ASIGNATURA FST41: Muestreo de Distribuciones Normales

Cirilo Alvarez Rojas

Universidad Nacional De Ingeniería
Facultad De Ingeniería Económica,Ingeniería Estadística
Y Ciencias Sociales
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA

Demostremos el siguiente teorema que será de mucha utilidad en las siguientes lecciones.

Proposición 1

Las variables aleatorias X y Y son independientes si y solo si, la función característica del vector (X, Y) se puede factorizar;

$$X, Y \text{independientes} \Longleftrightarrow \Phi_{X,Y}(s,t) = \Phi_X(s)\Phi_Y(t)\ \forall (s,t) \in \mathbb{R}^2$$

Prueba.

$$\Phi_{X,Y}(s,t) = \mathrm{E}[e^{isX+itY}] = \mathrm{E}[e^{isX}e^{itY}] = \mathrm{E}[e^{isX}]\mathrm{E}[e^{itY}] = \Phi_X(s)\Phi_Y(t)$$

por independencia de X y Y.

□

Recíprocamente, si $\Phi_{X,Y}(s,t) = \Phi_X(s)\Phi_Y(t)$, entonces, en el caso continuo, se tiene

$$\int\int e^{isx+ity} f(x,y)dxdy = \left[\int e^{isx} f(x)dx\right]\left[\int e^{ity} f(y)dy\right]$$

esto es,

$$\int\int e^{isx+ity} f(x,y)dxdy = \int e^{isx+ity} f(x)f(y)dxdy.$$

Por teorema de la unicidad de la función característica se debe tener

$$f(x,y) = f(x)f(y) \quad \text{para todo } (x,y) \in \mathbb{R}^2$$

Resulta que X y Y son independientes. Una prueba similar se da en el caso donde (X,Y) es de tipo discreto.

DISTRIBUCIONES CHI-CUADRADO, t y F: DISTRIBUCIONES EXACTAS DE MUESTREO

Introducción:

Se investiga ciertas distribuciones que surgen en el muestreo de una población normal. Sea $X_1, X_2, \ldots, X_n$ sea una muestra de la $\mathcal{N}(\mu, \sigma^2)$. Entonces se sabe que $X \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. Además, $\{\sqrt{n}(X - \mu)/\sigma\}^2 \sim \chi^2(1)$. Se Determina la distribución de S^2 seguidamente. Aquí definimos principalmente las distribuciones chi-cuadrado, t y F y estudiamos sus propiedades. Su importancia se hará evidente en lo siguiente y más adelante en la prueba de hipótesis estadísticas.

La primera distribución de interés es la distribución chi-cuadrado, como un caso especial de distribución gamma. Sea $n > 0$ un número entero. Entonces $\Gamma\left(\frac{n}{2}, 2\right) \sim \chi^2(2)$.

Distribución Chi-cuadrado

Definición:

Una variable aleatoria X tiene una distribución chi-cuadrado (χ^2) con n grados de libertad si su función de densidad de probabilidad está dado por

$$f(x;\mathrm{n}) = \begin{cases} \dfrac{x^{\frac{\mathrm{n}}{2}-1} e^{-\frac{1}{2}x}}{2^{\frac{\mathrm{n}}{2}}\Gamma(\frac{\mathrm{n}}{2})}, & \text{si } x > 0, \\ 0, & \text{si } x \leq 0 \end{cases} \qquad (*)$$

Notación: $X \sim \chi^2(\mathsf{n})$ que se lee: la variable aleatoria X se distribuye como una v.a. Chi-cuadrada con n grados de libertad.

Nota: Si $X \sim \chi^2(\mathsf{n})$, entonces

Media: $E(X) = \mathsf{n}$

Varianza: $Var(X) = 2\mathsf{n}$

Asimetría: $\beta_1 = 2\sqrt{\frac{2}{\mathsf{n}}}$

Curtosis: $\beta_2 = 3 + \frac{12}{\mathsf{n}}$

Función Característica (F.C.) $\Phi_X(t) = (1 - 2it)^{-\frac{\mathsf{n}}{2}}$

La distribución $\chi^2(\mathsf{n})$ se tabula para valores de $\mathsf{n} = 1, 2, \ldots$. Las tablas suelen ir hasta $\mathsf{n} = 30$, ya que para $\mathsf{n} > 30$ es posible utilizar una aproximación normal. En la Fig.1 se presenta la función de densidad de $(*)$ para valores seleccionados de n.

