqrt{\frac{mn}{n+m}} \frac{\left(\overline{\xi} - \overline{\eta}\right) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{m+n-2}\left(\frac{(n-1)S_\xi^2}{\sigma_1^2} + \frac{(m-1)S_\eta^2}{\sigma_2^2}\right)}} \sim \mathfrak{t}_{m+n-2}.$$

4.2.2. Distribución $\mathcal{F}$ de Fisher-Snedecor.

Añadimos a nuestra colección de estadísticos otros cuyos distribución es la denominada de Fisher-Snedecor, su definición es la siguiente: sean las variables aleatorias independientes $X \sim \mathcal{X}_n^2$ e $Y \sim \mathcal{X}_m^2$; entonces la v.a. $Z = \dfrac{X/n}{Y/m}$ se dice que tiene una distribución $\mathcal{F}$ de Fisher-Snedecor con n, m grados de libertad, se escribe

$$Z = \frac{X/n}{Y/m} \sim \mathcal{F}_{n,m}$$

y toda v.a. construida como cociente de una $\mathcal{X}_n^2$ entre una $\mathcal{X}_m^2$ tendrá dicha distribución.

Es obvio que si $\xi \sim \mathcal{F}_{n,m}$ entonces $\frac{1}{\xi} \sim \mathcal{F}_{m,n}$.

NOTA 4.1 *El cálculo de probabilidades para las distribuciones* $\mathcal{X}^2$, $\mathfrak{t}$ *o* $\mathcal{F}$ *se hace con la ayuda de tablas.*

4.3. Convergencia en ley.

Tratamos con el estudio de la convergencia saber cuál es la distribución de estadísticos definidos con la ayuda de una muestra aleatoria simple de gran tamaño. Precisamos del concepto de sucesión de variables aleatorias.

Entenderemos por una sucesión de v.a. a una colección infinita de variables aleatorias que denotaremos por $\xi_1, ..., \xi_n, ...$, o de manera abreviada por $\{\xi_n\}$

(y a veces simplemente por ξ_n). Tales v.a. están definidas desde un mismo espacio de probabilidad $(\Omega, \mathcal{A}, P)$ a valores en $\mathbb{R}$, i.e.

$$\xi_n : (\Omega, \mathcal{A}, P) \to \mathbb{R}.$$

Supondremos que cada una de estas variables ξ_n tiene una función de distribución $F_n(x)$ y nuestro objetivo es conocer si estas distribuciones de probabilidad convergen en algún sentido a alguna función de distribución de probabilidad $F(x)$.

Se dirá que $\{\xi_n\}$ es una sucesión de v.a. independientes si al tomar cualquier número finito de v.a. de la sucesión éstas resultan ser independientes. Si además cada una de ellas tiene asociada una misma función de distribución entonces se dirá que están idénticamente distribuidas.

DEFINICIÓN 4.4 *e dice que la sucesión de v.a.* $\{\xi_n\}$, *donde la distribución de* ξ_n *es* $F_n(x)$, converge en ley *a la v.a.* ξ, *siendo* $F(x)$ *la distribución de* ξ, *si para cada punto* x *donde* F *es continua se cumple*

$$\lim_{n\to\infty} F_n(x) = F(x).$$

La notación empleada para indicar que la sucesión $\{\xi_n\}$ converge en ley a ξ es $\xi_n \overset{ley}{\to} \xi$ (o $F_n(x) \overset{ley}{\to} F(x)$).

Puede darse la circunstancia en que $\lim_{n\to\infty} F_n(x) = F(x)$ sin ser F una función de distribución de probabilidad, en tal caso no hay convergencia en ley.

Observamos que si $\xi_n \overset{ley}{\to} \xi$ y $\xi_n \overset{ley}{\to} \eta$ entonces $P\{\omega \in \Omega : \xi(\omega) = \eta(\omega)\} = 1$, es decir las v.a. ξ y η son iguales con probabilidad uno - 'son iguales casi seguro' -.

El siguiente resultado relaciona este tipo de convergencia con las funciones características asociadas a las v.a. que definen la sucesión y la v.a. límite:

TEOREMA 4.4 (DE CONTINUIDAD) *Sea* ξ_n *sucesión de v.a. y* $\varphi_{\xi_n}(t) = \varphi_n(t)$ *sus correspondientes funciones características. Entonces:*

1. *Si* $\xi_n \overset{ley}{\to} \xi$ *entonces* $lim_{n\to\infty}\varphi_n(t) = \varphi(t)$ *para cada* $t \in \mathbb{R}$, *donde* $\varphi(t) = \varphi_\xi(t)$.

