La distribución F

Ernesto Barrios

3 de julio de 2019

1. Preliminares

La *ley de probabilidades* F es una distribución de probabilidades definida sobre los reales no negativos, resultado, muchas veces, del cociente de dos variables aleatorias independientes distribuidas ji-cuadrada χ^2. Para la deducción de la correspondiente función de densidad de probabilidad (*f. d. p.*) se utilizarán las siguientes dos proposiciones que se presentan sin demostración pero se refieren al texto de Hoerl, Port & Stone (*Introduction to Probability Theory*. Houghton Mifflin Company. Boston. 1971).

Proposición 1. Sea X una variable aleatoria (*v. a.*) continua con *f. d. p.* f_X y sea $k > 0$ una constante, entonces la *f. d. p.* de la *v. a.* $Y = kX$ está dada por

$$f_Y(y) = \frac{1}{k} f_X\left(\frac{1}{k} y\right)$$

Demostración. Se sigue inmediatamente del teorema de transformación, Teorema 1, sección 5.2, HP&S, p.119.

Proposición 2. Sean X_1 y X_2 *v. a.*'s continuas con *f. d. p.* conjunta f. Para la *v. a.* $Y = X_1/X_2$, su *f. d. p.* está dada por

$$f_{\frac{X_1}{X_2}}(y) = \int_{-\infty}^{\infty} |x| f(yx, x) dx$$

Demostración. Vea expresión (22) de la sección 6.2, de HP&S, p.151.

Proposición 3. Sean X_1 y X_2 *v. a.*'a independientes con $X_i \sim \text{Gamma}(\alpha_i, \beta)$, para $i = 1, 2$, entonces

$$f_{\frac{X_1}{X_2}}(y) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{y^{\alpha_1 - 1}}{(1 + y)^{\alpha_1 + \alpha_2}} \mathbb{I}_{(0,\infty)}(y)$$

Demostración. Sea $y > 0$. Entonces, por independencia de las *v. a.*'s $f(x1, x2) =$

$f_1(x_1)f_2(x_2)$, donde f_i es la *f. d. p.* de X_i. Luego,

$$\begin{aligned}
f_{\frac{X_1}{X_2}}(y) &= \int_{-\infty}^{\infty} |x| f(yx, x)dx \\
&= \int_0^{\infty} x f_1(xy) f_2(x) dx \\
&= \int_0^{\infty} x \frac{(xy)^{\alpha_1-1}}{\beta^{\alpha_1}\Gamma(\alpha_1)} e^{-(xy)/\beta} \frac{x^{\alpha_2-1}}{\beta^{\alpha_2}\Gamma(\alpha_2)} e^{-x/\beta} dx \\
&= \frac{y^{\alpha_1-1}}{\beta^{\alpha_1+\alpha_2}\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{1}{K} \int_0^{\infty} K x^{(\alpha_1+\alpha_2)-1} e^{-\frac{1+y}{\beta}x} dx \\
& \qquad\qquad\qquad\qquad \text{con } K = \frac{[(1+y)/\beta]^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)} \\
&= \frac{y^{\alpha_1-1}}{\beta^{\alpha_1+\alpha_2}} \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{\beta^{\alpha_1+\alpha_2}}{(1+y)^{\alpha_1+\alpha_2}} \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \frac{y^{\alpha_1-1}}{(1+y)^{\alpha_1+\alpha_2}}
\end{aligned}$$

donde la última integral vale 1 pues es la integral sobre todo el soporte de una densidad gamma con K la constante normalizadora precisamente.

Proposición 4. Sean $X_i \sim \chi^2_{n_i}$, *v. a.*'s independientes para $i = 1, 2$. Entonces, X_1/X_2 tiene *f. d. p.* dada por

$$f_{\frac{X_1}{X_2}}(x) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} \frac{x^{\frac{n_1}{2}-1}}{(1+x)^{\frac{n_1+n_2}{2}}} \mathbb{I}_{(0,\infty)}(x)$$

Demostración. Recuerde que si $X_i \sim \chi^2_{n_i}$, entonces $X_i \sim \text{Gamma}(n_i/2, 2)$ y aplique la proposición anterior.

