High-statistics behaviour of power-estimate ratios

Hans H. Brunner and Stefano Bettelli

January 16, 2019

Abstract

The ratio of two quantities that are built as the sum of a certain number of squares of independent standard normal variables is an interesting estimator in power-calibration theory. The statistical behaviour of such estimator is captured by the Fisher-Snedecor probability density function. More formally, if the random variables A and B are distributed as $\mathcal{N}(0, \sigma_A^2)$ and $\mathcal{N}(0, \sigma_B^2)$, respectively, and a certain number of variates a_n and b_n are sampled, then

$$X = \frac{s_A^2/n_A}{s_B^2/n_B} \qquad \text{where} \;\; s_A^2 = \sigma_A^{-2} \sum_n^{n_A} a_n^2 \;\; \text{and} \;\; s_B^2 = \sigma_B^{-2} \sum_n^{n_B} b_n^2$$

is distributed as $F(n_A, n_B)$, wheras s_A^2 and s_B^2 are distributed as $\chi^2(n_A)$ and $\chi^2(n_B)$, respectively. It is well known that for large values of n the distribution of s^2 approximates $\mathcal{N}(n, 2n)$. In this report we show that also X approaches a normal distribution.

The definitions and basic properties of the special functions cited in this report are taken from standard textbooks.[1] The main proof is inspired by a similar work on the beta distribution.[2] We also use the notation $\varphi(z)$ to mean $1 + O(z^{-1})$, with the property $\lim_{z\to\infty} \varphi(z) = 1$.

Theorem (Fisher-Snedecor convergence to the normal distribution).
If a random variable X is distributed according to $F(d_1, d_2)$, then the distribution of the scaled variable $Y = (X - 1)\,\sigma^{-1}$ with $\sigma^2 = \frac{2(d_1+d_2)}{d_1 d_2}$ converges pointwise to $\mathcal{N}(0, 1)$ for $d_1, d_2 \to \infty$.

Corollary 1 (Fisher-Snedecor behaviour for large values of the degrees of freedom).
If a random variable X is distributed according to $F(d_1, d_2)$, then X is approximately distributed as $\mathcal{N}(1, \sigma^2)$ for large values of d_1 and d_2. This follows directly from the main theorem. ■

The treatment of the distribution of Y in the main theorem provides as a byproduct the first-order approximation of the distribution of X as $f_X(x) = \mathcal{N}(1, \sigma^2) \cdot (1 - (x - 1) + O(\sigma^2))$.

Corollary 2 (Fisher-Snedecor behaviour for large and equal values of the degrees of freedom).
If a random variable X is distributed according to $F(d, d)$, then X is approximately distributed as $\mathcal{N}(1, \frac{4}{d})$ for large values of d. Ditto. ■

Corollary 3 (Beta prime behaviour for large and equal values of the degrees of freedom).
If a random variable X is distributed according to $\beta'(\frac{d}{2}, \frac{d}{2})$, then X is approximately distributed as $\mathcal{N}(1, \frac{4}{d})$ for large values of d. This follows from the relation $F(d, d) = \beta'(\frac{d}{2}, \frac{d}{2})$. ■

---

[1] *E.g.*, A. M. Mathai and P. N. Rathie, *Probability and Statistics*, Springer, 1977.
[2] *The beta(b, b) distribution converges to the normal distribution when $b \to \infty$* by Robin Ryder; available at: `http://www.math.wm.edu/~leemis/chart/UDR/PDFs/BetaNormal.pdf`.

Proof. A real positive-definite random variable X distributed according to the F-distribution $F(d_1, d_2)$ with parameters $d_1, d_2 \in \mathbb{R}^+$ has the probability density function

$$f_X(x) = \frac{\Gamma(\frac{1}{2}d_1 + \frac{1}{2}d_2)}{\Gamma(\frac{1}{2}d_1)\Gamma(\frac{1}{2}d_2)} \sqrt{\frac{(xd_1)^{d_1} d_2^{d_2}}{x^2(xd_1 + d_2)^{d_1+d_2}}} \qquad x \in \mathbb{R}^+. \tag{1}$$

