

Article

A New R-Function to Estimate the PDF of the Product of Two Uncorrelated Normal Variables

Antonio Seijas-Macias ^{1,2,*} , Amílcar Oliveira ^{3,†}  and Teresa A. Oliveira ^{2,3,†} 

¹ Departamento de Economía, Faculdade de Economía e Empresa, Universidade da Coruña, 15071 Coruña, Spain

² CEAUL, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal; teresa.oliveira@uab.pt

³ Departamento de Ciência e Tecnologia, Universidade Aberta, 1269-001 Lisboa, Portugal; amilcar.oliveira@uab.pt

* Correspondence: antonio.smacias@udc.gal

† These authors contributed equally to this work.

Abstract: This paper analyses the implementation of a procedure using the software R to calculate the Probability Density Function (PDF) of the product of two uncorrelated Normally Distributed Random Variables. The problem of estimating the distribution of the product of two random variables has been solved for some particular cases, but there is no unique expression for all possible situations. In our study, we chose Rohatgi's theorem as a basis for approximating the product of two uncorrelated Normally Distributed Random Variables. The numerical approximation of the product PDF was calculated using a function that we implemented in R. Several numerical examples show that the approximations obtained in R fit the theoretical values of the product distributions. The results obtained with our R function are very positive when we compare them with the Monte Carlo Simulation of the product of the two variables.

Keywords: Rohatgi's theorem; numerical integration; normal distribution; product normal variables

MSC: 60E05; 62E10; 62E17



Citation: Seijas-Macias, A.; Oliveira, A.; Oliveira, T.A. A New R-Function to Estimate the PDF of the Product of Two Uncorrelated Normal Variables. *Mathematics* **2023**, *11*, 3515. <https://doi.org/10.3390/math11163515>

Academic Editors: M. Filomena Teodoro and Marina Alexandra Pedro Andrade

Received: 7 July 2023

Revised: 4 August 2023

Accepted: 9 August 2023

Published: 14 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The product of two normally distributed variables was studied in the early years of the 20th Century [1–3]. Although there is no exact expression for the product, authors produce several expressions for different situations. The results do not follow the parameters of any known distribution, but there exist some results for some specific cases [2–4]. However, these approaches are only for some specific values since we can only apply them to some cases of the product, and we do not yet have a general expression for that product. In 1981, a book with tables for several cases was edited by AMS [5]. The most recent study of this problem [6] establishes the exact formula for the PDF of the product of two correlated standard normal random variables using an order zero-second kind modified Bessel function.

Another previous work tried to use Rohatgi's theorem [7] to obtain an exact distribution of the product of two uncorrelated Normally Distributed Random Variables; the authors developed an expression of the integral using an infinite series expansion and modified Bessel functions of the second kind [8,9]. Although there are several possibilities to determine the PDF of the product of two normally uncorrelated distributed variables, there is not a closed expression of it. We only have partial approximations [2,3] and some algorithms [10] with limited use of general formulas that we have to calculate numerically [6,8]. More recently, [11–13] proposed a study of the product of zero-mean correlated random variables.

The PDF of the product of two normal variables is still an open problem. The most common approach is using the Monte Carlo Simulation. Our approach is based on the

previous work of [9], but the authors propose a general expression for the product of two random variables, with limited application to normal variables. We have continued this work and we have generalized this approach to the whole product of two uncorrelated normal variables, obtaining a good approximation of the PDF using numerical integration.

We used the analysis development in [9] and we have written a function, using the software R (<https://www.r-project.org/>), accessed 31 July 2023, that calculates the PDF of the product of two normally distributed variables using Rohatgi's theorem. To study the precision of our function, we present the graphics of these PDFs and compare these results with a Monte Carlo Simulation of the product of the same variables.

To the best of our knowledge, the calculation of the PDF of the product of two normal variables is only performed via partial approximations for concrete examples or expressions involving integrals that do not have a closed form. This function, designed for the R program, allows one to quickly and easily calculate the estimated value of the PDF of the product of any two uncorrelated normal variables. Although the formula works well in general, it may have some bugs or errors when the value of the product is very close to zero.

2. The Product of Two Normal Random Variables

Let X and Y be two continuous random variables with joint PDF $f_{XY}(x, y)$. When the independence of the two variables is considered, the PDF of Z can be defined as:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y\left(\frac{z}{x}\right) \frac{1}{|x|} dx. \quad (1)$$

And then, the PDF of Z will be defined on a rectangular product space. In this paper, we consider that X and Y are independent variables.

In [9], a case of Rohatgi's Theorem is considered for determining the PDF of the product of two independent random variables. They describe the product $Z = XY$, where X is a random variable $f(x)$ on the interval (a, b) with $a, b > 0$ and Y is defined as $g(y)$ in an interval (c, d) with $c, d > 0$. Then, Z , X and Y are in the first quadrant. Following this, the theorem is defined.

Theorem 1. Let X be a random variable of continuous type with PDF $f(x)$, which is defined and positive on the interval (a, b) , where $0 < a < b < \infty$. Similarly, let Y be a random variable of the continuous type with PDF $g(y)$, which is defined and positive on the interval (c, d) , where $0 < c < d < \infty$.