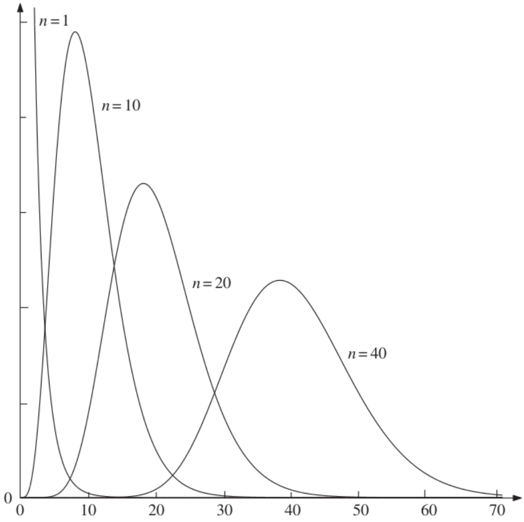

Figura 1: Densidades chi-cuadradas

Ejemplo

Sea n = 25. Entonces utilizando el aplicativo de Matthew Bognar calculamos

$$P(\chi^2(25) \leq 34{,}382) = 0{,}90.$$

Aproximemos esta probabilidad usando **TCL**. Tenemos $E(\chi^2(25)) = 25$ y $Var(\chi^2(25)) = 50$, de modo que

$$\begin{aligned} P(\chi^2(25) \leq 34{,}382) &= P\left(\frac{\chi^2(25) - 25}{\sqrt{50}} \leq \frac{34{,}382 - 25}{5\sqrt{2}}\right) \\ &\approx p(Z \leq 1{,}3268) \\ &= 0{,}90771 \end{aligned}$$

Teorema 1

Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas y sea $S_n = \sum_{j=1}^{n} X_j$. Entonces

(a) $S_n \sim \chi^2(n) \Longleftrightarrow X_1 \sim \chi^2(1)$ y

(b) $X_1 \sim \mathcal{N}(0,1) \Longrightarrow \sum_{j=1}^{n} X_j^2 \sim \chi^2(n)$

Prueba del ítem (a)

$\Longrightarrow$) Supongamos que $S_n \sim \chi^2(n)$, entonces la función característica de S_n es $\Phi_{S_n}(t) = (1-2it)^{-\frac{n}{2}} = \prod_{J=1}^{n} (1-2it)^{-\frac{1}{2}}$, entonces cada factor es una función característica de una variable X_j, esto es $\Phi_{X_j}(t) = (1-2it)^{-\frac{1}{2}}$ para $j = 1, 2, \ldots, n$ y así por el teorema de la unicidad de la función característica $X_1 \sim \chi^2(1)$.

Recíprocamente:

($\Longleftarrow$ si $X_1 \sim \chi^2(1)$, entonces la función característica de X_1 es $\Phi_{X_1}(t) = (1-2it)^{-\frac{1}{2}}$ y como las variables $X_1, \ldots, X_n$ son iid se tiene que,

$$\Phi_{S_n}(t) = \Phi_{\sum_{j=1}^{n} X_j}(t) = \prod_{j=1}^{n}(1-2it)^{-\frac{1}{2}} = (1-2it)^{-\frac{n}{2}}$$

y nuevamente por el teorema de la unicidad de la función característica se concluye que $S_n \sim \chi^2(n)$.

Prueba del ítem (b)

Si $X_1 \sim \mathcal{N}(0,1) \Longrightarrow X_1^2 \sim \chi^2(1)$, y como las variables aleatorias $X_1, X_2, \ldots, X_n$ son independientes e idénticamente distribuidas, por el ítem (a) se concluye que $\sum_{j=1}^{n} X_j^2 \sim \chi^2(n)$.