Let us introduce the variables $b = \frac{1}{2}d_1 > 0$ and $\rho = d_2/d_1 > 0$; the idea behind them is that b will grow to infinity, while ρ will be considered as a constant. The expression becomes then

$$f_X(x) = \frac{\Gamma((1+\rho)b)}{\Gamma(b)\Gamma(\rho b)} \frac{1}{x} \left(\frac{x \rho^\rho}{(x+\rho)^{1+\rho}} \right)^b . \tag{2}$$

Stirling's formula $\Gamma(z) = \sqrt{\frac{2\pi}{z}} \left(\frac{z}{\mathrm{e}}\right)^z \varphi(z)$ allows to rewrite the normalisation factor as

$$\frac{\Gamma((1+\rho)b)}{\Gamma(b)\Gamma(\rho b)} = \frac{\sqrt{\frac{2\pi}{(1+\rho)b}} \left(\frac{(1+\rho)b}{\mathrm{e}}\right)^{(1+\rho)b}}{\frac{2\pi}{\sqrt{\rho b}} \left(\frac{b}{\mathrm{e}}\right)^b \left(\frac{\rho b}{\mathrm{e}}\right)^{\rho b}} \varphi(b) = \sqrt{\frac{b}{2\pi} \frac{\rho}{1+\rho}} \left(\frac{(1+\rho)^{1+\rho}}{\rho^\rho} \right)^b \varphi(b), \tag{3}$$

and therefore the previous expression as

$$f_X(x) = \frac{\sqrt{b}}{\gamma\sqrt{2\pi}} \left(\frac{1+\rho}{x+\rho} \right)^{b(1+\rho)} x^{b-1} \varphi(b), \qquad \text{where } \gamma = \sqrt{\frac{1+\rho}{\rho}} > 0 \tag{4}$$

Let us now introduce the parameter $\sigma^2 = \frac{\gamma^2}{b} = \frac{2(d_1+d_2)}{d_1 d_2}$ and the normalised random variable Y (note that σ scales with b, hence it will be expanded in the following when taking a limit):

$$Y = \frac{X-1}{\sigma} \in (-\sigma^{-1}, \infty) \qquad \Rightarrow \qquad X = 1 + \sigma Y \quad \text{and} \quad J = \frac{\mathrm{d}X}{\mathrm{d}Y} = \sigma. \tag{5}$$

Y has the probability density function $f_Y(y) = J \cdot f_X(1+\sigma y)$, that is

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \left(1 + \frac{\sigma y}{1+\rho} \right)^{-b(1+\rho)} (1+\sigma y)^{b-1} \varphi(b) \tag{6}$$

The only terms of $f_Y(y)$ that do not have a trivial behaviour when $b \to \infty$ (with γ and ρ constant) are those with a b exponent. Their limit can be calculated with the L'Hôpital rule:

$$\lim_{b\to\infty} \left(\frac{1+\sigma y}{\left(1 + \frac{\sigma y}{(1+\rho)}\right)^{1+\rho}} \right)^b = \exp \lim_{b\to\infty} b \left[\ln(1+\sigma y) - (1+\rho) \ln\left(1 + \frac{\sigma y}{1+\rho}\right) \right]$$
$$= \exp \lim_{b\to\infty} \frac{1}{\frac{\mathrm{d}(b^{-1})}{\mathrm{d}b}} \frac{\mathrm{d}}{\mathrm{d}b} [\ldots] = \exp \lim_{b\to\infty} \frac{-(\gamma y)^2}{2} \frac{\frac{\rho}{1+\rho}}{\varphi(\sqrt{b})} = \exp\left(\frac{-y^2}{2} \frac{\gamma^2 \rho}{1+\rho} \right) = \exp(-\tfrac{1}{2}\, y^2). \tag{7}$$

Inserting the previous result into Eq. 6 one obtains the standard normal distribution.

$$\lim_{b\to\infty} f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}\, y^2) = \mathcal{N}(0,1). \qquad \blacksquare \tag{8}$$

In fact, it is easy to show that the asymptotic expansion of $f_Y(y)$ is dominated by the first-order correction of the term $(1+\sigma y)^{-1}$ in Eq. 6, that is

$$f_Y(y) = \mathcal{N}(0,1) \left(1 - \sigma y + O(\sigma^2)\right) \tag{9}$$
$$\Rightarrow f_X(x) = \mathcal{N}(1,\sigma^2) \left(1 - (x-1) + O(\sigma^2)\right). \tag{10}$$