2. *Si* $lim_{n\to\infty}\varphi_n(t) = \varphi(t)$ *y* $\varphi(t)$ *es continua en* $t = 0$ *entonces* $\xi_n \overset{ley}{\to} \xi$, *donde* ξ *es cualquier v.a. que tiene como función característica a* $\varphi(t)$.

Con el siguiente resultado se satisface en gran parte el objetivo marcado al principio de la sección, se trata de uno de los resultados más importantes de la Inferencia Estadística:

TEOREMA 4.5 (TEOREMA CENTRAL DEL LÍMITE O DE LEVY-LINDEBERG) *Sea* ξ_n *sucesión de v.a. independientes e idénticamente distribuidas con*

$$E[\xi_n] = \mu$$

y

$$V[\xi_n] = \sigma^2,$$

entonces

$$\frac{\sum_{i=1}^{n}\xi_i - n\mu}{\sqrt{V\left[\sum_{i=1}^{n}\xi_i\right]}} \overset{ley}{\to} \eta$$

con $\eta \sim N(0,1)$.

COROLARIO 4.1 (TEOREMA DE MOIVRE) *Sea* ξ_n *sucesión de v.a. independientes e idénticamente distribuidas con* $\xi_i \sim B(1,p)$ *para todo* i. *Entonces*

$$\frac{\sum_{i=1}^{n}\xi_i - np}{\sqrt{np(1-p)}} \overset{ley}{\to} \eta$$

con $\eta \sim N(0,1)$.

4.4. Ejemplos.

En esta sección nos dedicaremos a ejercitar los conceptos expuestos en secciones precedentes y al mismo tiempo familiarizarnos con el manejo de las tablas en el cálculo de probabilidades de las nuevas distribuciones definidas.

EJEMPLO 4.4 *Se define para todo lo que sigue a*

$$\mathcal{X}^2_{n,\alpha}$$

como aquél número para el cual

$$P\left\{\xi(\omega) \geq \mathcal{X}^2_{n,\alpha}\right\} = \alpha$$

siendo ξ una v.a. cuya distribución es una $\mathcal{X}_n^2$.

Con $n = 14$ nos proponemos encontrar a $\mathcal{X}_{14,0,95}^2$. En las tablas encontramos que $P\{\xi(\omega) \leq 6{,}57\} = 0{,}05$, de esto se desprende que $P\{\xi(\omega) \geq 6{,}57\} = 0{,}95$ y que $\mathcal{X}_{0,95,14}^2 = 6{,}57$.

Tomando $n = 15$, $\xi_1, ..., \xi_{15}$ una m.a.s. $N(\mu, \sigma^2)$ y con la ayuda de los cálculos anteriores demostramos que

$$P\left\{0{,}47\sigma^2 \leq S^2 \leq 1{,}69\sigma^2\right\} \approx 0{,}9, \tag{4.1}$$

lo cual informa de que en el 90 % de las muestras de tamaño 15 que realicemos S^2 estará entre $0{,}47\sigma^2$ y $1{,}69\sigma^2$. Demostramos (4.1):

$$\begin{aligned} P\left\{0{,}47\sigma^2 \leq S^2 \leq 1{,}69\sigma^2\right\} &= P\left\{\frac{6{,}57}{14}\sigma^2 \leq S^2 \leq \frac{23{,}7}{14}\sigma^2\right\} \\ &= P\left\{6{,}57 \leq \frac{(15-1)}{\sigma^2}S^2 \leq 23{,}7\right\} \\ &= F_\xi(23{,}7) - F_\xi(6{,}57) \end{aligned}$$

donde $\xi \sim \mathcal{X}_{14}^2$, basta pues mirar a las tablas de esta distribución y comprobar que esta probabilidad coincide con 0.9 ($F_\xi(23{,}7) = 0{,}95$ y $F_\xi(6{,}57) = 0{,}05$).

