Capítulo 4

Muestreo y Estadísticas

4.1 Población, Muestra y Estadísticas

Si por Población entendemos una colección bien definida de elementos en los cuales se mide o registra alguna(s) característica(s) de interés, en el caso de la Economía las poblaciones están constituidas por los llamados "agentes económicos" - consumidores, empresas ,etc.- y en ellos se registra numéricamente características relativas a su comportamiento en la asignación y distribución de recursos; estas características reciben nombre específico, como "Consumo" "Ingreso", "Precio", etc. y el análisis cuantitativo usa las cifras registradas para hacer explícitas las relaciones entre ellas y lograr predicciones confiables. Las características mencionadas antes, cambian de agente en agente, esto es, son "variables" y podemos tratarlas como tales en el análisis formal.

Registrar el valor de una variable económica X en cada elemento de una Población es algo que raramente se hace o se puede hacer. La cantidad de elementos suele ser demasiado grande o estar muy dispersa, para hacer un "levantamiento" total de datos - o sea un Censo -, de modo que uno se tiene que contentar con registrar datos no en toda, sino en parte de la población. Esto es, uno debe trabajar con muestras, con datos muestrales, y confiar en que éstos representen bien a la población.

Las muestras pueden ser generadas de diversas maneras, y una de ellas es la selección "objetiva" de elementos, que consiste en dejar que sea el azar quien defina cuáles elementos constituirán la muestra. Este sistema es objetivo pues podemos confiar en que el azar no tiene favoritos y que si algunos valores en la población son más frecuentes que otros - por ejemplo, las empresas pequeñas en relación a las grandes - esto se verá reflejado a la hora de hacer un sorteo de modo que en última instancia, serán las relaciones entre las fuerzas de la economía las que determinen las componentes de la muestra.

El tomar muestras hace posible el análisis económico, pero también lo complica: las cifras obtenidas no sólo reflejan las relaciones económicas; también tienen una componente aleatoria. Esto es, las variables económicas registradas se convierten en variables aleatorias, debido al mecanismo de sorteo usado para la selección. Entonces, el economista necesita alguna herramienta que le permita separar los "efectos económicos" de los "efectos del azar", que se consideran residuales, de poca importancia relativa, pero que de no ser considerados en el análisis pueden inducir ciertas discrepancias en los pronósticos que podrían ser tomadas como error en el análisis económico de base. Por otro lado, el fracaso de un análisis económico -esto es, la discrepancia grave entre el pronóstico derivado del análisis y la realidad observada- puede ser encubierto por el "efecto del azar" si no nos tomamos la molestia de separar éste último. En resumen, el economista necesita herramientas de trabajo que sean eficientes y formalmente convincentes, con sustento racional.

La Estadística Inferencial trata de las técnicas racionales de análisis de datos provenientes de muestras. Estas técnicas se basan en la Teoría de Probabilidad -las "Leyes del Azar"- y para estudiarlas necesitamos formalizar algunos conceptos, antes de hacer derivaciones lógicas que proporcionen las técnicas que buscamos.

4.1.1 Población

Sea X una variable (o vector) aleatoria con rango R_X y sea $f_X(x)$ su función de densidad o de probabilidad según sea el caso. **La Población de X se define como el conjunto $\{(x, f_X(x)) \mid x \in R_X\}$.**

4.1.2 Muestra Aleatoria

Si X es variable aleatoria, una muestra aleatoria de tamaño n (m.a.) es un vector aleatorio n-dimensional $(X_1, X_2, ..., X_n)$ cuyas componentes representan el proceso de repetir n veces, y de manera independiente, el experimento aleatorio que genera a X, registrando los valores obtenidos.

Observaciones:

(1) Conceptualmente la componente X_j es la misma v.a. X, el subíndice sólo indica la repetición en la cual se registra el valor de X. Por tanto se puede escribir:

$$R_X = R_{X_j}\text{;}\; f_{X_j}(x_j) = f_X(x_j) \text{ y } F_{X_j}(x_j) = F_X(x_j) \qquad \forall j$$

(2) Por construcción hay independencia entre las componentes, por tanto:

$$f_{X_1X_2\ldots X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n) = f_X(x_1) f_X(x_2) \ldots f_X(x_n) \equiv \prod_{j=1}^{n} f_X(x_j)$$

que se conoce como **"distribución de la muestra"** y que es la base para hacer inferencias (derivación de conclusiones Muestra ⇒ Población).

(3) Se resume (1) y (2) diciendo **que $X_1, X_2, \ldots, X_n$ *son independiente e idénticamente distribuidas*, lo que se denota i.i.d.**

(4) Toda la información disponible acerca de la distribución de X está en la muestra, y para procesarla exitosamente por lo general se procede a resumir dicha información con uno o dos índices que representen características importantes de la muestra aleatoria.

4.1.3 Estadística

Una estadística es una función que sólo depende de las componentes de la muestra aleatoria $(X_1, X_2, ..., X_n)$.

Ejemplos:

$$\overline{X} := \frac{\sum_{j=1}^{n} X_j}{n}$$ ("Media muestral": es el equivalente muestral de μ_X)

$$S^2 := \frac{\sum_{j=1}^{n} (X_j - \overline{X})^2}{n-1}$$ ("Varianza muestral": es el equivalente muestral de σ_X^2)

$$M_k := \frac{\sum_{j=1}^{n} X_j^k}{n}$$ ("k-ésimo momento muestral": es el equivalente muestral de m_k)

$$Y_n \equiv X_{(n)} := Máx\{X_1, X_2, ..., X_n\}$$ ("n-ésima estadística de orden" o "Máximo muestral")

Observaciones:

(1) Una estadística es una función que "caracteriza" determinado aspecto de la muestra, que mide una tendencia presente en ella. Así, la media muestral caracteriza la "tendencia central" de los datos muestrales, y la varianza - o mejor aún, su raíz cuadrada S - caracteriza la "tendencia a la dispersión" de los datos. **Como se ve, las**

estadísticas son maneras de procesar la información presente en la muestra, que permiten obtener datos relativos a la población de donde proviene ésta y hacer inferencias de la Muestra a la Población.

