Economics Bulletin

Volume 38, Issue 1

A note on the absolute moments of the bivariate normal distribution

Markus Haas
Institute for Quantitative Business and Economics Research (QBER), University of Kiel

Abstract

A short and simple calculation of the expected absolute value of the product of two correlated zero-mean normal variables is provided.

I thank an anonymous referee for thoughtful comments and suggestions. Financial support by the German Research Foundation (DFG) via grant HA 5391/2-1 is gratefully acknowledged.
Citation: Markus Haas, (2018) "A note on the absolute moments of the bivariate normal distribution", *Economics Bulletin*, Volume 38, Issue 1, pages 650-656
Contact: Markus Haas - markus.haas@qber.uni-kiel.de.
Submitted: June 21, 2017. **Published:** March 23, 2018.

1. Introduction

Consider random variables X and Y following a bivariate standard normal distribution with correlation ρ, with joint density (pdf)

$$\phi(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} \quad (1)$$

$$= \phi(x)\frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right), \quad (2)$$

where ϕ with a single argument denotes the univariate standard normal pdf. The corresponding cdf will be denoted by Φ, i.e. $\Phi(z) = \int_{-\infty}^{z} \phi(\xi) d\xi$.

It is known that the expectation of the product of the absolute values of X and Y is

$$\mathrm{E}\,|XY| = \frac{2}{\pi}(\rho \arcsin \rho + \sqrt{1-\rho^2}). \quad (3)$$

A special case of formula (3) dates back to Helmert (1876), who was concerned with calculating the variance of the mean absolute deviation $D = n^{-1}\sum_{i=1}^{n} |X_i - \overline{X}|$ for a random sample from a normal population, where $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$. Assuming unit variance, variables $Z_i = X_i - \overline{X}$ are multivariate normal with mean zero, variance $(n-1)/n$, and $\mathrm{Corr}(Z_i, Z_j) = -1/(n-1)$ for $i \neq j$, and so

$$\mathrm{E}(D^2) = \frac{1}{n^2}\left[n\,\mathrm{E}(Z_1^2) + n(n-1)\,\mathrm{E}(|Z_1 Z_2|)\right]$$

$$= \frac{n-1}{n^2}\left[1 + \frac{2}{\pi}\left(\arcsin\left(\frac{1}{n-1}\right) + \sqrt{n(n-2)}\right)\right].$$

In a more recent application in financial econometrics, the expected absolute value of the product of two zero–mean bivariate normal variables is required to compute the covariance matrix forecasts implied by a certain multivariate GARCH process (Pelletier, 2006; see Section 3 for a short account).

Nabeya (1951) devised a technique for calculating moments of the form $\mathrm{E}\,|X^m Y^n|$ for non–negative integers m and n. The calculations involved are rather cumbersome in particular when both m and n are odd, as in (3). Kamat (1953) suggested a method to compute the incomplete moments

$$I_{m,n}(\rho) = \int_0^\infty \int_0^\infty x^m y^n \phi(x, y; \rho) dy dx, \quad (4)$$

from which the absolute moments can be obtained via

$$\mathrm{E}\,|X^m Y^n| = 2(I_{m,n}(\rho) + I_{m,n}(-\rho)). \quad (5)$$

Expression (3) was also obtained by Li and Wei (2009) as a special case of a general formula for the expectation of the absolute value of a Gaussian quadratic form, writing XY as $\frac{1}{2}(X, Y)\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)(X, Y)'$. An alternative (infinite series) representation of $\mathrm{E}\,|XY|$

was derived by Wellner and Smythe (2002) by writing $\mathrm{E}\,|XY| = \mathrm{E}\{|X|\,\mathrm{E}(|Y|\,|X\,)\}$ and then using the fact that Y^2 given X has a noncentral χ^2 distribution.

In this note, an alternative derivation of (3) is provided which is very short and simple. Like the derivation in Wellner and Smythe (2002), it is based on conditioning, but rather than using the conditional noncentral χ^2 distribution of Y^2, it combines Kamat's (1953) use of (4) and (5) with conditional normality of Y.