Ejemplo 4.5 *Sea $\xi_1, ..., \xi_{10}$ una m.a.s. $N(\mu, \sigma^2)$. Se pide hallar un número A tal que*

$$P\left\{\frac{10}{\sigma^2}\left(\frac{\sum_{i=1}^{10}\xi_i}{n} - \mu\right)^2 \leq A\right\} = 0{,}9. \tag{4.2}$$

Sabemos que

$$\frac{\frac{\sum_{i=1}^{10}\xi_i}{n} - \mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

con lo cual

$$\left(\frac{\frac{\sum_{i=1}^{10}\xi_i}{n} - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \mathcal{X}_1^2.$$

Por lo tanto (4.2) se escribe como

$$P\left\{\frac{10}{\sigma^2}\left(\frac{\sum_{i=1}^{10}\xi_i}{n} - \mu\right)^2 \leq A\right\} = P\{Z \leq A\}$$

siendo Z una v.a. tal que $Z \sim \mathcal{X}_1^2$; basta usar las tablas y según éstas $A = \mathcal{X}_{0,1,1}^2 = 2{,}71$.

Ejemplo 4.6 *Sea* $\xi_1, ..., \xi_n$ *una m.a.s.* $N(\mu, \sigma)$ *y* η *una v.a* $N(0,1)$ *independiente de* $\overline{\xi}$. *Se pide demostrar que*

$$\eta^2 + \frac{n\left(\overline{\xi} - \mu\right)}{\sigma^2} \sim \mathcal{X}_2^2.$$

Se trata de una aplicación directa de algunas definiciones y resultados anteriores, los detalles se dejan al lector.

Ejemplo 4.7 *Sean* $\xi_1, ..., \xi_n$ *y* $\eta_1, ..., \eta_m$ *muestras aleatorias simples independientes de distribuciones* $N(\mu_1, \sigma_1)$ *y* $N(\mu_2, \sigma_2)$ *respectivamente. Se definen las* varianzas muestrales *de éstas como*

$$\begin{aligned} V_\xi^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2, \\ V_\eta^2 &= \frac{1}{m}\sum_{j=1}^{m}\left(\eta_j - \overline{\eta}\right)^2. \end{aligned}$$

Encontrar la ley de probabilidad para el estadístico cociente

$$Y = \frac{V_\xi^2}{V_\eta^2}.$$

Se tiene que

$$\begin{aligned} \frac{V_\xi^2}{V_\eta^2} &= \frac{\frac{n-1}{n}S_\xi^2}{\frac{m-1}{m}S_\eta^2} = \frac{\frac{\sigma_1^2}{n}\frac{n-1}{\sigma_1^2}S_\xi^2}{\frac{\sigma_2^2}{n}\frac{m-1}{\sigma_2^2}S_\eta^2} \\ &= \frac{\frac{\sigma_1^2}{n}}{\frac{\sigma_2^2}{n}}\frac{\mathcal{X}_{n-1}^2}{\mathcal{X}_{m-1}^2} = \frac{\frac{(n-1)\sigma_1^2}{n}}{\frac{(m-1)\sigma_2^2}{n}}\frac{\mathcal{X}_{n-1}^2/\left(n-1\right)}{\mathcal{X}_{m-1}^2/\left(m-1\right)} \\ &= \frac{\frac{(n-1)\sigma_1^2}{n}}{\frac{(m-1)\sigma_2^2}{n}}\mathcal{F}_{n-1,m-1}. \end{aligned}$$

Por tanto

$$\begin{aligned} P\left\{\frac{V_\xi^2}{V_\eta^2} \leq x\right\} &= P\left\{\frac{\frac{(n-1)\sigma_1^2}{n}}{\frac{(m-1)\sigma_2^2}{n}}\mathcal{F}_{n-1,m-1} \leq x\right\} \\ &= P\left\{\mathcal{F}_{n-1,m-1} \leq \frac{\frac{(m-1)\sigma_2^2}{n}}{\frac{(n-1)\sigma_1^2}{n}}x\right\} \\ &= F\left(\frac{\frac{(m-1)\sigma_2^2}{n}}{\frac{(n-1)\sigma_1^2}{n}}x\right) \end{aligned}$$

con F una función distribución de Fisher-Snedecor de parámetros $n-1, m-1$.

