The 48th International Biometrical Colloquium
in Honour of 90th Birthday of Professor Tadeusz Caliński
and
VI Polish – Portuguese Workshop on Biometry
Szamotuły, Poland, 9 – 13 September 2018

Distribution of the product of two normal variables. A state of the Art

Amílcar Oliveira [2,3] Teresa Oliveira [2,3] Antonio Seijas-Macías [1,3]

[1]Department of Economics. Universidade da Coruña (Spain)

[2]Department of Sciences and Technology. Universidade Aberta (Lisbon), Portugal.

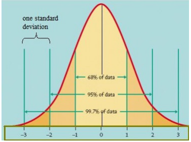
[3]Center of Statistics and Applications, University of Lisbon (Portugal).

September, 2018

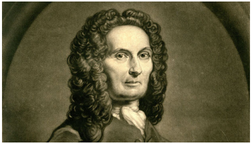
Outline

Introduction

- Normal distribution: the most common in Theory of Probability.
- Applications to the real world: biology, psychology, physics, economics,... .
- Density function (PDF): $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(x-\mu)^2}{2\sigma^2}$,where μ is the mean and σ is the standard deviation (σ^2 is the variance).
- Distribution function (CDF): $F(x) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$, where the error function is: $\mathrm{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-y^2} \mathrm{d}y$.



Normal Distribution $N(0, 1)$



Abraham de Moivre
(1667-1754)

Carl F. Gauss (1777-1855)

Introduction

- Several distributions are derived from normal distribution: Chi-square or t distribution are the most famous.
- Relation with other distributions (exponential, uniform, ...) is known.
- Let X and Y be two normally distributed variables with means μ_x and μ_y and variances σ_x^2, σ_y^2,
- Sum $X + Y$ is normally distributed with mean $\mu_x + \mu_y$ and variance $\sigma_x^2 + \sigma_y^2$, when there is no correlation.
- When there exists correlation (ρ), variance of the sum is $\sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y$.
- The product of two variables was not be able to characterize like the sum and remains like an open problem.

First Historical Approach

- Wishart and Bartlett (1932): The product of two independent normal variables is directly proportional to a second class Bessel function with a zero-order pure imaginary argument [WB32]
- Craig (1936): Let be two normal variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$, and correlation coefficient ρ_{xy} and the inverse of the variation coefficient: $r_x = \frac{\mu_x}{\sigma_x}$ and $r_y = \frac{\mu_y}{\sigma_y}$. Then we could deduce the moment-generating function.[Cra36]

$$M_{xy}(t) = \frac{\exp\left[\frac{(r_x^2+r_y^2-2\rho_{xy}r_x r_y)t^2+2r_x r_y t}{2(1-(1+\rho_{xy})t)(1-(1-\rho_{xy})t))}\right]}{((1-(1+\rho_{xy})t)(1-(1-\rho_{xy})t))^{1/2}} \tag{1}$$

- The product of two normal variables might be a non-normal distribution
- Skewness is $(-2\sqrt{2}, +2\sqrt{2})$, maximum kurtosis value is 12
- The function of density of the product is proportional to a Bessel function and its graph is asymptotical at zero.

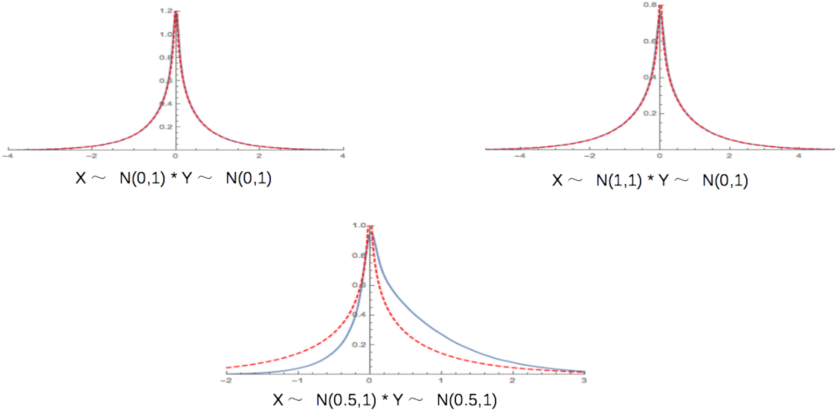


Figure: Examples of the product of two Normal Variables with $\rho = 0$ Craig (red - dashed) and MonteCarlo Simulation (blue)

Advances in 50's in 20th Century

- Aroian (1947): Type III Pearson function or Gram-Charlier Type A series ([Aro47]) .
- Limitations: $\rho = 0$, the Type III Pearson requires $\mu_x \neq 0$ or $\mu_y \neq 0$, Gram-Charlier approach has a very limited range of applicability.
- Advantages: There is no discontinuity at zero.