Definición 1

Sean $X_1, \ldots, X_{\mathrm{n}}$ variables aleatorias normales independientes con $E(X_j) = \mu_j$ y varianza $var(X_j) = \sigma^2, j = 1, 2, \ldots, \mathrm{n}$. También, sea $Y = \sum_{j=1}^{\mathrm{n}} X_i/\sigma_j^2$. La variable aleatoria Y se dice que se distribuye en forma chi-cuadrado no-central con parámetro de no-centralidad $\sum_{j=1}^{\mathrm{n}} \mu_j/\sigma^2$ y n grados de libertad. Se escribe como $Y \sim \chi^2(\mathrm{n}, \delta)$, donde $\delta = \sum_{j=1}^{\mathrm{n}} \mu_j/\sigma^2$. La función de densidad de probabilidad de una variable aleatoria $\chi^2(\mathrm{n}, \delta)$ se puede demostrar que ses igual a

$$f_{\mathrm{n}}(y, \delta) = \begin{cases} \pi^{-\frac{1}{2}} 2^{-\mathrm{n}/2} \exp\left\{-\frac{1}{2}(\delta + y)\right\} y^{(\mathrm{n}-2)/2} \\ \displaystyle\sum_{j=0}^{\infty} \frac{(\delta y)^j (\Gamma(j+1/2))}{(2j)!\Gamma(j+1/2n)} & ; \quad y > 0 \\ 0 & ; \quad y \leq 0 \end{cases} \tag{1}$$

donde $\delta = \sum_{j=1}^{n} \mu_j^2/\sigma^2$. Haciendo $\delta = 0$, observamos que ecuación (1) se reduce a la función de densidad de probabilidad $\chi^2(n)$ central.

Función característica

Tenemos para $Y = \frac{1}{\sigma^2}\sum_{j=1}^{n} X_i^2$

$$
\begin{aligned}
\Psi_Y(t) = \mathrm{E}\left(e^{itY}\right) &= \mathrm{E}\left(e^{it\frac{1}{\sigma^2}\sum_{j=1}^{n} X_j^2}\right) \\
&= \mathrm{E}\left(\prod_{J=1}^{n} e^{it\frac{1}{\sigma^2}X_j^2}\right) \\
&= \prod_{J=1}^{n} \mathrm{E}\left(e^{it\frac{1}{\sigma^2}X_j^2}\right) \qquad \text{por independencia}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left(e^{it\frac{1}{\sigma^2}X_j^2}\right) &= \int_{-\infty}^{\infty} e^{it\frac{1}{\sigma^2}x_j^2} f_{X_j}(x_j)dx_j \\
&= \int_{-\infty}^{\infty} e^{it\frac{1}{\sigma^2}x_j^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_j-\mu_j}{\sigma}\right)^2} dx_j \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}\left\{(x_j-\mu_j)^2-2itx_j^2\right\}} dx_j \quad (*)
\end{aligned}
$$

Ahora desarrollemos el exponente de e en la integral $(*)$

$$
\left\{(x_j-\mu_j)^2 - 2itx_j^2\right\} = \left\{(1-2it)x_j^2 - 2x_j\mu_j + \mu_j^2\right\}
$$

completando al cuadrado se tiene

$$
= \frac{1}{1-2it}\left\{(1-2it)x_j - \mu_j\right\}^2 - \frac{2it}{1-2it}\mu_j^2
$$

reemplazando éste última igualdad dentro de la llave en la integral del $(*)$ resulta

$$E\left(e^{it\frac{1}{\sigma^2}X_j^2}\right) = e^{\left\{\frac{it\mu_j^2}{(1-2it)\sigma^2}\right\}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2(1-2it)\sigma^2}\left\{(1-2it)x_j-\mu_j\right\}^2}dx_j$$

$$= e^{\left\{\frac{it\mu_j^2}{(1-2it)\sigma^2}\right\}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left\{\frac{(1-2it)x_j-\mu_j}{\sqrt{(1-2it)}\sigma}\right\}^2}dx_j$$

haciendo cambio de variable

$$Z = \frac{(1-2it)x_j-\mu_j}{\sqrt{(1-2it)}\sigma}$$

e integrando se obtiene que

$$E\left(e^{it\frac{1}{\sigma^2}X_j^2}\right) = \frac{1}{\sqrt{1-2it}}e^{\left\{\frac{it\mu_j^2}{(1-2it)\sigma^2}\right\}}$$

Luego, la función característica de la distribución chi-cuadrado no central es

$$\begin{aligned}\Psi_Y(t) &= \prod_{J=1}^{\mathsf{n}} \frac{1}{\sqrt{1-2it}} e^{\left\{\frac{it\mu_j^2}{(1-2it)\sigma^2}\right\}} \\ &= (1-2it)^{-\frac{\mathsf{n}}{2}} e^{\frac{1}{(1-2it)\sigma^2}\sum_{j=1}^{\mathsf{n}} \mu_j^2}\end{aligned}$$

Ejercicio 1.