(2) Toda estadística es una variable aleatoria: su valor cambia de muestra en muestra y lo hace según el azar, correspondiendo a la distribución que está asociada a la v.a. X que genera la muestra. Esto es, toda estadística tiene una distribución de probabilidades, que por lo general es diferente de la de X **, pero que hereda sus parámetros.**

Ejemplo 1

De una distribución uniforme $U(x;0,\beta)$ se toma una m.a. $(X_1, X_2,..., X_n)$ de tamaño n y se define la estadística $Y := Máx\{X_1,X_2,...,X_n\}$. Hallar la distribución acumulativa $F_Y(y)$ de la estadística Y, su función de densidad $f_Y(y)$ y su valor esperado $E(Y)$

Solución:

Como $\forall X_j \quad 0 \le X_j \le \beta \Rightarrow 0 \le Máx\{X_j\} \le \beta \Rightarrow R_Y = [0,\beta]$. Sea ahora y un elemento cualquiera de R_Y, entonces, en general:

$$F_Y(y) = P[(Y \le y)] = P[Máx\{X_j\} \le y] = P[(X_1 \le y) \cap (X_2 \le y) \cap ... \cap (X_n \le y)] = P\left[\bigcap_{j=1}^{n}(Xj \le y)\right] =$$

$$= \prod_{j=1}^{n} P(X_j \le y) = \prod_{j=1}^{n} F_{X_j}(y) = \prod_{j=1}^{n} F_X(y) = [F_X(y)]^n .$$

Pero por dato $X \sim U(x;0,\beta)$, luego, en este caso particular:

$F_X(y) = \int_0^y \frac{1}{\beta}dx = \frac{y}{\beta}$ y $F_Y(y) = [F_X(y)]^n = \left(\frac{y}{\beta}\right)^n$, es decir, la distribución acumulativa de Y es:

$$F_Y(y) = \left(\frac{y}{\beta}\right)^n \text{ para } 0 \le y \le \beta$$

Derivando con respecto a y, obtenemos $f_Y(y) = \frac{d}{dy}F_Y(y) = n\left(\frac{y}{\beta}\right)^{n-1}(\frac{1}{\beta}) = n\frac{y^{n-1}}{\beta^n}$, es decir, la función de densidad de Y es:

$$f_Y(y) = n\frac{y^{n-1}}{\beta^n} \text{ para } 0 \le y \le \beta$$

Finalmente:

$$E(Y) = \int_0^\beta y f_Y(y)dy = \int_0^\beta yn\frac{y^{n-1}}{\beta^n}dy = \int_0^\beta n\frac{y^n}{\beta^n}dy = \frac{n}{\beta^n}\int_0^\beta y^n dy = \frac{n}{\beta^n}\left[\frac{y^{n+1}}{n+1}\right]_0^\beta = \frac{n}{\beta^n}\left(\frac{\beta^{n+1}}{n+1}\right) = \frac{n}{n+1}\beta ,$$

o sea, el valor esperado de Y es:

$$E(Y) = \frac{n}{n+1}\beta$$

Observaciones:

(1) La densidad de Y no es uniforme como la de X, pero tiene el mismo parámetro β de la "distribución madre" (o sea de la uniforme $U(x;0,\beta)$).

(2) $E(Y) = \frac{n}{n+1}\beta \Leftrightarrow \frac{n+1}{n}E(Y) = \beta \Leftrightarrow E\left(\frac{n+1}{n}Y\right) = \beta$. Por tanto si definimos la estadística $\hat{\beta} := \frac{n+1}{n}Máx\{X_j\}$, aunque $\hat{\beta}$ es variable aleatoria y toma diversos valores, tiene la curiosa propiedad de "coincidir en promedio" con el parámetro β, sea cual fuere el valor de éste. Por tanto, si tomamos una m.a.

de esta distribución, podemos confiar que, de alguna manera, el valor de la v.a. $\hat{\beta}$ estará "alrededor" de β, o sea **podemos 'inferir' que** $\hat{\beta} \cong \beta$

Ejemplo 2

La duración X de una conexión a Internet es una v.a. con distribución exponencial $Exp(x;1/\theta)$ donde $\theta > 0$ es parámetro desconocido. Una institución reguladora piensa tomar una muestra al azar de n consumidores, tomados por sorteo del registro de abonados a un servicio de banda ancha, y registrar los respectivos tiempos de conexión $(X_1, X_2, ..., X_n)$. Se definen las estadísticas:

$Y =$ Duración de la conexión más breve, i.e. $Y := Mín\{X_1, X_2, ..., X_n\}$ y

$W =$ Duración total de las conexiones en la muestra, i.e. $W := \sum_{i=1}^{n} X_i$

a) Calcule la función generatriz de momentos de W y pruebe que W tiene distribución Gamma.

b) Hallar la distribución acumulativa $F_Y(y)$ de la estadística Y, su función de densidad $f_Y(y)$ y pruebe que su valor esperado es β.

Solución:

a) Según la definición, la función generatriz de momentos de W, denotada $M_W(t)$, es $M_W(t) = E\left(e^{tW}\right)$

$$M_W(t) = E\left(e^{tW}\right) = E\left(e^{t\sum_{j=1}^{n} X_j}\right) = E\left(e^{\sum_{j=1}^{n} tX_j}\right) = E\left(\prod_{j=1}^{n} e^{tX_j}\right) = \prod_{j=1}^{n} E\left(e^{tX_j}\right) = \prod_{j=1}^{n} M_{X_j}(t) = \prod_{j=1}^{n} M_X(t) = \left(M_X(t)\right)^n$$

$= \left(\dfrac{1/\theta}{1/\theta - t}\right)^n = \left(\dfrac{1}{1-\theta t}\right)^n$ que en efecto, corresponde a la función generatriz de una distribución Gamma $\Gamma(w; \alpha = n, \beta = \theta)$, es decir $W \sim \Gamma(w; \alpha = n, \beta = \theta)$.

b) $$F_Y(t) = P(Y \le y) = 1 - P(Y > y) = 1 - \prod_{j=1}^{n} P(X_j > y) = 1 - \prod_{j=1}^{n} \int_y^{\infty} \frac{1}{\theta} e^{-\frac{1}{\theta} x_j} dx_j = 1 - \prod_{j=1}^{n} \left[-e^{-\frac{1}{\theta} x_j}\right]_y^{+\infty} = 1 - \prod_{j=1}^{n} \left[e^{-\frac{y}{\theta}}\right]$$

$= 1 - e^{-\frac{ny}{\theta}}$. Luego $F_y(y) = 1 - e^{-\frac{ny}{\theta}}$ $0 < y < \infty$ y derivando tenemos que la función de densidad de Y es $f_y(y) = \dfrac{n}{\theta} e^{-\frac{n}{\theta} y}$ $0 < y < \infty$, es decir $Y \sim Exp(\beta = \dfrac{n}{\theta})$

Nótese que $E(W) = \alpha\beta = n\theta$ y en cambio $E(Y) = \dfrac{1}{\beta} = \left(\dfrac{\theta}{n}\right)$

4.2 Propiedades de estadísticas importantes

Las dos estadísticas más importantes son $\overline{X}$ y S^2, pues por lo general sirven para representar de manera eficiente a la muestra completa:

- $\overline{X}$, la media de la muestra, es el promedio de los n valores obtenibles al tomar la m.a. y sirve para representar a la muestra completa. Juega a nivel de m.a. un papel análogo a la media poblacional μ_X
- S^2, es la varianza muestral y de modo similar a su contrapartida poblacional σ_X^2, mide la dispersión o variabilidad presente en la muestra. Cuantifica qué tanta diferencia hay entre los elementos de la m.a.. Esta característica (lo que se llama "variabilidad") es muy importante porque mucho del esfuerzo desplegado en el análisis económico gira alrededor de la idea de 'explicar' las diferencias observadas entre los agentes económicos en una característica de interés Y, a partir de las diferencias inducidas por otra u otras variables X o Z. Por ejemplo, un buen porcentaje de las diferencias en los ingresos de las personas (Y) se explican por los diversos niveles educativos alcanzados (X) y por el capital o riqueza (Z) de la familia de origen. Es claro que

de acuerdo a esto, podemos descomponer o desagregar la varianza de Y en partes atribuibles a X y a Z respectivamente.

Observación:

Hay otras estadísticas importantes, como la desviación estándar muestral $S := \sqrt{S^2}$ (que mide la distancia promedio entre las componentes de la muestra) o la correlación muestral r_{XY}, pero su tratamiento es más complejo y será objeto de otros cursos.

Propiedad 1

Si $(X_1, X_2, ..., X_n)$ es m.a. tomada de la población de una v.a. X con media μ y varianza σ^2, entonces:

$$E[\overline{X}] = \mu \text{ y } V[\overline{X}] = \frac{\sigma^2}{n}$$

Demostración:

- $n\overline{X} = n\sum_{j=1}^{n} X_j \Rightarrow E[n\overline{X}] = E[\sum_{j=1}^{n} X_j] = \sum_{j=1}^{n} E[X_j] = \sum_{j=1}^{n} E[X] = \sum_{j=1}^{n} \mu = n\mu$; pero $E[n\overline{X}] = nE[\overline{X}]$ (pues E[•] es "operador lineal"). Por tanto $E[\overline{X}] = \mu$.
- También: $V[\overline{X}] = V[\frac{\sum_{j=1}^{n} X_j}{n}] = \frac{1}{n^2} V[\sum_{j=1}^{n} X_j] = \frac{1}{n^2}\{\sum_{j=1}^{n} V(X_j) + 2\sum_{i<j} Cov(X_i X_j)\}$, pero por tratarse de una m.a., sabemos que $V[X_j] = \sigma^2 \forall j$ y $Cov(X_i X_j) = 0 \ \forall i \neq j$ ya que los elementos de una m.a. son variables aleatorias i.i.d (***independiente e idénticamente distribuidas***), es decir $V[\overline{X}] = \frac{1}{n^2} \sum_{j=1}^{n} \sigma^2 = \frac{1}{n^2} n\sigma$ y tenemos que $V[\overline{X}] = \frac{\sigma^2}{n}$

Observación:

Análogamente se puede probar que $E(M_k) = m_k$ y $V(M_k) \frac{m_{2k} - (m_k)^2}{n}$

Lema

En general se cumple que $\sum_{j=1}^{n} (X_j - \overline{X})^2 = \sum_{j=1}^{n} X_j^2 - n\overline{X}^2$

Demostración:

$$\sum_{j=1}^{n}(X_j - \overline{X})^2 = \sum_{j=1}^{n}(X_j^2 + \overline{X}^2 - 2\overline{X}X_j^2) = \sum_{j=1}^{n} X_j^2 + \sum_{j=1}^{n} \overline{X}^2 - 2\overline{X}\sum_{j=1}^{n} X_j = \sum_{j=1}^{n} X_j^2 + n\overline{X}^2 - 2\overline{X}n\overline{X} =$$

$$= \sum_{j=1}^{n} X_j^2 + n\overline{X}^2 - 2n\overline{X}^2 = \sum_{j=1}^{n} X_j^2 - n\overline{X}^2$$

Propiedad 2

En el contexto anterior, se cumple $E(S^2)=\sigma^2$

Demostración:

Recordemos que $S^2 := \dfrac{\sum_{j=1}^{n}(X_j-\overline{X})^2}{n-1} \Rightarrow (n-1)S^2 = \sum_{j=1}^{n}(X_j-\overline{X})^2 = \sum_{j=1}^{n}X_j^2 - n\overline{X}^2$ y tomando valor esperado:

$E[\sum_{j=1}^{n}X_j^2 - n\overline{X}^2] = E[\sum_{j=1}^{n}X_j^2] - E[n\overline{X}^2] = \sum_{j=1}^{n}E[X_j^2] - nE[\overline{X}^2]$. Pero, como sabemos, para cualquier v.a. W se cumple $E(W^2)=V(W)+[E(W)]^2$, y aplicando esta propiedad tanto a X_j^2 como a $\overline{X}^2$ tenemos

$$\sum_{j=1}^{n}E[X_j^2] - nE[\overline{X}^2] = \sum_{j=1}^{n}(\sigma^2+\mu^2) - n(\frac{\sigma^2}{n}+\mu^2) = n\sigma^2 + n\mu^2 - \sigma^2 - n\mu^2 = (n-1)\sigma^2$$

Es decir $E[(n-1)S^2] = (n-1)E[\sum_{j=1}^{n}(X_j-\overline{X})^2] = (n-1)\sigma^2$ y por tanto, como se afirmó. $E[S^2]=\sigma^2$

Observación:

Las propiedades dicen que "en promedio" estas dos estadísticas coinciden con los respectivos parámetros poblacionales. Y en el caso de la media muestral, su varianza demuestra que si el tamaño de muestra es suficientemente grande, prácticamente no hay mayor diferencia entre μ y $\overline{X}$. **Como en un ejemplo anterior, pero de modo más completo, podemos 'inferir' que dada la muestra, se cumple que** $\mu \cong \overline{X} \pm \sqrt{V(\overline{X})} = \overline{X} \pm \frac{\sigma}{\sqrt{n}}$ **Nótese que si n es grande, entonces** $V[\overline{X}] \to 0$ **y 'prácticamente'** $\mu = \overline{X}$ **. La formalización de este hecho tiene un nombre, la Ley de los Grandes Números para** $\overline{X}$ **.**

Propiedad 3 (Ley de los Grandes Números para $\overline{X}$)

Si $(X_1, X_2,..., X_n)$ es m.a. tomada de la población de una v.a. X con media μ y varianza σ^2, entonces:

$$\lim_{n\to\infty} P(|\overline{X}-\mu| \le \varepsilon) = 1$$

Demostración:

- Recordemos la Desigualdad de Tchebychev, válida para toda v.a. W con media μ_W y varianza σ_W^2 finitas:

$$P(|W-\mu_W| < k\sigma_W) \ge 1 - \frac{1}{k^2} \quad \forall k>0$$

- En el caso de $\overline{X}$ **,** ya vimos que $\mu_{\overline{X}} \equiv E[\overline{X}] = \mu$ y $\sigma_{\overline{X}}^2 \equiv V(\overline{X}) = \frac{\sigma^2}{n}$

- Deseamos probar que $\lim_{n\to\infty} P(|\overline{X}-\mu| \le \varepsilon) = 1$. Dado un ε > 0, tomemos k tal que $\varepsilon = k\frac{\sigma}{\sqrt{n}}$ (o sea $k = \frac{\varepsilon\sqrt{n}}{\sigma}$); aplicando Tchebychev $P(|\overline{X}-\mu| \le \varepsilon) \ge 1 - \frac{\sigma^2}{\varepsilon^2 n}$ y como $P(|\overline{X}-\mu| \le \varepsilon) \le 1$, podemos escribir $1 - \frac{\sigma^2}{\varepsilon^2 n} \le P(|\overline{X}-\mu| \le \varepsilon) \le 1$; tomando límite cuando n tiende al infinito:

$$\lim_{n\to\infty} 1 - \frac{\sigma^2}{\varepsilon^2 n} \le \lim_{n\to\infty} P(|\overline{X}-\mu| \le \varepsilon) \le \lim_{n\to\infty} 1 \Leftrightarrow 1 \le \lim_{n\to\infty} P(|\overline{X}-\mu| \le \varepsilon) \le 1.$$

O sea hemos probado que $\lim_{n\to\infty} P(|\overline{X}-\mu|\le\varepsilon)=1$

Observaciones:

(1) Esta Ley garantiza que si no conocemos el valor real μ, es posible "aproximarlo" (estimarlo en jerga técnica) con el margen de error ε que deseemos, usando su equivalente muestral $\overline{X}$ **, si usamos un tamaño de muestra n suficientemente grande.**

(2) **Un teorema similar se cumple en el caso de M_k como estimación de m_k**

Ejemplo:

Se desea aproximar (estimar) el ingreso medio μ de los microempresarios del sector metal mecánico, a partir de la media $\overline{X}$ de una muestra aleatoria de n microempresarios tomados al azar. Se desea que con una probabilidad de 95% o más, el error de estimación $|\overline{X}-\mu|$ sea inferior a las 5 unidades monetarias. De estudios previos, se sabe que la desviación estándar poblacional del ingreso es $\sigma = 25$ u.m. ¿Cuál es el tamaño de muestra n que logra esto?

Solución:

Partiendo de $P(|\overline{X}-\mu|\le\varepsilon)\ge 1-\frac{\sigma^2}{\varepsilon^2 n}$, tomemos $\varepsilon\le 5$ y $1-\frac{25^2}{\varepsilon^2 n}\ge 0.95$ Entonces se cumple

$P(|\overline{X}-\mu|\le 5)\ge 0.95 \Leftrightarrow \varepsilon\le 5$ y $1-\frac{25^2}{\varepsilon^2 n}\ge 0.95 \Leftrightarrow \varepsilon\le 5$ y $\frac{25^2}{\varepsilon^2 n}\le 0.05$ y si tomamos

$n\ge\frac{25^2}{0.05\varepsilon^2}\ge\frac{25^2}{0.05(5)^2}=500$, se cumplen las desigualdades anteriores, por tanto el tamaño mínimo de muestra es n = 500

Propiedad 4 (Teorema Central del Límite para $\overline{X}$)

Sea $(X_1, X_2,..., X_n)$ m.a. tomada de la población de una v.a. X con media μ y varianza σ^2 finitas, entonces si n es suficientemente grande $(n\ge 30)$, se cumple que:

$$Z=\sqrt{n}\left(\frac{\overline{X}-\mu}{\sigma}\right)=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$$

Demostración:

- Recordemos el Teorema general del Límite Central:

Sean $X_1, X_2,\cdots, X_n,\cdots$ variables aleatorias **independientes** con medias $\mu_1,\mu_2,\cdots,\ \mu_n,\cdots$ finitas y varianzas $\sigma_1^2,\sigma_2^2,\cdots,\sigma_n^2,\cdots$ finitas tales que $\sigma_j^2>0\ \forall j$ y $E(|X_j-\mu_j|^3)\equiv\beta_j$ existe $\forall j$. Además supongamos que $B_n = \left(\sum_{j=1}^n\beta_j\right)^{1/3}$ y $C_n=\left(\sum_{j=1}^n\sigma_j^2\right)^{1/2}$ son tales que $\lim_{n\to\infty}(B_n/C_n)=0$.

Sea la v.a. $T=\sum_{j=1}^n X_j$; entonces si n es suficientemente grande $(n\ge 30)$ se cumple que:

$$Z=\frac{T-\mu_T}{\sigma_T}\sim N(0,1) \text{ donde } \mu_T=\sum_{j=1}^n\mu_j \text{ y } \sigma_T^2=\sum_{j=1}^n\sigma_j^2$$

- Aplicando el Teorema general al caso de una m.a., es evidente que $\mu_j=\mu$ y $\sigma_j^2=\sigma^2$ para todo j. Entonces $\mu_T=n\mu$ y $\sigma_T^2=n\sigma^2$ y tenemos que $Z=\frac{T-\mu_T}{\sigma_T}$ se transforma en

$$Z=\frac{T-\mu_T}{\sigma_T}=\frac{T-n\mu}{\sqrt{n\sigma^2}}=\frac{\sum_{j=1}^{n}X_j-n\mu}{\sqrt{n\sigma^2}}=\frac{\dfrac{\sum_{j=1}^{n}X_j-n\mu}{n}}{\dfrac{\sqrt{n\sigma^2}}{n}}=\frac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}$$ que sigue una distribución $N(0,1)$.