Theorem ([ATC78], p. 167)

Let X and Y be two normally distributed variables with mean μ_x, μ_y, variances σ_x^2, σ_y^2 and correlation coefficient ρ. Let be $r_x = \frac{\mu_x}{\sigma_x}$ and $r_y = \frac{\mu_y}{\sigma_y}$. Distribution function of $Z = \frac{xy}{\sigma_x \sigma_y}$ is

$$F_Z(z) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \phi(z, r_x, r_y, \rho, t) \mathrm{d}t, \qquad (2)$$

where $\phi(z, r_x, r_y, \rho, t) =$
$\frac{1}{t} \frac{1}{G} \exp\left(-\frac{(H+4\rho r_x r_y)t^2+(1-\rho^2)Ht}{2G^2}\right) * \left\{\left[\left(\frac{G+I}{2}\right)^{1/2} \sin A\right] - \left[\left(\frac{G-I}{2}\right)^{1/2} \cos A\right]\right\}$, *with*
$A = \left(t\left(y - \frac{r_1 r_2 I - \rho H t^2}{G^2}\right)\right)$, $G^2 = (1 + (1-\rho^2)t^2)^2 + 4\rho^2 t^2$, $H = r_1^2 + r_2^2 - 2\rho r_1 r_2$ *and*
$I = 1 + (1-\rho^2)t^2$.

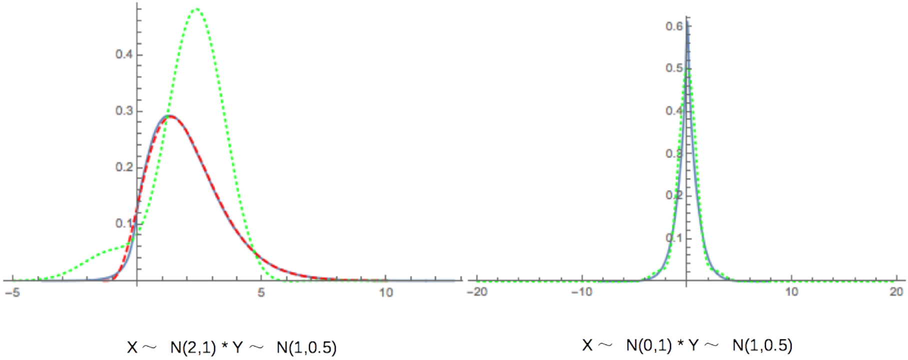


Figure: Examples of the Product of two normal variables no correlated: Gram-Charlier (green - pointed), Pearson Type III (red - dashed) y MonteCarlo simulation (blue)

Rohatgi's Theorem [Roh76]

Theorem ([GLD04], p.452-453)

Let X be a continuous random variable with PDF $f(x)$ definite and positive in (a,b), with $0 < a < b < \infty$. Let Y be a random variable with PDF $g(y)$, definite and positive in (c,d), with $0 < c < d < \infty$. Then, PDF of $Z = XY$ is

- *When $ad < bc$:*

$$h(z) = \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ac < z < ad \\ \int_{z/d}^{z/c} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ad < z < bc \\ \int_{z/d}^{b} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & bc < z < bd \end{cases}$$

- *When $ad = bc$*

$$h(z) = \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ac < z < ad \\ \int_{z/d}^{b} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ad < z < bd \end{cases}$$

- *When $ad > bc$*

$$h(z) = \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ac < z < ad \\ \int_a^{b} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & bc < z < ad \\ \int_{z/d}^{b} g\left(\frac{z}{x}\right) f(x)\frac{1}{x}\,dx & ad < z < bd \end{cases}$$

Application Rohatgi's Theorem

- Only for PDF of random variables in first quadrant, but generalization to other quadrants is straightforward.
- The PDF of the product is not defined at zero.
- Range for normal distribution must be bounded.
- Very good approach for the product of two independent $N(0,1)$ distributions:

$$h(z) = \begin{cases} \frac{K_0(-z)}{\pi} & -\infty < z < 0 \\ \frac{K_0(z)}{\pi} & 0 < z < \infty \end{cases}$$

where $K_0(\cdot)$ is the modified second class Bessel function.