EJEMPLO 4.8 *Sean $\xi_1, ..., \xi_{25}$ una m.a.s. $N(\mu, \sigma)$ de tamaño 25. Determinar la probabilidad de que la varianza muestral sea* 0,03 *veces mayor que la varianza poblacional σ^2.*

Se pide evaluar $P\{V^2 \geq 0{,}72\sigma^2\}$; como

$$\begin{aligned} V^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2 \\ &= \frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n}\left(\xi_i - \overline{\xi}\right)^2 \\ &= \frac{n-1}{n}S^2 \end{aligned}$$

entonces

$$\begin{aligned} P\left\{V^2 \geq (0{,}72)\,\sigma^2\right\} &= P\left\{\frac{24}{25}S^2 \geq 0{,}72\sigma^2\right\} \\ &= P\left\{\frac{24}{\sigma^2}S^2 \geq (25)\,(0{,}72)\right\} \\ &= P\left\{\mathcal{X}^2_{24} \geq 18\right\} = 0{,}8. \end{aligned}$$

EJEMPLO 4.9 *Sea el triángulo rectángulo formado por las estrellas A, B y C. Los catetos de dicho triángulo son AC y AB. Las mediciones de tales catetos son de carácter aleatorio de forma que X, la v.a. que mide el cateto AC, e Y, la de AB, están distribuidas según una $N(0,1)$. Se supone que tales mediciones son independientes. Calcular la probabilidad de que una medición de la hipotenusa BC sea superior a $\sqrt{\frac{16}{5}}$.*

La probabilidad pedida es $P\left\{Z \geq \sqrt{\frac{16}{5}}\right\}$, con Z la v.a. que mide la distancia de dicha hipotenusa. Sabemos por el Teorema de Pitágoras que $Z = \sqrt{X^2+Y^2}$, entonces

$$\begin{aligned} P\left\{\sqrt{X^2+Y^2} \geq \sqrt{\frac{16}{5}}\right\} &= P\left\{X^2+Y^2 \geq \frac{16}{5}\right\} \\ &= P\left\{\mathcal{X}^2_2 \geq \frac{16}{5}\right\} = 0{,}83. \end{aligned}$$

EJEMPLO 4.10 *Calcular la probabilidad de que el mayor valor de una m.a.s. de tamaño 10 de distribución* $U(0,1)$, *exceda de* 0,9. *¿Cuál sería la probabilidad de que el menor valor fuese menor que* 0,5*?.*

La probabilidad pedida es

$$\begin{aligned} P\{\max\{\xi_1,...,\xi_{10}\} > 0{,}9\} &= 1 - P\{\max\{\xi_1,...,\xi_{10}\} \leq 0{,}9\} \\ &= 1 - P\{\xi_1 \leq 0{,}9,\ \xi_2 \leq 0{,}9,...,\xi_{10} \leq 0{,}9\} \\ &= 1 - P\{\xi_1 \leq 0{,}9\} P\{\xi_2 \leq 0{,}9\} ... P\{\xi_{10} \leq 0{,}9\} \\ &= 1 - (0{,}9)^{10}. \end{aligned}$$

El siguiente cálculo es

$$\begin{aligned} P\{\min\{\xi_1,...,\xi_{10}\} < 0{,}5\} &= 1 - P\{\min\{\xi_1,...,\xi_{10}\} \geq 0{,}5\} \\ &= 1 - P\{\xi_1 \geq 0{,}5,\ \xi_2 \geq 0{,}5,...,\xi_{10} \geq 0{,}5\} \\ &= 1 - P\{\xi_1 \geq 0{,}5\} ... P\{\xi_{10} \geq 0{,}5\}, \end{aligned}$$

y como $P\{\xi_1 \geq 0{,}5\} = \int_{0,5}^{1} dt = 0{,}5,$

$$P\{\xi_1 \geq 0{,}5\} = 1 - (0{,}5)^{10}.$$

EJEMPLO 4.11 *Sea una m.a.s. de tamaño 16 de una población de tipo normal de media* $\mu = 1$. *Determinar aquél número c para el que*

$$P\left\{\overline{\xi} - \mu \leq c \cdot S\right\} = 0{,}9,$$

donde $\overline{\xi}$ *es la media muestral y* S *la desviación standard de la muestra.*

Observamos que

$$\begin{aligned} P\left\{\overline{\xi} - \mu \leq c \cdot S\right\} &= P\left\{\frac{\overline{\xi} - \mu}{S/\sqrt{16}} \leq c \cdot \sqrt{16}\right\} \\ &= P\left\{t_{16-1} \leq c \cdot \sqrt{16}\right\} = 0{,}9, \end{aligned}$$

y esto según las tablas implica que $4c = 1{,}341$, *es decir* $c = 0{,}33525$.