Observaciones:

- Informalmente, se escribe que n "grande" $(n\geq 30)\Rightarrow \overline{X}\sim N(\mu,\frac{\sigma^2}{n})$.
- El $n\geq 30$ se ha convertido en un valor típico de "muestra grande", pero hay que tener cuidado, pues si la distribución madre de X es muy asimétrica, n tendrá que ser mayor.
- Nótese que el quid del teorema es que $\overline{X}$ puede ser vista como una suma de variables a la cual es aplicable el Teorema General del Límite Central. Es natural extender esta situación a otras estadísticas que básicamente son sumas; por ejemplo $M_k=\sum_{j=1}^{n}X_j^k/n$ tiene distribución normal si n es grande.
- El teorema hace de $\overline{X}$ una estadística muy especial pues, sin importar la distribución madre de X **si** n **es grande, la distribución de** $\overline{X}$ **es normal.**

Ejemplo:
En el ejercicio sobre la estimación del ingreso medio μ de los microempresarios, asumiendo que n será "grande", podemos refinar el cálculo del tamaño de muestra usando el TLC:

Queremos que $P(|\overline{X}-\mu|\leq 5)\geq 0.95$ pero como $\overline{X}\sim N(\mu,\frac{\sigma^2}{n}=\frac{25^2}{n})$, podemos usar tablas de la distribución normal y $P(|\overline{X}-\mu|\leq 5)\geq 0.95\Leftrightarrow P(\frac{|\overline{X}-\mu|}{\frac{25}{\sqrt{n}}}\leq\frac{5}{\frac{25}{\sqrt{n}}})\geq 0.95\Leftrightarrow P(|Z|\leq\frac{5\sqrt{n}}{25})\geq 0.95$.

Igualando a 0.95 tenemos $P(|Z|\leq\frac{5\sqrt{n}}{25})=0.95\Leftrightarrow P(Z\leq\frac{5\sqrt{n}}{25})=0.975\Leftrightarrow\frac{5\sqrt{n}}{25}=1.96\Leftrightarrow$

$\sqrt{n}=\frac{25(1.96)}{5}=5(1.96)=9.8$

$\Leftrightarrow n=(9.8)^2=96.04\cong 97$ (redondeamos "hacia arriba" para asegurar una probabilidad de 95% o más).

Nótese que el valor de n es mucho menor que el obtenido usando la Ley de Grandes Números, esto se debe a que hemos aplicado un "modelo de datos", más refinado.

4.3 Distribuciones Asociadas al Muestreo de una Distribución Normal

Según vimos en el numeral anterior, la distribución normal es de uso frecuente y por eso vale la pena investigar la distribución de sus dos estadísticos principales.

4.3.1 Distribución de $\overline{X}$

Sea $(X_1,X_2,...,X_n)$ m.a. tomada de la población de una v.a. $X\sim N(\mu,\sigma^2)$. En este contexto se cumple que

$$\overline{X}\sim N(\mu,\frac{\sigma^2}{n})\quad\forall n$$

Demostración:
Calcularemos la función generatriz de momentos de $\overline{X}$ y verificaremos que corresponde a la F.G.M. de una distribución normal:

- $M_{\overline{X}}(t)=E[e^{t\overline{X}}]=E[e^{t\sum_{j=1}^{n}\frac{1}{n}X_j}]=E[e^{\sum_{j=1}^{n}\frac{t}{n}X_j}]=E[\prod_{j=1}^{n}e^{\frac{t}{n}X_j}]=\prod_{j=1}^{n}E[e^{\frac{t}{n}X_j}]=\prod_{j=1}^{n}M_{X_j}(\frac{t}{n})=\prod_{j=1}^{n}M_X(\frac{t}{n})$

  $=[M_X(\frac{t}{n})]^n$, pues al ser las $\{X_j\}$ i.i.d., $M_{X_j}(t)=M_X(t)$ $\forall$ X_j y además el esperado del producto es producto de esperados.
- Como $M_X(t)= e^{t\mu+\frac{t^2}{2}\sigma^2}$, entonces $M_{\overline{X}}(t)=\left[e^{\frac{t}{n}\mu+\frac{t^2}{2n^2}\sigma^2}\right]^n=e^{t\mu+\frac{t^2}{2}\frac{\sigma^2}{n}}$ que corresponde a la función generatriz de momentos de una $N(\mu,\frac{\sigma^2}{n})$, **i.e.** $\overline{X}\sim N(\mu,\frac{\sigma^2}{n})$.

  Nótese que la demostración se cumple para cualquier n, es decir, **si la distribución madre es una normal, $\overline{X}$ tendrá distribución normal de todos modos, para cualquier tamaño de muestra n, no necesariamente grande.**

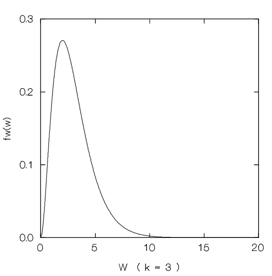
4.3.2 Distribución asociada a S^2.

Si se toma una m.a. de una distribución normal, la varianza muestral S^2 es una estadística que tiene distribución Gamma, pero por razones de comodidad al tabular probabilidades, se trabaja con otra distribución asociada a la Gamma: la llamada distribución Ji-Cuadrado.

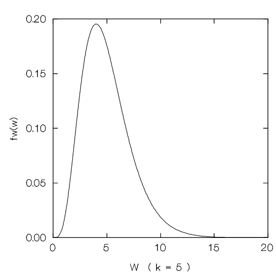
Distribución Ji-Cuadrado $\chi^2(k)$ y propiedades

Sea W v.a. con distribución Gamma $\Gamma(\alpha=\frac{k}{2},\beta=2)$ donde *k* es entero positivo, diremos entonces que W tiene distribución Ji-Cuadrado de parámetro *k*. Este caso particular de la Gamma, se denota $\chi^2(k)$, i.e. $\chi^2(k)\equiv\Gamma(\alpha=\frac{k}{2},\beta=2)$. **El entero k es el único parámetro y se llama "Grados de libertad" de la distribución.** Como se trata de un caso particular de la Gamma, tiene todas sus propiedades, en particular:

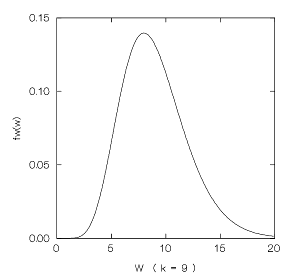
- $f_W(w)=\Gamma(\alpha=\frac{k}{2},\beta=2)=\chi^2(k)=\dfrac{w^{\frac{k}{2}-1}e^{-w/2}}{2^{\frac{k}{2}}\Gamma(\frac{k}{2})}$ $W>0$
- $E[W]=\alpha\beta=\frac{k}{2}2=k$, $V[W]=\alpha\beta^2=\frac{k}{2}2^2=2k$ y $M_W(t)=\left(\dfrac{1}{1-2t}\right)^{\frac{k}{2}}$ si $t<\frac{1}{2}$
- La gráfica de la distribución $\chi^2(k)$ es asimétrica a la derecha, y conforme k crece, la asimetría se va suavizando. En realidad, cada valor de k define una distribución específica, y sería más correcto hablar de la familia de distribuciones $\chi^2(k)$.