Pruebe que, si $X_1, X_2, \ldots, X_k$ son independientes, $X_i \sim \chi^2(\mathsf{n}_j, \delta_j), j = 1, 2, \ldots, k$, entonces $\sum_{j=1}^{\mathsf{n}} X_j \sim \chi^2\left(\sum_{j=1}^{k} \mathsf{n}_j, \sum_{j=1}^{k} \delta_j\right)$

Ejercicio 2.

Encuentre la media y la varianza de $Y \sim \chi^2(\mathsf{n}, \delta)$

Definición

Considere que la variable aleatoria X se distribuye en forma normal estándar ($X \sim \mathscr{N}(0,1)$) y la variable aleatoria Y se distribuye en forma de distribución chi-cuadrado con n grados de libertad ($Y \sim \chi^2(\mathrm{n})$) y además, las variables X y Y son independientes. Entonces el estadístico

$$T = \frac{X}{\sqrt{Y/\mathrm{n}}} \tag{2}$$

se dice que tiene una distribución-t con n grados de libertad y se escribe $T \sim t(\mathrm{n})$.

Distribución t de "student"

Una variable aleatoria X tiene distribución t de "student" si su función de densidad de probabilidad está dada por

$$f_X(x) = \frac{\Gamma\left(\left(\frac{n+1}{2}\right)\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < x < \infty$$

donde n es el parámetro de distribución y recibe el nombre de "número de grados de libertad".

Notación $X \sim t(n)$

Propiedades

Si $X \sim t(\mathrm{n})$, entonces tiene las siguientes propiedades:

(a) La distribución de t es simétrica respecto de $x = 0$.

(b) $\max f_X$ se da en $x = 0$ y es igual a $\frac{\Gamma((\frac{n+1}{2}))}{\Gamma(\frac{n}{2})\sqrt{n\pi}}$

(c) Función de distribución:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{\Gamma((\frac{n+1}{2}))}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt.$$

(d) Por la simetría se tiene $F_X(-a) = 1 - F_X(a)$

(e) $\mathrm{E}(X) = 0$ para todo $\mathrm{n} \in \mathbb{N}$ y $Var(X) = \frac{n}{n-2}$ para $n > 2$.

(f) Cuando $n \to \infty$ la distribución t se aproxima para para la distribución normal estándar.

Comentario 1

Para $\mathrm{n} = 1$, T *es una variable Cauchy. por consiguiente se sumirá* $\mathrm{n} > 1$. *para cada* n *se tiene las diferentes densidades de probabilidad. En la Figura 2 se muestra* $f_{\mathrm{n}}(t)$ *para algunos valores seleccionados de* n.

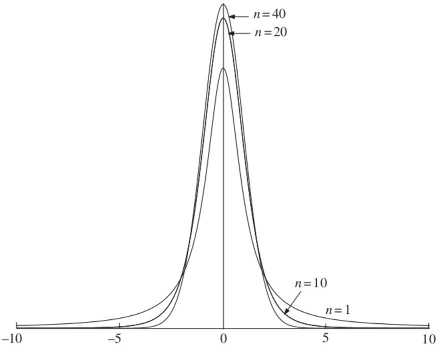

Figura 2: Densidades t-student

Como la distribución normal, la distribución t es importante en la teoría de la estadística y, por tanto, se tabula (ahora, se puede hallar cualquier probabilidad utilizando el aplicativo).