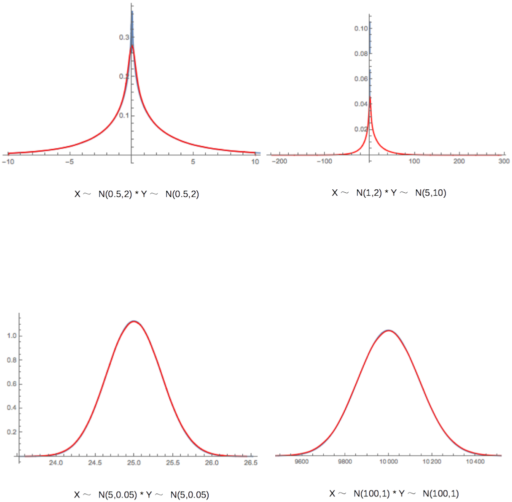


Figure: Examples of Product of two independent normal distributions: Rohatgi's Theorem Approach (blue) and MonteCarlo Simulation (red)

Advances in early 21st Century

Ware and Lad (2003) : A bivariate independent normal distribution [WL03]:

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \right]$$

Marginal density $f(z)$ would be:

$$f(z) = \int_{-\infty}^{\infty} f(z|y)f(y)\mathrm{d}y = \int_{-\infty}^{\infty} f(z,y)\mathrm{d}y \tag{3}$$

- Approach using numerical integration: Newton-Cotes
- Simulation with MonteCarlo method
- Analytical approach using normal distribution: Moment-generating Function:

$$\mu_z = \mu_x\mu_y + \rho\sigma_x\sigma_y \tag{4}$$

$$\sigma_z^2 = \mu_x^2\sigma_y^2 + \mu_y^2\sigma_x^2 + \sigma_x^2\sigma_y^2 + 2\rho\mu_x\mu_y\sigma_x\sigma_y + \rho^2\sigma_x^2\sigma_y^2 \tag{5}$$

For the case of two independent normally distributed variables, the limit distribution of the product is normal. These approach follows the evolution of ratio (mean/standard deviation), but some important questions remain open [WL03]:

- When the ratio mean/standard deviation is enough to guarantee the normal approach for the product.
- Approximation to normality is more sensitive for individual ratios or combined ratio.
- How is the evolution of the skewness of the product, when is null? Is there a level for skewness and normality of product

Approach to the Product of Two Normal Variables

Let X and Y be two variables normales with parameter: μ_x, σ_x^2 and $r_x = \frac{\mu_x}{\sigma_x}$ and μ_y, σ_y^2 and $r_y = \frac{\mu_y}{\sigma_y}$. Then [SMO12] :

- When two variables have unit variance ($\sigma^2 = 1$), with different mean, normal approach is a good option for means greater than 1. But, when the mean is lower, normal approach is not correct.
- When two variables have unit mean ($\mu = 1$), with different variance, normal approach requires that, at least, one variable has a variance lower than 1.
- When, at least, one of the inverse of the variation coefficient δ_x or δ_y is high, then normal approach is correct.
- When two normal distributions have same variance $\sigma_x^2 = \sigma_y^2 = \sigma^2$, we define combined ratio as $\frac{\mu_x \mu_y}{\sigma}$, then a high value for combined ratio produce a good normal approach for product, but when combined ratio is lower than 1, the normal approach fails [OOSM13].

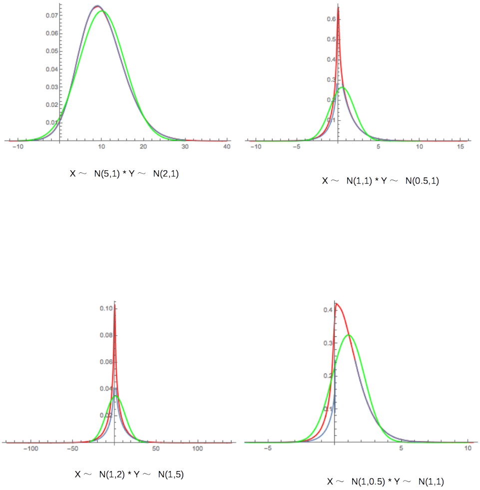


Figure: Examples of Product of two independent normal distributions: Numerical Integration (blue), MonteCarlo Simulation (red), Normal Approach (green)

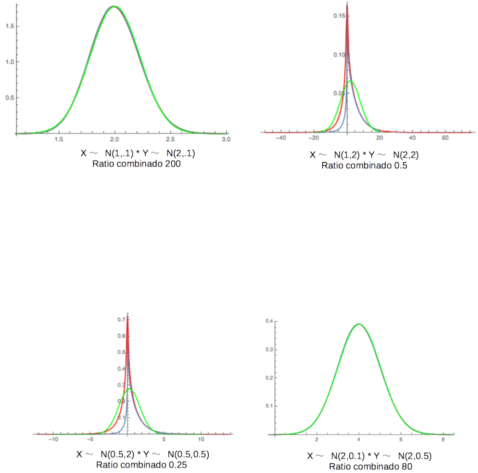


Figure: Examples of Product of two independent normal distributions: Numerical Integration (blue), MonteCarlo Simulation (red), Normal Approach (green)

Recent Publications

Theorem ([NP16] p. 202)

Let (X, Y) be a bivariate normal distribution random vector with mean zero and variance one and correlation coefficient ρ. Then, PDF of $Z = XY$ is

$$f_Z(z) = \frac{1}{\pi\sqrt{1-\rho^2}} \exp\left[\frac{\rho z}{1-\rho^2}\right] K_0\left(\frac{|z|}{1-\rho^2}\right) \tag{6}$$

for $-\infty < z < \infty$, where $K_0(\cdot)$ is second class zero order modified Bessel function.