Distribución $\chi^2(3)$



Distribución $\chi^2(5)$



Distribución $\chi^2(9)$

Observaciones:

(1) Esta distribución se conoce como distribución Ji-Cuadrado por la letra χ **(**"Ji")que la representa y por el cuadrado del exponente. También la llaman distribución "Chi-Cuadrado" por una desafortunada traducción del nombre de la letra griega χ**.**

(2) **Uso de la Tabla Ji-Cuadrado.** Algunos percentiles de esta distribución han sido tabulados para distintos grados de libertad k . La estructura de la tabla es opuesta a la de la Tabla Z: En el interior de la tabla están los valores de la variable (los percentiles); en el borde izquierdo están los Grados de Libertad k y en la línea superior están las probabilidades, escritas como sub-índice.

Ejemplo

Si $W \sim \chi^2(4)$, entonces por lectura directa de la Tabla tenemos:

$P(W \le 14.9) = 0.995$; $P(W \le 7.78) = 0.90$; $P(W \le 0.297) = 0.01$

Propiedad 1 (Relación con la distribución normal estándar)

Si $Z \sim N(0,1)$ y definimos $W = Z^2$, entonces se cumple que $W \sim \chi^2(k=1)$

Demostración:

$F_W(w) = P(W \le w) = P(Z^2 \le w) = P(-\sqrt{w} \le Z \le \sqrt{w}) = F_Z(\sqrt{w}) - F_Z(-\sqrt{w})$ **y derivando**

$f_W(w) = \frac{1}{2\sqrt{w}} F_Z^{'}(\sqrt{w}) + \frac{1}{2\sqrt{w}} F_Z^{'}(-\sqrt{w}) = \frac{1}{2\sqrt{w}} f_Z(\sqrt{w}) + \frac{1}{2\sqrt{w}} f_Z(-\sqrt{w})$**.**

Pero $f_Z(z) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}$, luego $f_Z(\sqrt{w}) = f_Z(-\sqrt{w}) = \frac{e^{-\frac{w}{2}}}{\sqrt{2\pi}}$ y así $f_W(w) = \frac{1}{\sqrt{w}} f_Z(\sqrt{w}) =$

$\frac{1}{\sqrt{w}} \frac{e^{-\frac{w}{2}}}{\sqrt{2\pi}} = \frac{w^{\frac{1}{2}-1} e^{-w/2}}{2^{\frac{1}{2}} \sqrt{\pi}} = \frac{w^{\frac{1}{2}-1} e^{-w/2}}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})}$ que corresponde a una $\chi^2(k=1)$.

Propiedad 2 (Propiedad Reproductiva)

Si $W_1 \sim \chi^2(k_1)$ y $W_2 \sim \chi^2(k_2)$ son independientes y definimos $W = (W_1 + W_2)$ entonces se cumple que $W \sim \chi^2(k = k_2 + k_2)$

Demostración:

$$M_W(t) = E(e^{tW}) = E(e^{t(W_1+W_2)}) = E(e^{tW_1}e^{tW_2}) = E(e^{tW_1})E(e^{tW_2}) = M_{W_1}(t)M_{W_2}(t)$$

Como $W_1 \sim \chi^2(k_1) \Rightarrow M_{W_1}(t) = \left(\frac{1}{1-2t}\right)^{\frac{k_1}{2}}$ y $W_2 \sim \chi^2(k_2) \Rightarrow M_{W_2}(t) = \left(\frac{1}{1-2t}\right)^{\frac{k_2}{2}}$

entonces $M_W(t) = \left(\frac{1}{1-2t}\right)^{\frac{k_1}{2}+\frac{k_2}{2}} = \left(\frac{1}{1-2t}\right)^{\frac{k_1+k_2}{2}}$ que corresponde a una $\chi^2(k = k_2 + k_2)$,

Se deduce entonces que $W \sim \chi^2(k = k_2 + k_2)$

Ejemplo

Si $Z \sim N(0,1)$, halle c tal que $P(Z^2 \le c) = 0.95$

Solución:

Sea $W = Z^2$. De acuerdo a la Propiedad 1 anterior: $W \sim \chi^2(k=1)$, luego:

$0.95 = P(Z^2 \le c) = P(W \le c)$. Entonces, entrando a la Tabla $\chi^2(k)$ con $k = 1$, vemos en la intersección con la columna $\chi^2_{0.95}$ el valor 3.84, o sea $c = 3.84$

Ejemplo

Si $W_1 \sim \chi^2(5)$ y $W_2 \sim \chi^2(7)$. Halle el percentil 99 de la distribución de la suma $(W_1 + W_2)$

Solución:

Sea $W = (W_1 + W_2)$, aplicando la propiedad reproductiva $W \sim \chi^2(k = 5 + 7 = 12)$ y de la tabla $\chi^2_{0.99} = 26.2$, que ha sido encontrado entrando a la tabla con 12 grados de libertad y una probabilidad de 0.99. Aquí hemos usado $\chi^2_{0.99}$ como notación para representar al percentil 99 de la distribución, siguiendo una costumbre de los textos estadísticos.

Ejemplo

Si $X \sim N(0,2)$ e $Y \sim N(0,4)$ son independientes, halle un valor c tal que $P(2X^2 + Y^2 \le c) = 0.95$

Solución:

Estandarizando primero, tenemos: $Z_1 = X/\sqrt{2} \sim N(0,1)$) y $Z_2 = Y/2 \sim N(0,1)$. Sean ahora, $W_1 = X^2/2$ y $W_2 = Y^2/4$. De la Propiedad 1 se deduce que $W_1 \sim \chi^2(1)$ y que $W_2 \sim \chi^2(1)$ también. Por tanto $(W_1 + W_2) = (X^2/2) + (Y^2/4) \sim \chi^2(2)$. Aplicando lo anterior a los datos:
$0.95 = P(2X^2 + Y^2 \le c) = P((X^2/2) + (Y^2/4)) \le c/4) = P(W_1 + W_2 \le c/4)$. De la tabla Ji-Cuadrado y con $k = 2$ grados de libertad, tenemos $c/4 = \chi^2_{0.95} = 5.99$. Por tanto $c = 4 * 5.99 = 23.96$ es la respuesta al problema.