The PDF of $Z = XY$ is divided into three sections:

$$\begin{aligned} 1. \quad ad < bc \quad h(z) &= \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ac < z < ad \\ \int_{z/d}^{z/c} g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ad < z < bc \\ \int_{z/d}^b g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & bc < z < bd \end{cases} \\ 2. \quad ad = bc \quad h(z) &= \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ac < z < ad \\ \int_{z/d}^b g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ad < z < bd \end{cases} \\ 3. \quad ad > bc \quad h(z) &= \begin{cases} \int_a^{z/c} g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ac < z < ad \\ \int_a^b g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & bc < z < ad \\ \int_{z/d}^b g\left(\frac{z}{x}\right) f(x) \frac{1}{x} dx & ad < z < bd \end{cases} \end{aligned}$$

Proof. See [9]. $\square$

This theorem can be generalised for variables that lie entirely in the other quadrants. (II, III and IV). When the variable contains the value 0, it is more difficult since some of the rectangular product spaces would lie in two quadrants or not be entirely inside one of them. One solution would be to consider 0 at an endpoint of the interval. This approach results in more efficiency than other approaches. In [14], a formula for calculating the PDF of the product of n uniform independently and identically distributed random variables in

the interval $[0, 1]$ is presented, using mostly Laplace and Merlin transformation techniques. This procedure was improved in [15] to consider non-identically distributed variables. More recently, in [16], the authors use Fourier analysis to produce the PDF of a product of n independent and identically distributed uniform random variables in the interval $[a, b]$.

We developed a function for R [17] that is an implementation of the procedure described in [9] for two normally independently distributed variables, $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$.

We have to adapt the range of the normal distribution to a bounded range to be able to use the function defined in R. Thus, for the normal distribution, we have to consider two limits for the range of the variable. Fortunately, 99% of the range of the normal distribution is within the interval $(\mu - 4\sigma, \mu + 4\sigma)$. When 0 is in the range of some of the two normally distributed variables, we have to adopt a strategy to avoid the value 0, because the Expression (1) is not defined when $x = 0$. We can limit the value of the normal variable to two values near zero; for example, we can use a limit value $\epsilon \approx 0$. The use of numerical integration requires the definition of an interval and a step between two points and then we have to fix a barrier to use values of x too near to zero because they produce an error and misvalue results.

The limits of integration are $N(\mu_x, \sigma_x)$ in the range $\{\mu_x - 4\sigma_x, \mu_x + 4\sigma_x\}$. If this interval is not only positive or negative, we must consider two intervals: $\{\mu_x - 4\sigma_x, -\epsilon\}$ and $\{+\epsilon, \mu_x + 4\sigma_x\}$. If the normal variable is quite concentrated near 0, we can use a scale effect to avoid these values.

We developed a function that calculates the PDF of n points of the product of two normal uncorrelated variables $X \sim N(\mu_x, \sigma_x)$ and $Y \sim N(\mu_y, \sigma_y)$, where $a = \mu_x - 4\sigma_x$, $b = \mu_x + 4\sigma_x$, $c = \mu_y - 4\sigma_y$ and $d = \mu_y + 4\sigma_y$.

The function, named “product.Normal”, has five parameters: the mean and variance of the two normal functions involved in the product (mx, sx, my, sy) and n , that is, the step between two consecutive points. The result of the function is a list of numbers that represents several values of the PDF of the product into the range of the variable $Z = XY$. To calculate the results, we considered six cases (or scenarios) and defined six internal procedures (“product. α normal”) where α is the number of the scenario to be applied. (See Listing 1).

The different scenarios were motivated by the range of values of the two normal variables considered: only positive, only negative or mixed values. In the last case, the value zero is in the range of the normal variable.

Listing 1. A Function to calculate the PDF of the product of two normally uncorrelated distributed variables—R code.

```
product.Normal <- function(n,mx,sx,my,sy){
  a<- mx-4*sx; b<-mx+4*sx; c<-my-4*sy; d<-my+4*sy;
  if (a>0 && b>0 && c>0 && d>0){
    product.1normal(n,a,b,c,d,mx,sx,my,sy)}
  else if (a<0 && b<0 && c<0 && d<0) {
    product.2normal(n,a,b,c,d,mx,sx,my,sy)}
  else if (a>0 && b>0 && c<0 && d<0){
    product.3normal(n,a,b,c,d,mx,sx,my,sy)}
  else if (a>0 && b>0 && c<0 && d>0){
    product.4normal(n,a,b,c,d,mx,sx,my,sy)}
  else if (a<0 && b>0 && c<0 && d>0){
    product.5normal(n,a,b,c,d,mx,sx,my,sy)}
  else {product.6normal(n,a,b,c,d,mx,sx,my,sy)}
}
```

The R code is available in the Supplementary Material.

In this function, we have the following variables: n is the step to consider between two consecutive points (values like 10 or 100 are the most common), mx is the mean of the X variable, my is the mean of the Y variable and sx and sy are the variance of the two normally distributed variables X and Y , respectively.

This function considers the values of the two uncorrelated normal variables in the range $X \in [a, b]$ and $Y \in [c, d]$ and considers the product $Z = XY$ in the range $[li, ls]$, where li and ls are a function of the values of a, b, c, d . The function develops a vector with $\frac{ls-li}{1/n}$ values of the PDF of the product Z .