Recent Publications

Theorem ([Cui+16], pp.1662-1663)

Let X and Y two real Gaussian random variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$ with ρ the correlation coefficient. Then the exact PDF $f_Z(z)$ of the product $Z = XY$ is given by:

$$\exp\left\{-\frac{1}{2(1-\rho^2}\left(\frac{\mu_x^2}{\sigma_x^2}+\frac{\mu_y^2}{\sigma_y^2}-\frac{2\rho(x+\mu_x\mu_y)}{\sigma_x\sigma_y}\right)\right\}$$

$$\times\sum_{n=0}^{\infty}\sum_{m=0}^{2n}\frac{x^{2n-m}|x|^{m-n}\sigma_x^{m-n-1}}{\pi(2n)!(1-\rho^2)^{2n+1/2}\sigma_y^{m-n+1}}\left(\frac{\mu_x}{\sigma_x^2}-\frac{\rho\mu_y}{\sigma_x\sigma_y}\right)^m$$

$$\binom{2n}{m}\times\left(\frac{\mu_y}{\sigma_y^2}-\frac{\rho\mu_x}{\sigma_x\sigma_y}\right)^{2n-m}K_{m-n}\left(\frac{|x|}{(1-\rho^2)\sigma_x\sigma_y}\right)$$

where $K_v(\cdot)$ denotes the modified Bessel function of the second kind and order v.

Final Summary

- First Approaches: Bessel Function - Product of two independent standard normal distributions.
- Moment-generating function of the product
- New Options: Pearson Function Type III - Gram-Charlier Series Type A.
- Rohatgi's Theorem.
- Alternatives approaches:
  - Approach using functions: Bessel, Pearson, Gram-Charlier Series, ...
  - Approach to normal distribution: mean and variance of the product, skewness and kurtosis.
  - Approach using numerical integration methods.
- Future: Alternative distributions: Skew-Normal, Extended Skew-Normal, ...

References: I

[Aro47] Leo A. Aroian. "The probability function of the product of two normally distributed variables". In: *The Annals of Mathematical Statistics* 18.2 (1947), pp. 265–271.

[ATC78] Leo A. Aroian, Vidya S. Taneja, and Larry W. Cornwell. "Mathematical Forms of the Distribution of the Product of Two Normal Variables". In: *Communications in Statistics - Theory and Methodology* A7.2 (1978), pp. 165–172.

[Cra36] Cecil C. Craig. "On the Frequency of the function xy". In: *The Annals of Mathematical Statistics* 7 (1936), pp. 1–15.

[Cui+16] Guolong Cui et al. "Exact Distribution for the Product of Two Correlated Gaussian Random Variables". In: *IEEE Signal Processing Letters* 23.11 (2016), pp. 1662–1666.

[GLD04] Andrew G. Glen, Lawrence M Lemmis, and John H. Drew. "Computing the distribution of the product of two continuous random varialbes". In: *Computational Statistics & Data Analysis* 44 (2004), pp. 451–464.

[NP16] Saralees Nadarajah and Tibor K. Pogány. "On the distribution of the product of correlated normal random variables". In: *Comptes Research de la Academie Sciences Paris, Serie I* 354 (2016), pp. 201–204.

References: II

[OOSM13] A. Oliveira, T. Oliveira, and A. Seijas-Macías. "The influence of ratios and combined ratios on the distribution of the product of two independent gaussian random variables". In: *Proceedings of the 59th World Statistics Congress of the International Statistical Institute*. Ed. by Proceedings of the 59th World Statistics Congress of the International Statistical Institute. International Statistical Institute. 2013.

[Roh76] Vijay K. Rohatgi. *An Introduction to Probability Theory Mathematical Studies*. New York: Wiley, 1976.

[SMO12] Antonio Seijas-Macías and Amílcar Oliveira. "An approach to distribution of the product of two normal variables". In: *Discussiones Mathematicae. Probability and Statistics* 32.1-2 (2012), pp. 87–99.

[WB32] J. Wishart and M. S. Bartlett. "The distribution of second order moment statistics in a normal system". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 28.4 (1932), pp. 455–459.

[WL03] Robert Ware and Frank Lad. *Approximating the Distribution for Sums of Products of Normal Variables*. Tech. rep. The University of Queensland, 2003.

Acknowledgments

This research is based on the partial results of the funded by FCT - Fundação Nacional para a Ciência e Tecnologia, Portugal, through the project UID/MAT/00006/2013.

Thank you for your attention

Email: antonio.smacias@udc.gal