Distribución asociada a S²

Si de una distribución normal $N(\mu, \sigma^2)$ se toma una m.a. $(X_1, X_2, ..., X_n)$ de tamaño n y se define la variable W mediante $W = (n-1)S^2/\sigma^2$, entonces se cumple:

$$W = (n-1)S^2/\sigma^2 \sim \chi^2(k = n-1)$$

Demostración:

La verificación de esta afirmación, pasa por reconocer que si definimos las variables Z_j, mediante $Z_j = \dfrac{X_j - \mu_j}{\sigma}$, entonces:

- $Z_j \sim N(0,1)$ y aplicando las propiedades 1 y 2 anteriores, tenemos que $Z_j^2 = (\dfrac{X_j - \mu_j}{\sigma})^2 \sim \chi^2(1)$ y por tanto $\sum_{j=1}^{n} Z_j^2 \sim \chi^2(k = n)$
- Por otra parte, $\sum_{j=1}^{n}(X_j - \mu)^2 = \sum_{j=1}^{n}(X_j - \overline{X})^2 + n(\overline{X} - \mu)^2$, ya que

$$\sum_{j=1}^{n}(X_j - \mu)^2 = \sum_{j=1}^{n}(X_j - \overline{X} + \overline{X} - \mu)^2 = \sum_{j=1}^{n}(X_j - \overline{X})^2 + \sum_{j=1}^{n}(\overline{X} - \mu)^2 + 2\sum_{j=1}^{n}(X_j - \overline{X})(\overline{X} - \mu) =$$

$$\sum_{j=1}^{n}(X_j - \overline{X})^2 + \sum_{j=1}^{n}(\overline{X} - \mu)^2 + 2(\overline{X} - \mu)\sum_{j=1}^{n}(X_j - \overline{X}) = \sum_{j=1}^{n}(X_j - \overline{X})^2 + n(\overline{X} - \mu)^2$$, pues como sabemos

$$\sum_{j=1}^{n}(X_j - \overline{X}) = 0$$

- Luego, dividiendo por σ^2 obtenemos: $\dfrac{\sum_{j=1}^{n}(X_j-\mu)^2}{\sigma^2}=\dfrac{\sum_{j=1}^{n}(X_j-\overline{X})^2}{\sigma^2}+\dfrac{n(\overline{X}-\mu)^2}{\sigma^2}$ o equivalentemente $\sum_{j=1}^{n}Z_j^2=(n-1)S^2/\sigma^2+Z^2$, pues como $\overline{X}\sim N(\mu,\dfrac{\sigma^2}{n})$, entonces $Z=\dfrac{(\overline{X}-\mu)}{\dfrac{\sigma}{\sqrt{n}}}\sim N(0,1)$ y por tanto $Z^2=\dfrac{n(\overline{X}-\mu)^2}{\sigma^2}$.
- Como ya vimos $Z^2\sim\chi^2(1)$ y $\sum_{j=1}^{n}Z_j^2\sim\chi^2(k=n)$; por tanto, en virtud de la propiedad reproductiva, tenemos: $W=(n-1)S^2/\sigma^2\sim\chi^2(k=n-1)$

Ejemplo
De una distribución $N(\mu,\sigma^2)$ se piensa tomar una m.a. de tamaño $n=12$. Halle un valor b tal que $P(S^2\le b)=0.95$, si se sabe que $\sigma^2=9$.

Solución:
En este caso $n=12\Rightarrow n-1=11$ y por tanto $(n-1)S^2/\sigma^2=11S^2/9\sim\chi^2(k=11)$. De la probabilidad pedida:

$$0.95=P(S^2\le b)=P(11S^2/9\le 11b/9)\Rightarrow 11b/9=\chi^2_{0.95}(11)=19.7\Rightarrow b=16.12$$

Proposición (Distribución de S^2)
Si de una distribución normal $N(\mu,\sigma^2)$ se toma una m.a. $(X_1,X_2,...,X_n)$ de tamaño n, entonces se cumple que

$$S^2\sim\Gamma(\alpha=\frac{n-1}{2},\beta=\frac{2\sigma^2}{n-1})$$

Demostración:
Como $W=\dfrac{(n-1)S^2}{\sigma^2}$, despejando tenemos $S^2=\dfrac{\sigma^2}{(n-1)}W$ y por tanto $M_{s^2}(t)=E[e^{ts^2}]$

$=E[e^{t(\frac{\sigma^2}{n-1})W}]=E[e^{(\frac{t\sigma^2}{n-1})W}]=M_W(\dfrac{t\sigma^2}{n-1})=\left(1-2\left(\dfrac{t\sigma^2}{n-1}\right)\right)^{-\frac{n-1}{2}}$ que corresponde a la F.G.M de una Gamma de parámetros $\alpha=\dfrac{n-1}{2}$ y $\beta=\dfrac{2\sigma^2}{n-1}$, es decir $S^2\sim\Gamma(\alpha=\dfrac{n-1}{2},\beta=\dfrac{2\sigma^2}{n-1})$

Corolario:

$$E[S^2]=\alpha\beta=\left(\frac{n-1}{2}\right)\left(\frac{2\sigma^2}{n-1}\right)=\sigma^2 \text{ y } V[S^2]=\alpha\beta^2=\left(\frac{n-1}{2}\right)\left(\frac{2\sigma^2}{n-1}\right)^2=\frac{2\sigma^4}{n-1}$$

Observación:
Aplicando la D. de Tchebychev: $n\to\infty\Rightarrow P(|S^2-\sigma^2|\le\varepsilon)\to 1$ $\forall\varepsilon>0$. Es decir S^2 es una buena estimación de σ^2 cuando usamos una muestra de tamaño grande.

4.4 Otras distribuciones importantes.

4.4.1 Distribución t de Student

Definición (Variable T de Student)

Sean $Z \sim N(0,1)$ y $W_1 \sim \chi^2(k)$ independientes. Se define la variable T, llamada variable t de Student, mediante: $T = \dfrac{Z}{\sqrt{\dfrac{W}{k}}}$

Proposición (Distribución t de Student)

En el contexto de la definición anterior, la función de densidad de T es

$$f_T(t) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\,\Gamma(\frac{k}{2})}(1+\frac{t^2}{k})^{-\frac{k+1}{2}} \quad -\infty < t < \infty$$

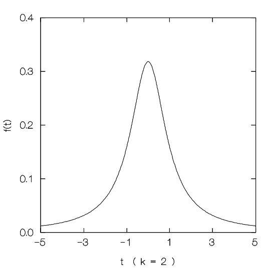
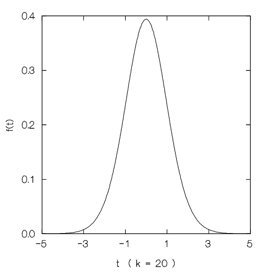
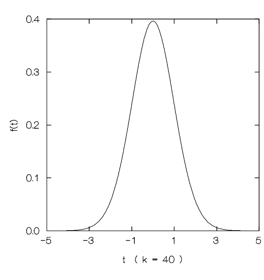
siendo k parámetro de la distribución.