There are 6 Scenarios:

- Scenario 1—When $a, b, c, d > 0$. In this case, the PDF of the product contains only positive values.
- Scenario 2—When $a, b, c, d < 0$. In this case, the PDF of the product contains only positive values.
- Scenario 3—When $a, b > 0$ and $c, d < 0$. In this case, the PDF of the product contains only negative values.
- Scenario 4—When $a < 0, b > 0, c, d > 0$. The X variable has positive and negative values and contains a value of zero. The Y variable is only positively valued. We have to divide the range of X into two sub-ranges, $(a, -0.001)$ and $(+0.001, b)$, to avoid the zero value since, at this point, the integrals are not defined.
- Scenario 5—When $a < 0, b > 0, c, d < 0$. The X variable has positive and negative values and contains a value of zero. The Y variable is only negatively valued. We have to divide the range of X into two sub-ranges $(a, -0.001)$ and $(+0.001, b)$ to avoid the zero value since at this point the integrals are not defined.
- Scenario 6—When $a < 0, b > 0$ and $c < 0, d > 0$. Then, we have to consider several subranges for X and Y and estimate the values of the PDF of the product considering the following sub-ranges: $(a, -0.001)$, $(+0.001, b)$, $(c, -0.001)$ and $(+0.001, d)$.

Using the values calculated with our function, we could obtain the mean, variance, skewness and kurtosis of the estimated function of the product of the two normal variables.

Although we do not know the distribution function of the product, we can calculate the statistics of the distribution—mean, variance, skewness and kurtosis—using the moment-generating function (MGF) of the product. The MGF of $Z = XY$ is given by

$$\begin{aligned} M_U(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(y-\mu_y)^2}{2\sigma_y^2}\right) \exp(xyt) dx dy \\ &= \frac{1}{(1-\sigma_x^2\sigma_y^2t^2)^{1/2}} \exp\left(\frac{(\sigma_x^2\mu_y^2 + \sigma_y^2\mu_x^2)t^2 + 2t\mu_x\mu_y}{2(1-\sigma_x^2\sigma_y^2t^2)}\right). \end{aligned} \quad (2)$$

Then, we have

- Mean:

$$\mu_x\mu_y \quad (3)$$

- Variance:

$$\mu_y^2\sigma_x^2 + (\mu_x^2 + \sigma_x^2)\sigma_y^2 \quad (4)$$

- Coefficient of skewness:

$$\frac{6\mu_x\mu_y\sigma_x^2\sigma_y^2}{(\mu_y^2\sigma_x^2 + (\mu_x^2 + \sigma_x^2)\sigma_y^2)^{3/2}} \quad (5)$$

- Excess of kurtosis:

$$\frac{6\sigma_x^2\sigma_y^2(2\mu_y^2\sigma_x^2 + (2\mu_x^2 + \sigma_x^2)\sigma_y^2)}{(\mu_y^2\sigma_x^2 + (\mu_x^2 + \sigma_x^2)\sigma_y^2)^2} \quad (6)$$

3. Results

We used the function to estimate the PDF of the product of two uncorrelated normal variables X and Y . The range of each variable was computed as a function of $\mu \pm 4\sigma$. In cases where the variable took the null value, only values of ± 0.01 were considered.

The procedure implemented in R was used to calculate and represent the product density function (red/dark grey). To make a comparison, a Monte Carlo Simulation with 10^6 points of the same two normal variables was used and their density function was plotted on the same graph (green/light grey). As we did not know the shape of the distribution of the product, we could simulate the product using the Monte Carlo Simulation. We considered a large number of products and drew a histogram of the points to obtain an estimation of the shape of the distribution.

A comparative table of the values of the product statistics—mean, variance, skewness and kurtosis—has also been included. Three values were considered: the theoretical values of these statistics, which were calculated using the moment-generating function; the Monte Carlo Simulation, where the values of the statistics were calculated based on the values obtained in the Simulation process; and the values obtained for the density function calculated using Rohatgi's theorem.

3.1. Scenarios 1, 2 and 3

For Scenario 1, we considered three examples of two positive normal variables, with rank (a, b) and (c, d) and three different situations with $a, b, c, d > 0$, with $a < b$ and $c < d$:

Example 1 (a): $ad = bc$: $X \sim N(5, 0.5)$ and $Y \sim N(5, 0.5)$.

Example 1 (b): $ad < bc$: $X \sim N(5, 0.75)$ and $Y \sim N(1, 0.1)$.

Example 1 (c): $ad > bc$: $X \sim N(5, 0.05)$ and $Y \sim N(1, 0.1)$.

The graphical result of Example 1 (a) is in Figure 1. We can observe a very good approach for the Simulation of the product (using the Monte Carlo Simulation) and the estimation of the product using the R function defined in the previous section. The two approaches have the same shape and there is a very strong coincidence between them. This graphic and the following graphics in the paper try to show the coincidence of the shapes of the two approximations, but they are not proof of the goodness of the approximation to the real shape of the product.