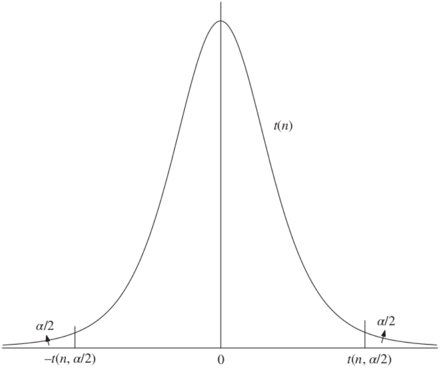

Figura 3:

Comentario 2

La fdp $f_{\mathrm{n}}(t)$ *es simétrica en* t, *y* $f_{\mathrm{n}}(t) \to 0$ *cuando* $t \to +\infty$. *Para* n *grande, la distribución* t *está cerca de la distribución normal. De hecho,* $(1 + t^2/\mathrm{n})^{-(n+1)/2} \to e^{-t^2/2}$ *cuando* $\mathrm{n} \to \infty$. *Además, cuando* $t \to \infty$ *o* $t \to -\infty$, *las extremos de* $f_{\mathrm{n}}(t) \to 0$ *mucho más lentamente que los extremos de la fdp* $\mathscr{N}(0,1)$. *Así, para n pequeño y* t_0 *grande se cumple la siguiente relación*

$$P(|T| > t_0) \geq P(|Z| > t_0), \qquad Z \sim \mathscr{N}(0,1),$$

es decir, hay más probabilidad en el extremo de la distribución t *que en el extremo cola de la distribución normal estándar. En lo que sigue escribiremos* $t(n, \frac{\alpha}{2})$ *para el valor de* T *(Figura 3) para el cual*

$$P(|T| > t(n, \frac{\alpha}{2})) = \alpha$$

Los valores positivos de $t(n, \alpha)$ se tabulan para algunos valores seleccionados de n y α. Los valores negativos se pueden obtener de la simetría, $t(\mathrm{n}, 1-\alpha) = -t(\mathrm{n}, \alpha)$.

Ejemplo

Sea $\mathrm{n} = 5$ y $\alpha = 0{,}05$ entonces utilizando el aplicativo, se obtiene $t(5, 0{,}025) = 2{,}57058$ y $t(5, 0{,}05) = 2{,}01505$.

Los correspondientes valores según la distribución normal $\mathscr{N}(0, 1)$ son $\Phi^{-1}(0{,}025) = Z = 1{,}95996$ y $\Phi^{-1}(0{,}05) = Z = 1{,}64485$

Para $n = 30$

$$F_T^{-1}(30, 0{,}95) = t(30, 0{,}05) = 1{,}69726 \quad \text{y} \quad \Phi^{-1}(0{,}95) = z = 1{,}64485$$

Teorema 2

Si $X \sim t(n), n > 1$. Entonces $\mathrm{E}(X^r)$ existe para $r < n$. En particular, si $r < n$ es impar

$$\mathrm{E}(X^r) = 0$$

y si $r < n$ es par

$$\mathrm{E}(X^r) = n^{\frac{r}{2}} \frac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}$$

Distribución T no central

Definición 2

Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ *y* $\frac{Y}{\sigma^2} \sim \chi^2(n)$ *y* X *y* Y *son independientes,la variable aleatoria definida por*

$$T = \frac{X}{\sqrt{\frac{y}{n}}}$$

La función de densidad de una distribución t *no-central está dada por*

$$f_{\mathrm{n}}(t, \delta) = c \sum_{j=0}^{\infty} \frac{\Gamma\left(\frac{n+k+1}{2}\right) \delta^k 2^{\frac{k}{2}} t^k}{\Gamma(k+1)(n+t^2)^{\frac{k}{2}}}$$

donde

$$c = \frac{n^{\frac{n}{2}}}{\sqrt{\pi}\Gamma\left(\frac{n}{2}\right)} \times \frac{e^{-\frac{\delta^2}{2}}}{(n+t^2)^{\frac{n+1}{2}}}.$$

Si T tiene una distribución t no-central con n grados de libertad y prámetro de no-centralidad δ, entonces:

$$E(T) = \delta\sqrt{\frac{n}{2}}\,\Gamma\left(\frac{n-1}{2}\right)\left[\Gamma\left(\frac{n}{2}\right)\right]^{-1}$$

$$Var(T) = \sigma_T^2 = \frac{n}{n-2}(1+\delta^2) - [E(T)]^2,\ n > 2$$

Distribución F

Definición 3

Sean X y Y variables aleatorias independientes chi-cuadras com m y n grados de libertad respectivamente. Entonces se dice que la variable aleatoria

$$f = \frac{X/m}{Y/n} \tag{3}$$

tiene distribución-F con (m, n) grados de libertad, y se escribe $f \sim F(m, n)$.