Observaciones:

(1) $f_T(t)$ se llama "distribución t de Student de parámetro k" o también "Distribución t de Student con k Grados de Libertad". La variable T que genera la distribución se llama "Variable T de Student" . El parámetro k de la distribución $\chi^2(k)$ es "heredado" por $f_T(t)$

(2) $E[T] = 0$; $V[T] = \dfrac{k}{k-2}$ para $k > 2$

(3) La gráfica de $f_T(t)$ es similar a la gráfica de la $N(0,1)$, aunque algo más aplanada; conforme k crece, $f_T(t)$ y la distribución $N(0,1)$ se acercan y finalmente coinciden. Dada la simetría con respecto al origen de la gráfica, es posible usar ésta para deducir ciertas propiedades útiles.

		
Distribución t(2)	Distribución t(20)	Distribución t(40)

Notación:

$f_T(t)$ se denota $t(k)$ y se escribe $T \sim t(k)$ para decir que T tiene una distribución t-Student de parámetro k. En verdad, como k puede tomar cualquier valor entero positivo, debiera hablarse de "distribuciones" t-Student, pues cada valor de k define una distribución. Ciertos percentiles han sido tabulados, para distintos valores de k.

Uso de la Tabla t-Student

La estructura de la Tabla es sencilla: Los valores de la variable o percentiles están en el interior; en la primera columna figuran los valores alternativos de k (los "grados de libertad") y en la primera fila están las probabilidades acumuladas.

Ejemplo:
Si $T \sim t(k=10)$, entonces por lectura directa de la Tabla :

$P(T \leq 1.3722) = 0.90$. O sea, el percentil 90, denotado $t_{0.90}$, es $t_{0.90} = 1.372$

También $P(T > 1.3722) = 0.10$ (por complemento) y finalmente, usando la simetría de la distribución: $P(T \leq -1.3722) = 0.10$

Propiedad
Si de una distribución normal $N(\mu,\sigma^2)$ se toma una m.a. $(X_1, X_2,..., X_n)$ de tamaño n y se define la variable t mediante $t = \dfrac{(\overline{X}-\mu)}{\dfrac{S}{\sqrt{n}}}$ entonces $t \sim t(k = n-1)$

Demostración:
La demostración de esta propiedad es directa, si recordamos que:

- $Z = \dfrac{(\overline{X}-\mu)}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$ y que $W = (n-1)S^2/\sigma^2 \sim \chi^2(k = n-1)$
- Aplicando la definición de la variable T: $T = \dfrac{Z}{\sqrt{\dfrac{W}{k}}} = \dfrac{\dfrac{(\overline{X}-\mu)}{\dfrac{\sigma}{\sqrt{n}}}}{\sqrt{\dfrac{(n-1)S^2/\sigma^2}{(n-1)}}} = \dfrac{(\overline{X}-\mu)}{\dfrac{S}{\sqrt{n}}}$ que tiene distribución $t(k = n-1)$

4.4.2 Distribución F de Fisher

Definición (Variable F)
Sean V con distribución $\chi^2(k_1)$ y W con distribución $\chi^2(k_2)$, variables independientes. Se define la variable F, mediante $F = \dfrac{V/k_1}{W/k_2}$.

Proposición (Distribución F de Fisher)

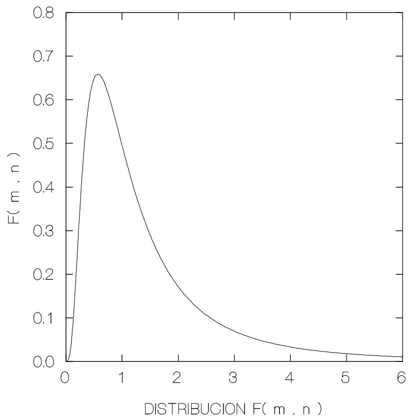
La función de densidad de F es:

$$g_F(f) = \frac{\Gamma(\frac{k_1+k_2}{2})}{\Gamma(\frac{k_1}{2})\Gamma(\frac{k_2}{2})} k_1^{\frac{k_1}{2}} k_2^{\frac{k_2}{2}} \frac{f^{\frac{k_1}{2}-1}}{(k_1+k_2 f)^{\frac{k_1+k_2}{2}}} \quad f > 0$$

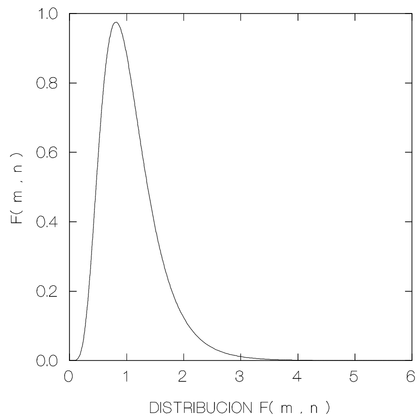
Nota:
k_1 es llamado "grados de libertad del numerador" y k_2 es llamado "grados de libertad del denominador".

La distribución de F se llama Distribución F de Fisher de parámetros k_1 y k_2. Se denota $F(k_1,k_2)$ y se escribe $F \sim F(k_1,k_2)$

En general, la gráfica de la distribución cambia con k_1 y k_2. Dos gráficos típicos de la función de densidad de F, $g_F(f)$, son:



Distribución F(5,15)



Distribucion F(15,30)

Como se ve, la distribución F es asimétrica a la derecha. Conforme k_1 y k_2 crecen, la asimetría se va perdiendo. Como consecuencia de la definición, se tiene la siguiente propiedad:

Propiedad
F tiene distribución $F(k_1,k_2)$ si y sólo si $1/F$ tiene distribución $F(k_2,k_1)$.

Algunos percentiles de esta distribución han sido tabulados, para ciertas parejas (k_1,k_2). Y otros percentiles se pueden encontrar aplicando la propiedad anterior.

Uso de la Tabla F
A la tabla F se entra con los valores de k_1 y k_2 (grados de libertad del numerador y del denominador respectivamente) en <u>ese orden</u>, y los percentiles están en el interior de la tabla.

Ejemplo
Si F tiene distribución $F(7,5)$, por lectura directa se tiene $P(F \leq 3.37) = 0.90$; $P(F \leq 10.5) = 0.975$; $P(F > 4.88) = 0.05$

También como $1/F$ tiene distribución $F(5,7)$ tenemos $P(1/F \leq 9.92) = 0.995$ y por tanto $P(F \geq 1/9.92) = 0.995$ y así $P(F < 1/9.92) = 0.005$

Observación:
Esta distribución es la base para comparar varianzas, a partir del cociente que forman y cuya distribución es una F de Fisher.

Propiedad
Sean Y_1 con distribución $N(\mu_1,\sigma_1^2)$ e Y_1 con distribución $N(\mu_2,\sigma_2^2)$ y sean S_1^2 y S_2^2 varianzas de respectivas muestras independientes de tamaños n_1 y n_2. Entonces $F := \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$.

Propiedad
Si $T \sim t(k)$, entonces $T^2 \sim F(1,k)$