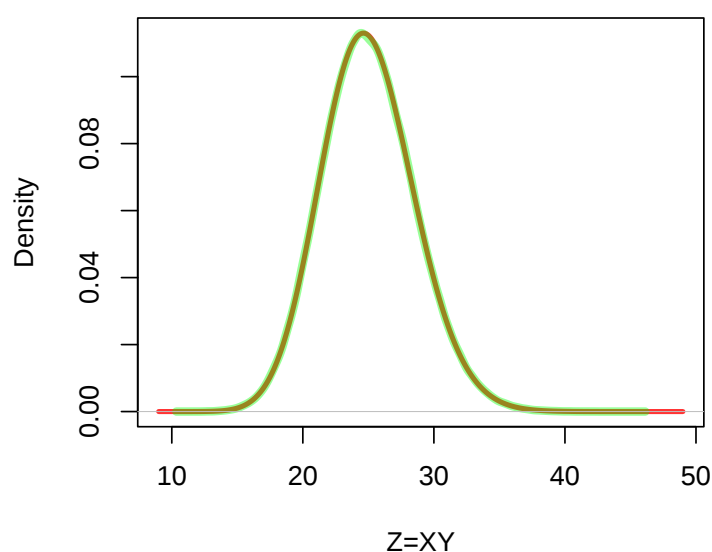


Figure 1. ($X \sim N(5, 0.5)$ and $Y \sim N(5, 0.5)$); Simulation (green/light grey); PDF (red/dark grey).

In Table 1, we compare the values of the statistics mean, variance, skewness and kurtosis (excess) for the theoretical values of the product distribution (first line), Monte

Carlo Simulation (second line) and Rohatgi's Theorem approximation (third line). We added the bias and the relative bias of the numerical approach values and the theoretical values of the product distribution calculated using the expression of the mean, variance, skewness and kurtosis defined in the previous section.

Table 1. Scenario1: $a, b, c, d > 0$ with $a < b$ and $c < d$.

Example 1(a)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	25	12.5625	0.2105	0.0595
MonteCarlo Sim.	25.0036	12.5552	0.2118	0.0666
<i>bias</i>	+0.0036	−0.073	+0.0013	+0.0071
<i>relative bias</i>	+0.011%	−0.061%	+0.612%	+11.932%
R function	25.0000	12.5489	0.2104	0.0521
<i>bias</i>	0	−0.0136	−0.0001	−0.0074
<i>relative bias</i>	0%	−0.108%	−0.047%	−12.433%
Example 1(b)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	5	0.8181	0.2280	0.0822
MonteCarlo Sim.	5.0010	0.8174	0.2307	0.0867
<i>bias</i>	+0.001	−0.0007	+0.0027	+0.0045
<i>relative bias</i>	+0.022%	−0.0851%	+1.183%	+5.475%
R function	5.0000	0.8172	0.2279	0.0736
<i>bias</i>	0	−0.0009	−0.0001	+0.0086
<i>relative bias</i>	0%	−0.113%	−0.0434%	−10.463%
Example 1(c)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	5	0.2525	0.0059	0.0011
MonteCarlo Sim.	5.0001	0.2530	0.0059	0.0038
<i>bias</i>	+0.0001	−0.0005	0	+0.0027
<i>relative bias</i>	+0.002%	+0.198%	0%	245.455%
R function	5.0000	0.2522	0.0059	−0.012
<i>bias</i>	0	−0.0003	0	−0.0131
<i>relative bias</i>	0%	−0.118%	0%	−1190.91%

For Scenario 2, we considered three examples of two negative normal variables, with rank (a, b) and (c, d) and $a, b, c, d < 0$, with $a < b$ and $c < d$:

Example 2 (a): $ad = bc$: $X \sim N(-5, 0.05)$ and $Y \sim N(-5, 0.05)$.

Example 2 (b): $ad > bc$: $X \sim N(-5, 0.75)$ and $Y \sim N(-1, 0.1)$.

Example 2 (c): $ad < bc$: $X \sim N(-3, 0.5)$ and $Y \sim N(-5, 1)$.

In Table 2, we compare the values of the statistics mean, variance, skewness and kurtosis (excess) for the theoretical values of the product distribution (first line), Monte Carlo Simulation (second line) and Rohatgi's Theorem approximation (third line). As we saw in Scenario 1, the values of the mean and variance statistics are in full agreement, and only in the case of skewness and kurtosis are there small differences, which, in general, can be considered as non-significant.

In Scenario 3, we considered two normally distributed variables; one of them was positive with $a, b > 0$ and the other one was negative with $c, d < 0$. For these values, we had $a < b$ and $c < d$. We considered three examples for the following situations: $ac < bd$, $ac > bd$ and $ac = bd$:

Example 3 (a): $ad > bc$ and $ac < bd$: $X \sim N(5, 0.5)$ and $Y \sim N(-2, 0.25)$.

Example 3 (b): $ad > bc$ and $ac = bd$: $X \sim N(5, 1)$ and $Y \sim N(-1, 0.1)$.

Example 3 (c): $ad > bc$ and $ac > bd$: $X \sim N(5, 1)$ and $Y \sim N(-1, 0.05)$.

Table 2. Scenario 2: $a, b, c, d < 0$ with $a < b$ and $c < d$.