EJEMPLO 4.12 *Sea una población representada por la v.a.* ξ *con* $E[\xi] = 25$ *y* $V[\xi] = 240$. *Se extrae una m.a.s. de tamaño* $n = 100$. *Determinar* $P\left\{\overline{\xi} \in [23{,}55, 28{,}1]\right\}$.

Se trata de un ejemplo típico en el que se usa el Teorema Central del Límite. Según este resultado

$$\frac{\sum_{i=1}^{n}\xi_i - n\mu}{\sqrt{V\left[\sum_{i=1}^{n}\xi_i\right]}} \overset{ley}{\to} N(0,1),$$

pero esto es lo mismo que

$$\begin{aligned}
\frac{\overline{\xi}-\mu}{\frac{1}{n}\sqrt{V\left[\sum_{i=1}^{n}\xi_i\right]}} &= \frac{\overline{\xi}-\mu}{\frac{1}{n}\sqrt{n\cdot V\left[\xi_i\right]}}\\
&= \frac{\overline{\xi}-25}{\frac{1}{100}\sqrt{100\cdot 240}}\\
&= \frac{\overline{\xi}-25}{\frac{1}{10}\sqrt{240}}.
\end{aligned}$$

Con esto vemos que

$$\begin{aligned}
P\left\{\overline{\xi}\in[23{,}55,28{,}1]\right\} &= P\left\{23{,}55\leq\overline{\xi}\leq 28{,}1\right\}\\
&= P\left\{\frac{23{,}55-25}{\frac{1}{10}\sqrt{240}}\leq\frac{\overline{\xi}-25}{\frac{1}{10}\sqrt{240}}\leq\frac{28{,}1-25}{\frac{1}{10}\sqrt{240}}\right\}\\
&= P\left\{\frac{-1{,}45}{1{,}55}\leq\frac{\overline{\xi}-25}{\frac{1}{10}\sqrt{240}}\leq\frac{3{,}1}{1{,}55}\right\}\\
&= P\left\{-0{,}935\leq\frac{\overline{\xi}-25}{\frac{1}{10}\sqrt{240}}\leq 2\right\}\\
&= F(2)-F(-0{,}935)\\
&= 0{,}9733-0{,}1749=0{,}798,
\end{aligned}$$

donde F es la distribución de una $N(0,1)$.

EJEMPLO 4.13 *Sea una población representada por una v.a. ξ $B(n,p)$. Determinar n y p para que $P\{\xi>n/2\}\geq 0{,}9$.*

Una ξ sigue la misma distribución que la suma de n v.a. $\xi_1,...,\xi_n$ de tipo

$B(1,p)$. Por tanto en lugar de escribir ξ tomaremos $\sum_{i=1}^{n}\xi_i$ y así

$$\begin{aligned} P\{\xi > n/2\} &= P\left\{\sum_{i=1}^{n}\xi_i > n/2\right\} \\ &= P\left\{\frac{\sum_{i=1}^{n}\xi_i - np}{\sqrt{V\left[\sum_{i=1}^{n}\xi_i\right]}} > \frac{n/2 - np}{\sqrt{V\left[\sum_{i=1}^{n}\xi_i\right]}}\right\} \\ &= P\left\{N(0,1) > \frac{n/2-np}{\sqrt{np(1-p)}}\right\} \geq 0{,}9 \end{aligned}$$

si y sólo si $\frac{n/2-np}{\sqrt{np(1-p)}} \leq -1{,}285$. Por ejemplo con $p = \frac{13}{20}$ queda

$$-\frac{3}{20}n = n/2 - n\,(13/20) \leq -1{,}285\sqrt{n\,(91/400)}$$

i.e.

$$(3n/20)^2 \geq \left(1{,}285\sqrt{n\,(91/400)}\right)^2 \approx (0{,}375)\,n$$

con lo cual bastará con elegir

$$n \geq \frac{400}{9}(0{,}375) > 17.$$

4.5. Problemas del Capítulo IV

1. Calcular los siguientes valores de la variable que sigue una $\mathcal{F}$ de Fisher-Snedecor: $\mathcal{F}_{0{,}05,5,21}$, $\mathcal{F}_{0{,}01,24,12}$, $\mathcal{F}_{0{,}05,8,4}$.