Teorema 3

La función de densidad de probabilidad del estadístico-F dada en al ecuación (3) está dada por,

$$g(f) = \begin{cases} \frac{\Gamma\left[\frac{(m+n)}{2}\right]}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right) \left(\frac{m}{n} f\right)^{\frac{m}{2}-1} \left(1 + \frac{m}{n} f\right)^{-\frac{(m+n)}{2}} & f > 0, \\ 0, & f \leq 0 \end{cases} \tag{4}$$

Comentario 3

Si $X \sim F(m, n)$*, entonces* $1/X \sim F(n, m)$*. Si hacemos* $m = 1$*, entonces* $F = [t(n)]^2$ *de modo que* $F(1, n)$ *y* $t^2(n)$ *tienen la misma distribución. También se continúa, que si* $Z \sim \mathscr{C}(0, 1)$ *[la cual es la misma como* $t(1)$*],* $Z^2 \sim F(1, 1)$.

Comentario 4

Como es natural, escribimos $F(m, n; \alpha)$ para el porcentaje de puntos supriores de la distribución $F(m, n)$, esto es,

$$P\left(F(m,n) > F(m,n;\alpha)\right) = \alpha \tag{5}$$

De la observación 3, se tiene la siguiente relación

$$F(m,n;1-\alpha) = \frac{1}{F(n,m;\alpha)} \tag{6}$$

Por tanto es suficiente tabular valores de F que sean mayores o iguales a 1.

Teorema 4

Si $X \sim F(m, n)$. Entonces, para $k > 0$, entero,

$$E(X^k) = \left(\frac{n}{m}\right)^k \frac{\Gamma\left[k + \frac{m}{2}\right]\Gamma\left[\frac{n}{2} - k\right]}{\Gamma\left[\frac{m}{2}\right]\Gamma\left[\frac{n}{2}\right]} \qquad n > 2k \tag{7}$$

En particular,

$$E(X) = \frac{n}{n-2}, \quad n > 2, \tag{8}$$

y

$$Var(X) = \frac{n^2(2m + 2n - 4)}{m(n-2)^2(n-4)}, \quad n > 4. \tag{9}$$

Teorema 5

Si $X \sim F(m, n)$, entonces, $Y = 1/[1 + (m/n)X]$ es $B(n/2, m/2)$.

Consecuentemente, para cada $x > 0$

$$F_X(x) = 1 - F_Y\left[\frac{1}{1 + (m/n)x}\right]$$

Demostración.

Se deja como ejercicio $\square$

Si en la definición (3) hacemos que $X \sim \chi^2(n, \delta)$ se obtiene la variable aleatoria F no-central.

Distribución F no-central

Definición 4

Sea $X \sim \chi^2(n, \delta)$ y $Y \sim \chi^2(n)$, y sean X yY independientes. Entonces la variable aleatoria

$$F = \frac{X/m}{Y/n} \tag{10}$$

se dice que tiene un a distribución-F no-central con grados de libertad (m, n) y parámetro de no-centralidad δ. La función de densidad de la distribución F no-central está definida por

Teorema 6

Sea $X \sim \chi^2(n, \mu)$ y $Y \sim \chi^2(m)$ dos variables aleatorias independientes. Entonces las variable aleatoria $Z = \frac{X/n}{Y/m}$ tiene función de densidad

$$f(z) = e^{-\mu} \sum_{k=0}^{\infty} \frac{\mu^k}{k!} \frac{\left(\frac{n}{m}\right)^{\frac{n+2k}{2}} \Gamma\left(\frac{n+m+2k}{2}\right)}{\Gamma\left(\frac{n+2k}{2}\right)\Gamma\left(\frac{m}{2}\right)} z^{\frac{n+2k}{2}-1} \left(1 + \frac{n}{m} z\right)^{-\frac{n+m+2k}{2}}$$

si $z > 0$ y 0 si $z \leq 0$.