Example 2(a)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	25	0.1250	0.0212	0.0005
MonteCarlo Sim.	24.9990	0.1249	0.0203	−0.0027
<i>bias</i>	−0.001	−0.0001	−0.0009	−0.0032
<i>relative bias</i>	−0.004%	−0.08%	−4.245%	−640%
R function	25.0000	0.1248	0.0212	−0.0638
<i>bias</i>	0	−0.0002	0	−0.0643
<i>relative bias</i>	0%	−0.16%	0%	−12860%
Example 2(b)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	5	0.8181	0.2280	0.0822
MonteCarlo Sim.	5.0011	0.8190	0.2251	0.0784
<i>bias</i>	+0.0011	+0.0009	−0.0029	−0.0038
<i>relative bias</i>	+0.022%	+0.110%	−1.271%	−4.622%
R function	5.0000	0.8181	0.2280	0.0821
<i>bias</i>	0	0	0	−0.0001
<i>relative bias</i>	0%	0%	0%	−0.121%
Example 2(c)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	15	15.5	0.3687	0.1919
MonteCarlo Sim.	15.0033	15.5262	0.3696	0.2011
<i>bias</i>	+0.0033	+0.0262	+0.0009	+0.0092
<i>relative bias</i>	+0.022%	+0.169%	+0.244%	+4.794%
R function	14.9998	15.4954	0.3676	0.1857
<i>bias</i>	−0.0002	−0.0046	−0.0011	−0.0062
<i>relative bias</i>	−0.001%	−0.029%	−0.298%	−3.23%

In Table 3, we compare the values of the statistics mean, variance, skewness and kurtosis (excess) for the theoretical values of the product distribution (first line), Monte Carlo Simulation (second line) and Rohatgi's Theorem approximation (third line). The same results as those in Scenarios 1 and 2 are observed.

Table 3. Scenario 3: $a, b > 0$ and $c, d < 0$ with $a < b$ and $c < d$.

Example 3(a)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−10	2.5781	−0.2264	0.0725
MonteCarlo Sim.	−9.9993	2.5835	−0.2302	0.0791
<i>bias</i>	+0.007	+0.0054	−0.0038	+0.0066
<i>relative bias</i>	−0.07%	−0.209%	+1.678%	+9.103%
R function	−10.0000	2.5753	−0.2263	0.0646
<i>bias</i>	0	−0.0028	+0.0001	−0.0079
<i>relative bias</i>	0%	−0.108%	−0.044%	−10.896%
Example 3(b)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−5	1.26	−0.2121	0.0948
MonteCarlo Sim.	−5.0015	1.2628	−0.2165	0.0927
<i>bias</i>	−0.0015	0.0028	−0.0044	−0.0021
<i>relative bias</i>	+0.03%	+0.222%	+2.07%	−2.215%
R function	−5.0000	1.2586	−0.2120	0.0847
<i>bias</i>	0	−0.0014	+0.0001	−0.0101
<i>relative bias</i>	0%	−0.111%	−0.047%	−10.654%

Table 3. Cont.

Example 3(c)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−5	1.065	−0.0682	0.0281
MonteCarlo Sim.	−4.9999	1.0647	−0.0696	0.0277
<i>bias</i>	0.001	−0.0003	−0.0014	−0.0004
<i>relative bias</i>	−0.02%	−0.028%	2.052%	−1.423%
R function	−5.0000	1.0638	−0.0682	0.0155
<i>bias</i>	0	−0.0012	0	−0.0126
<i>relative bias</i>	0%	−0.112%	0%	−44.83%

3.2. Scenario 4 and Scenario 5

In Scenario 4, we had a normally distributed variable with a zero value and range a, b with $a < 0 < b$ and a positive variable with range (c, d) with $c < d$. We considered Example 4 with $X \sim N(1, 0.1)$ and $Y \sim N(0.5, 2)$ (see Figure 2). The graphical result shows a strange result for the estimation of the product for the R function when the value of the product Z is very near zero. In this situation, the estimated value tends to $-\infty$ as a consequence of the instability of the very small values of the variables near zero. The same situation is observed for Scenario 5 (see Figure 3).

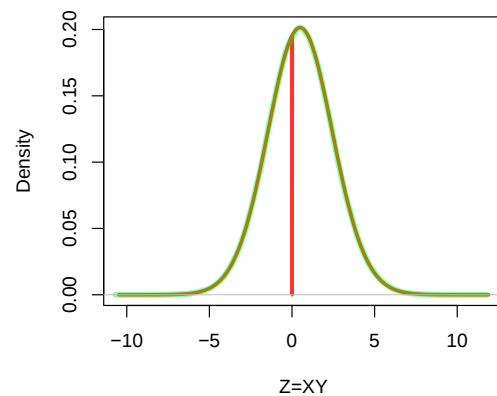


Figure 2. $X \sim N(1, 0.1)$ and $Y \sim N(0.5, 2)$; Simulation (green/light grey); PDF (red/dark grey).