2. Calcular las siguientes probabilidades:

   a) $P(\mathcal{X}_{10}^2 \leq 20{,}4)$

   b) $P(\mathcal{X}_{14}^2 > 23{,}2)$

   c) $P(11{,}9 \leq \mathcal{X}_{18}^2 \leq 32{,}4)$

3. Calcular $\mathcal{X}_{45,0{,}025}^2$ y $\mathcal{X}_{24,0{,}95}^2$

4. Dada una población que sigue una distribución normal con media conocida y desvicaión típica desconocida, se extrae una muesra de tamaño 16. Calcular probabilidad $P(0{,}5 < \dfrac{S^2}{\sigma^2} < 1{,}8)$.

5. Calcular $\mathfrak{t}_{15,0,1}$, $\mathfrak{t}_{25,0,2}$ y $\mathfrak{t}_{8,00,2}$.

6. La producción diaria de un determinado artículo se supone de tipo uniforme y oscila entre 6.000 y 10.000 unidades. Determinar la probabilidad de que la producción media supere las 8.100 unidades, habiéndose realizado observaciones durante 320 días, y supuesto que el número de unidades producidas en un día es independiente de los restantes.

7. Sea la variable aleatoria η_n que representa la suma de los puntos que salen al lanzar n veces un dado.

   a) Calcular $E[\eta_n]$ y $V[\eta_n]$.

   b) Utilizar el Teorema Central del Límite para encontrar un natural n, lo más pequeño posible, de modo que

$$P\left\{\left|\frac{\eta_n}{n} - 3{,}5\right| \geq 0{,}1\right\} \leq 0{,}1.$$

8. Encuéntrese un número k tal que

$$P(490 < \xi < k) = 0{,}5$$

   si ξ = número de caras obtenido al lanzar una moneda 1000 veces.

9. Sean $X_1, ..., X_{50}$ m.a.s. de distribución $U(0,1)$ (uniforme en $(0,1)$). Consideramos el estadístico $\overline{X} = \frac{1}{50}\sum_{i=1}^{50} X_i$. Enunciar con todo detalle el Teorema Central del Límite y aplicarlo al cálculo de

$$P\left\{\overline{X} < 0{,}4\right\}.$$

10. La longitud en milímetros ξ que se puede estirar un filamento de nylon sin ruptura sigue una distribución de probabilidad cuya densidad de probabilidad es

$$f_\xi(x) = \begin{cases} 5e^{-5x} & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$

   Emplear el Teorema Central del Límite para determinar de manera aproximada, la probabilidad de que la longitud promedio de 100 filamentos esté comprendida entre 0,18 y 0,22 (Ayuda: $\alpha_1 = 1/5$ y $\alpha_2 = \frac{2}{25}$).

11. Se sabe que en un banco la probabilidad de recibir un cheque sin fondos a lo largo de una semana es 0,15. Si durante una semana se espera recibir 1000 cheques, se pide:

    a) Calcular de manera aproximada, utilizando el Teorema Central del Límite para ello, la probabilidad de que en tal semana se reciban como máximo 125 cheques sin fondos.

    b) Si suponemos que por cada cheque sin fondos la entidad bancaria se beneficia en 200 pesetas, determinar el número de cheques que deben recibir en una semana para que con una probabilidad de 0,9, el banco, por ese concepto, se beneficie en 10000 pesetas.

12. De una población $N(\mu, \sigma)$ y una muestra aleatoria simple $\xi_1, ..., \xi_n$, se pretende estimar su varianza σ^2 a través del estadístico

    $$\xi^* = \frac{\sum_{i=1}^n \xi_i^2}{n}.$$

    Se pide:

    a) Calcular la esperanza de ξ^*.

    b) ¿En qué condiciones ξ^* es insesgado?

    c) ¿Cuándo es ξ^* eficiente?
    (Ayuda: usar que $E[\chi_1^2] = 1$ y que $V[\chi_1^2] = 2$)