Proposición 5. Sean $X \sim \chi^2_m$ y $Y \sim \chi^2_n$ independientes, entonces $W = \dfrac{X/m}{Y/n}$ tiene una *f. d. p.* dada por

$$f_{\frac{X_1}{X_2}}(w) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \left(\frac{m}{n}\right)^{m/2} \frac{w^{\frac{m}{2}-1}}{(1+w)^{\frac{m+n}{2}}} \mathbb{I}_{(0,\infty)}(w) \tag{1}$$

Demostración. Sea $W = kX_1/X_2$ y aplique la proposición 1 con $k = n/m$ al cociente de la proposición anterior.

2. Distribución F

Definición 6. Sea W la *v. a.* con *f. d. p.* dada por la expresión (1). Entonces, W se dice que sigue la *distribución F con m y n grados de libertad* y se denota por $W \sim F_{m,n}$.

Corolario 7. Si $X \sim \chi^2_m$ y $Y \sim \chi^2_n$ independientemente, entonces

$$W = \frac{X/m}{Y/n} \sim F_{m,n}$$

La figura 1 muestra la función de densidad de la distribución F para distintos valores de los grados de libertad. La flexibilidad de la función de densidad es resultado de los *dos parámetros* de la distribución.

El siguiente par de proposiciones muestra algunas propiedades de la distribución F. Su demostración se sigue, considerando que si $W \sim F_{m,n}$, W podría escribirse como $W = \frac{n}{m} XY^{-1}$, producto de *v. a.*'s independientes distribuidas ji-cuadrada con m y n grados de libertad respectivamente.

Proposición 8. Sea $W \sim F_{m,n}$, entonces

i) $\mathbb{E}[W] = \dfrac{n}{n-2}, \quad \text{si } n > 2.$

ii) $\text{var}(W) = \dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \quad \text{si } n > 4.$

Proposición 9. Sea $W \sim F_{m,n}$, entonces

1. $W^{-1} \sim F_{n,m}$.
2. Si $T \sim t_n$ (t-Student con n grados de libertad), entonces $T^2 \sim F_{1,n}$.
3. $\mathbb{P}(W \leq w) = 1 - \mathbb{P}(W^{-1} \leq w^{-1})$.
4. Sea $0 < p < 1$ y $F(p; m, n)$ al p-ésimo percentil de la distribución de W, esto es, $p = \mathbb{P}(W \leq F(p; m, n))$, entonces se tiene que

$$F(p; m, n) = \frac{1}{F(1-p; n, m)}$$

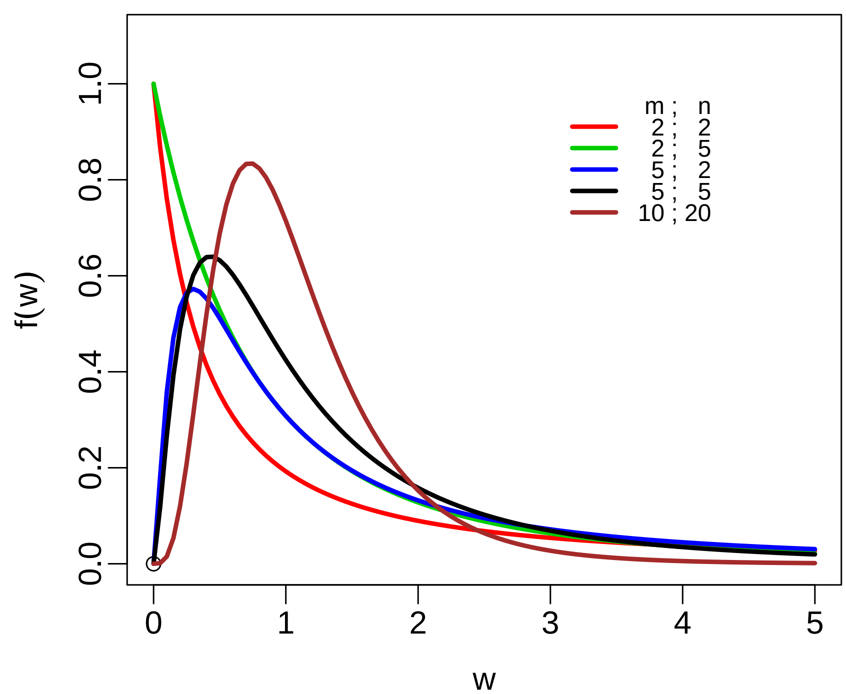

Figura 1: Función de densidad de la distribución F para distintos valores de los grados de libertad m y n.