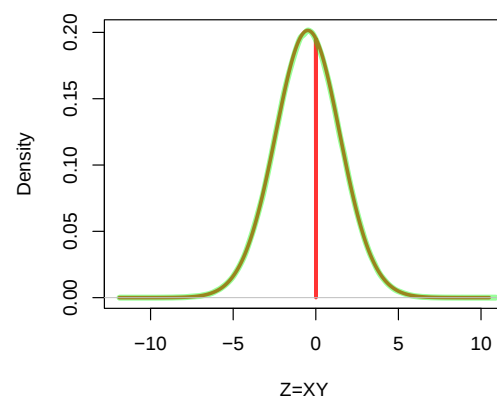


Figure 3. $X \sim N(0.5, 0.2)$ and $Y \sim N(-1, 0.1)$; Simulation (green/light grey); PDF (Red/dark grey).

In this example, there were greater differences than in the previous examples (see Table 4), although except in the case of kurtosis, they do not appear to be significant.

Table 4. Scenario 4: $a < 0 < b$ and $c, d > 0$ with $a < b$ and $c < d$.

Example 4	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	0.5	4.0425	0.0147	0.1181
MonteCarlo Sim.	0.4987	4.0427	0.0125	0.1202
<i>bias</i>	−0.0013	0.0002	−0.0022	0.0021
<i>relative bias</i>	−0.26%	−0.004%	−14.966%	−1.778%
R function	0.5044	4.0596	0.0111	0.0894
<i>bias</i>	+0.0044	+0.4671	−0.0036	−0.0287
<i>relative bias</i>	+0.88%	+11.554%	−24.4898%	−24.301%

In Scenario 5, we had a normally distributed variable with a zero value and range (a, b) with $a < 0 < b$ and a negative variable with range (c, d) and $c < d$. We considered Example 5 with $X \sim N(0.5, 0.2)$ and $Y \sim N(-1, 0.1)$ (see Figure 3). As in Scenario 4, there were greater differences than in the examples in Scenarios 1, 2 and 3 (see Table 5), although they do not appear to be significant.

Table 5. Scenario 5: $a < 0 < b$ and $c, d < 0$ with $a < b$ and $c < d$.

Example 5	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−0.5	4.0425	−0.0147	0.1181
MonteCarlo Sim.	−0.5011	4.0507	−0.0211	0.1214
<i>bias</i>	−0.0011	+0.0082	−0.0064	+0.0033
<i>relative bias</i>	+0.22%	+0.202%	+43.537%	+2.794%
R function	−0.4999	4.0605	−0.0118	0.1061
<i>bias</i>	+0.0001	+0.018	+0.0029	−0.012
<i>relative bias</i>	−0.02%	+0.445%	−19.727%	−10.169%

3.3. Scenario 6: $a < 0 < b$ and $c < 0 < d$ with $a < b$ and $c < d$

When the two normally distributed variables are zero-valued, the application of Rohatgi's theorem is more difficult because it is not defined for zero. In this case, we considered two subintervals for the two normal variables, $(a, -\epsilon)$ and (ϵ, b) , when the original range is (a, b) with $a < 0$ and $b > 0$. We used $\epsilon = 0.01$. We had nine situations:

Example 6 (a): $ac > bd$ and $ad > bc$: $X \sim N(0.5, 2)$ and $Y \sim N(-1, 0.5)$.

Example 6 (b): $ac > bd$ and $ad < bc$: $X \sim N(-1, 0.5)$ and $Y \sim N(-0.5, 0.5)$.

Example 6 (c): $ac > bd$ and $ad = bc$: $X \sim N(-1, 1)$ and $Y \sim N(-2, 2)$.

Example 6 (d): $ac < bd$ and $ad > bc$: $X \sim N(1, 2)$ and $Y \sim N(0, 2)$.

Example 6 (e): $ac < bd$ and $ad < bc$: $X \sim N(1, 2)$ and $Y \sim N(0.5, 0.5)$.

Example 6 (f): $ac < bd$ and $ad = bc$: $X \sim N(1, 2)$ and $Y \sim N(5, 10)$.

Example 6 (g): $ac = bd$ and $ad > bc$: $X \sim N(2, 2)$ and $Y \sim N(-0.5, 0.5)$.

Example 6 (h): $ac = bd$ and $ad < bc$: $X \sim N(-2, 2)$ and $Y \sim N(1, 1)$.

Example 6 (i): $ac = bd$ and $ad < bc$: $X \sim N(0, 2)$ and $Y \sim N(0, 1)$.

In Tables 6 and 7, we compare the values of the statistics. We can observe some significant differences between the theoretical values and the PDF function values. The differences are greater for skewness and kurtosis because of the strange results of the PDF function estimation when the value of the variable is near zero.

Table 6. Scenario 6: $a < 0 < b$ and $c < 0 < d$ with $a < b$ and $c < d$. Cases (a)–(f).