Notación

$Z \sim F(n.m, \mu)$

Si $Z \sim F(n.m, \mu)$ entonces

$$\mathrm{E}(Z) = \frac{n(m+\delta)}{m(n-2)}, \quad n > 2 \text{ y}$$

$$Var(Z) = \frac{2n^2}{m^2(n-4)(n-2)^2}[(m+\delta)^2 + (n-2)(m+2\delta)], \; n > 4.$$

Ejercicios propuestos

1. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de $X \sim \mathcal{N}(\mu, \sigma^2)$ y $\overline{X}$ y S^2, respectivamente, la media muestral y la varianza muestral. Sea $X_{n+1} \sim X \sim \mathcal{N}(\mu, \sigma^2)$, y suponga que $X_1, X_2, \ldots, X_n,\ X_{n-1}$ son independientes. Encuentre la distribución de muestreo de $[\frac{(X_{n+1}-\overline{X})}{S/\sqrt{\mathrm{n}/(n+1)}}]$
2. Sean $X_1, X_2, \ldots, X_\mathrm{m}$ y $Y_1, Y_2, \ldots, Y_n$ muestras aleatorias independientes de $\mathcal{N}(\mu_X, \sigma_X^2)$ y $\mathcal{N}(\mu_Y, \sigma_Y^2)$, respectivamente. Además, sean α y β dos números reales fijos. Si $\overline{X}$, $\overline{Y}$ denotan las medias muestrales correspondientes, ¿cuál es la distribución muestral de
$$\frac{\alpha(\overline{X}-\mu_X)+\beta(\overline{Y}-\mu_Y)}{\sqrt{\frac{(m-1)S_X^2+(n-1)S_Y^2}{m+n-2}}\sqrt{\frac{\alpha^2}{m}+\frac{\beta^2}{n}}}$$

donde S_X^2 y S_Y^2respectivamente denotan las varianzas muestrales de ls $X's$ y de las Y'.

3. Sea $X_1, X_2, \ldots, X_\mathrm{n}$ una muestra aleatoria de $X \sim \mathscr{N}(\mu, \sigma^2)$ y sea k es un número entero entero positivo. Encuentra $\mathrm{E}(S^{2k})$. En particular, encuentre $\mathrm{E}(S^2)$ y $\mathrm{V}ar(S^2)$.
4. Se toma una muestra aleatoria de 5 de una población normal con una media de 2,5 y una varianza $\sigma^2 = 36$.
   (a) Encuentre la probabilidad de que la varianza muestral se encuentre entre 30 y 44.
   (b) Encuentre la probabilidad de que la media muestral se encuentre entre 1.3 y 3.5, mientras que la varianza muestral se encuentra entre 30 y 44.

5. Se observó que la vida media de una muestra de 10 bombillas era de 1327 horas con una desviación estándar de 425 horas. Una segunda muestra de 6 bombillas elegidas de un lote diferente mostró una vida media de 1215 horas con una desviación estándar de 375 horas. Si se supone que las medias de los dos lotes son las mismas, ¿qué tan probable es la observación? diferencia entre las dos medias de muestra?
6. Sean S_X^2 y S_Y^2 las varianzas muestrales de dos muestras independientes de tamaños $n_x = 5$ y $n_Y = 4$ de dos poblaciones que tienen la misma varianza desconocida σ^2. Encuentre (aproximadamente) la probabilidad de que $\frac{S_X^2}{S_Y^2} < 1/5{,}2$ o $> 6{,}25$.

7. Sea $X \sim t(10)$ encuentre $P(|X| > 2{,}228)$
8. Sea $X \sim t(14)$ encuentre b tal que $P(|X| < b) = 0{,}90$
9. Si $X \sim F(5, 10)$, encuentre a y b de modo que $P(X \leq a) = 0{,}05$ y $P(X \leq b) = 0{,}95$ y en consecuencia $P(a < X < b) = 0{,}90$.
10. Sea $X_1, X_2, \ldots, X_5$ una muestra de la $\mathscr{N}(0, 4)$. Encuentre $P\left(\sum_{j=1}^{n} X_j^2 \geq 5{,}75\right)$
11. $X \sim \chi^2(61)$. Encuentra $P(X > 50)$