Example 6(a)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−0.5	5.0625	−0.2633	2.1362
MonteCarlo Sim.	−0.4980	5.0762	−0.2696	2.1206
<i>bias</i>	+0.002	+0.0137	−0.3344	−0.0156
<i>relative bias</i>	−0.4%	+0.270%	+2.392%	−0.730%
R function	−0.5140	5.0646	−0.3344	1.7838
<i>bias</i>	−0.014	0.0021	−0.0711	−0.3524
<i>relative bias</i>	+2.8%	+0.0414%	+27.003%	−16.496%
Example 6(b)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	0.5	0.375	0.8164	1.8333
MonteCarlo Sim.	0.5011	0.3749	0.8152	1.8057
<i>bias</i>	+0.0011	−0.0001	−0.0012	−0.0276
<i>relative bias</i>	+0.22%	−0.026%	−0.146%	−1.505%
R function	0.5022	0.3695	0.7384	1.4312
<i>bias</i>	+0.0022	−0.0055	−0.078	−0.4021
<i>relative bias</i>	+0.44%	−1.466%	−9.554%	−21.933%
Example 6(c)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	2	12	1.1547	3.3333
MonteCarlo Sim.	1.9995	12.0378	1.1686	3.4367
<i>bias</i>	−0.0005	+0.0378	+0.0139	+0.1034
<i>relative bias</i>	−0.025%	+0.315%	+1.203%	+3.102%
R function	1.9881	10.7059	0.9704	2.7137
<i>bias</i>	−0.0119	−1.2941	−0.1843	−0.6196
<i>relative bias</i>	−0.595%	−10.784%	−15.960%	−18.588%
Example 6(d)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	0	20	0	5.76
MonteCarlo Sim.	0.0011	19.960	−0.0075	5.8143
<i>bias</i>	+0.0011	−0.04	−0.0075	+0.0543
<i>relative bias</i>	− − %	−0.2%	− − %	+0.942%
R function	0.3783	21.8164	−0.0706	5.2635
<i>bias</i>	+0.3783	+1.8164	−0.0706	−0.4965
<i>relative bias</i>	− − %	+9.082%	− − %	−8.619%
Example 6(e)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	0.5	2.25	0.8889	4.1481
MonteCarlo Sim.	0.4999	2.2386	0.8842	4.0750
<i>bias</i>	−0.0001	−0.0114	−0.0047	−0.0731
<i>relative bias</i>	−0.02%	−0.506%	−0.527%	−1.762%
R function	0.5288	2.4004	0.8184	3.6308
<i>bias</i>	0.0288	0.1504	−0.0705	−0.5173
<i>relative bias</i>	+5.76%	+6.684%	−7.931%	−12.4708%
Example 6(f)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	5	600	0.8164	5.3333
MonteCarlo Sim.	5.0181	601.7577	0.8141	5.3873
<i>bias</i>	+0.0181	+1.7577	−0.0023	+0.054
<i>relative bias</i>	+0.362%	+0.292%	−0.281%	+1.012%
R function	9.9677	590.6478	0.9831	5.3607
<i>bias</i>	+4.9677	−9.3522	+0.1667	+0.0274
<i>relative bias</i>	+99.354%	−1.558%	+20.418%	+0.513%

Table 7. Scenario 6 (cont.): $a < 0 < b$ and $c < 0 < d$ with $a < b$ and $c < d$. Cases (g)–(i).

Example 6(g)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−1	3	−1.1547	3.3333
MonteCarlo Sim.	−0.9970	2.9958	−1.1529	3.3334
<i>bias</i>	+0.003	−0.0042	+0.0018	+0.0001
<i>relative bias</i>	−0.3%	−0.14%	−0.155%	+0.003%
R function	−1.0418	3.1180	−1.0821	3.0250
<i>bias</i>	−0.0418	+0.118	+0.0726	−0.308
<i>relative bias</i>	+4.18%	+3.933%	6.287%	−9.249%
Example 6(h)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	−2	12	−1.1547	3.3333
MonteCarlo Sim.	−2.0027	12.0018	−1.1482	3.3156
<i>bias</i>	−0.0027	+0.0018	−0.0065	−0.0177
<i>relative bias</i>	+0.135%	+0.015%	−0.562%	−1.787%
R function	−2.1344	10.0343	−0.9256	1.5456
<i>bias</i>	−0.1344	−1.957	−2.0803	−1.7877
<i>relative bias</i>	+6.72%	+16.381%	−180.159%	−53.6315%
Example 6(i)	Mean	Variance	Skewness	Kurtosis
Theoretical Val.	0	4	0	6
MonteCarlo Sim.	0.0033	3.9818	0.0012	5.9445
<i>bias</i>	+0.0033	−0.0182	+0.0012	−0.0555
<i>relative bias</i>	− − %	−0.0923%	− − %	−0.8812%
R function	0.0000	3.9077	0.0000	5.1188
<i>bias</i>	0	−0.0923	0	−0.8812
<i>relative bias</i>	− − %	−2.307%	− − %	−14.6867%

The examples in which the variables take null values show much worse results than in the cases analysed above. The greatest divergences occur in the value of kurtosis. When the value of the variable approaches zero, it presents density values very different from those obtained through the Simulation processes and the theoretical values.

4. Discussion

This work has focused on the estimation of the probability density function (PDF) of the product of two normal functions without correlation. Using Rohatgi's theorem, we can calculate the value of this function using an integral approximation. Unfortunately, the integral does not have an analytical expression, so it is necessary to use numerical integration to approximate the values of the integral at specific intervals. We have developed a function in R to estimate the PDF of the product of these two normal variables and obtain a vector of values.

As an example, we used a graphical representation of the product density function obtained with the values reported by the R function and we compared it with the probability density function of a Monte Carlo Simulation of the product of two variables. When we had two normally distributed variables without a zero value in their range, the two approximations showed a very high degree of agreement.

On the other hand, we also compared the values of the following distribution statistics: mean, variance, skewness and kurtosis. In this case, we used the following three values:

- (i) The values of the theoretical distribution of the product. These values were obtained through the moment-generating function.
- (ii) The values obtained in the Simulation process.
- (iii) The values of the estimation of the distribution through the R function.

In general, the results were very close in all the cases analysed. We only observed some small significant differences in those cases where the normal variables incorporated the null value in their range of variation.

In summary, the new R function presents an alternative solution for estimating the PDF of the product of two uncorrelated normal variables. Its implementation incorporates advanced statistical techniques and leverages the capabilities of the R programming language, making it an invaluable tool for researchers and analysts seeking precise and reliable results in complex statistical analyses.

Future work should focus on analysing the observed differences and try to approximate the estimated value of the density function for those situations where the normally distributed variable reaches zero. On the other hand, a more ambitious goal would be to generalise Rohatgi's theorem to the case of variables with a non-zero correlation. Moreover, another future challenge is to explore similar procedures for calculating the PDFs of some other non-Gaussian distributions.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/math11163515/s1>.

Author Contributions: Conceptualisation, A.O.; methodology, A.S.-M., A.O. and T.A.O.; software, A.S.-M.; validation, A.S.-M., A.O. and T.A.O.; formal analysis, A.S.-M., A.O. and T.A.O.; writing—original draft preparation, A.S.-M.; writing—review and editing, A.O. and T.A.O.; supervision, A.O. and T.A.O.; funding acquisition, A.S.-M., A.O. and T.A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by FCT—Fundação para a Ciência e a Tecnologia under the project—UIBD/00006/2020 and this research was partially funded by FUAC—Fundación Univesidade da Coruña by special funds of the Cátedra de Internacionalización Luis Fernández Somoza.

Acknowledgments: The authors are especially grateful for the work of the anonymous reviewers of this paper, whose suggestions and comments have made a substantial contribution to the improvement of it.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

PDF Probability Density Function
MGF Moment-Generating Function
AMS American Mathematical Society

References

1. Wishart, J.; Bartlett, M.S. The distribution of second order moment statistics in a normal system. *Math. Proc. Camb. Philos. Soc.* **1932**, *28*, 455–459. [[CrossRef](#)]
2. Craig, C.C. On the Frequency of the function xy . *Ann. Math. Stat.* **1936**, *7*, 1–15. [[CrossRef](#)]
3. Aroian, L.A. The probability function of the product of two normally distributed variables. *Ann. Math. Stat.* **1947**, *18*, 265–271. [[CrossRef](#)]
4. Aroian, L.A.; Taneja, V.S.; Cornwell, L.W. Mathematical Forms of the Distribution of the Product of Two Normal Variables. *Commun. Stat. Theory Methods* **1978**, *7*, 165–172. [[CrossRef](#)]
5. Meeker, W.Q.; Cornwell, L.W.; Aroian, L.A. *The Product of Two Normally Distributed Random Variables*; Number 7 in Selected Tables in Mathematical Statistics; American Mathematical Soc.: Providence, RI, USA, 1981.
6. Nadarajah, S.; Pogány, T.K. On the distribution of the product of correlated normal random variables. *Comptes Rendus Math.* **2016**, *354*, 201–204. [[CrossRef](#)]
7. Rohatgi, V.K. *An Introduction to Probability and Statistics*; Wiley: New York, NY, USA, 1976.
8. Cui, G.; Yu, X.; Iommelli, S.; Kong, L. Exact Distribution for the Product of Two Correlated Gaussian Random Variables. *IEEE Signal Process. Lett.* **2016**, *23*, 1662–1666. [[CrossRef](#)]
9. Glen, A.G.; Leemis, L.M.; Drew, J.H. Computing the distribution of the product of two continuous random variables. *Comput. Stat. Data Anal.* **2004**, *44*, 451–464. [[CrossRef](#)]
10. Meeker, W.Q.; Escobar, L.A. An algorithm to compute the CDF of the product of two normal random variables. *Commun.-Stat.-Simul. Comput.* **1994**, *23*, 271–280. [[CrossRef](#)]

11. Gaunt, R.E. A note on the distribution of the product of zero-mean correlated normal random variables. *Stat. Neerl.* **2019**, *73*, 176–179. [\[CrossRef\]](#)
12. Gaunt, R.E. Stein's method and the distribution of the product of zero mean correlated normal random variables. *Commun.-Stat.-Theory Methods* **2021**, *50*, 280–285. [\[CrossRef\]](#)
13. Gaunt, R.E. The basic distributional theory for the product of zero mean correlated normal random variables. *Stat. Neerl.* **2022**, *76*, 450–470. [\[CrossRef\]](#)
14. Springer, M. *The Algebra of Random Variables*; Wiley: New York, NY, USA, 1979.
15. Ishihara, T. The distribution of the sum and the product of independent uniform random variables distributed at different intervals. *Trans.-Jpn. Soc. Ind. Appl. Math.* **2002**, *12*, 197–208.
16. Dettmann, C.P.; Georgiou, O. Product of n independent uniform random variables. *Stat. Probab. Lett.* **2009**, *79*, 2501–2503. [\[CrossRef\]](#)
17. R Core Team. R Foundation for Statistical Computing. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.