2. Computation of $\mathrm{E}\,|XY|$

In the derivation, we use the well–known quadrant probability

$$\Pr(X>0, Y>0) \overset{(11)}{=} \int_0^\infty \phi(x)\Phi\left(\frac{\rho x}{\sqrt{1-\rho^2}}\right)dx = \frac{\arcsin\rho}{2\pi} + \frac{1}{4}. \tag{6}$$

For completeness, a straightforward derivation of (6) as in Owen (1956) is reproduced in the appendix. We also use the fact that $\phi'(z) = -z\phi(z)$ and the basic symmetry relations $\phi(z) = \phi(-z)$ and $\Phi(z) = 1 - \Phi(-z)$.

From (2), and using the notation in (4),

$$I_{1,1}(\rho) = \int_0^\infty x\phi(x)\int_0^\infty \frac{y}{\sqrt{1-\rho^2}}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)dydx,$$

where the inner integral becomes, upon substituting $z = (y-\rho x)/\sqrt{1-\rho^2}$, and defining

$$\gamma = \rho/\sqrt{1-\rho^2}, \tag{7}$$

$$\begin{aligned}\int_0^\infty \frac{y}{\sqrt{1-\rho^2}}\phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)dy &= \int_{-\gamma x}^\infty (\sqrt{1-\rho^2}z + \rho x)\phi(z)dz \\ &= \sqrt{1-\rho^2}\phi(\gamma x) + \rho x\Phi(\gamma x).\end{aligned}$$

The derivative of the function

$$\varphi(x) = \sqrt{1-\rho^2}\phi(\gamma x) + \rho x\Phi(\gamma x)$$

is

$$\varphi'(x) = -\sqrt{1-\rho^2}\gamma^2 x\phi(\gamma x) + \rho\Phi(\gamma x) + \rho\gamma x\phi(\gamma x) = \rho\Phi(\gamma x).$$

Since $\varphi(0) = \sqrt{1-\rho^2}\phi(0) = \sqrt{1-\rho^2}/\sqrt{2\pi}$, integration by parts shows that

$$\begin{aligned}I_{1,1}(\rho) &= \int_0^\infty \varphi(x)x\phi(x)dx = -\varphi(x)\phi(x)|_0^\infty + \int_0^\infty \phi(x)\varphi'(x)dx \\ &= \varphi(0)\phi(0) + \rho\int_0^\infty \phi(x)\Phi(\gamma x)dx \\ &\overset{(6)}{=} \frac{\sqrt{1-\rho^2}}{2\pi} + \frac{\rho\arcsin\rho}{2\pi} + \frac{\rho}{4}.\end{aligned} \tag{8}$$

Finally, since $\arcsin(-\rho) = -\arcsin(\rho)$, and using (5), we get (3).

Remark 1 *Probability (6) is used in (8) to compute the integral* $\int_0^\infty \phi(x)\Phi(\gamma x)dx$. *Alternatively, a perhaps still more straightforward derivation of this integral is via Azzalini's (1985) skew–normal (SN) distribution, which has density*

$$f_{SN}(z;\gamma) = 2\phi(z)\Phi(\gamma z), \quad \gamma \in \mathbb{R},$$

where γ *is the skewness parameter and determines the degree of asymmetry of the density. It is straightforward to check by differentiation that the cdf is (Azzalini, 1985)*

$$F_{SN}(z;\gamma) = \Phi(z) - \frac{1}{\pi}\int_0^\gamma \frac{\exp\left\{-\frac{z^2}{2}(1+x^2)\right\}}{1+x^2}dx,$$

and so

$$F_{SN}(0;\gamma) = \frac{1}{2} - \frac{1}{\pi}\int_0^\gamma \frac{dx}{1+x^2} = \frac{1}{2} - \frac{\arctan\gamma}{\pi} \overset{(7)}{=} \frac{1}{2} - \frac{\arcsin\rho}{\pi},$$

and

$$\int_0^\infty \phi(z)\Phi(\gamma z)dz = (1 - F_{SN}(0;\gamma))/2 = \frac{1}{4} + \frac{\arcsin\rho}{2\pi}.$$

3. Financial applications and nonnormal distributions

Many financial markets are characterized by volatility clustering, i.e. alternating periods of low and high volatility. This phenomenon is often modeled by specifying a generalized autoregressive conditional heteroskedasticity (GARCH) model for the unexpected shocks to financial returns.

One of the first multivariate GARCH models was the constant conditional correlation (CCC) GARCH model of Bollerslev (1990), which combines time–varying conditional variances with a constant conditional correlation matrix, $\boldsymbol{R}$. An N–dimensional time series $\{\boldsymbol{\epsilon}_t\} = \{(\epsilon_{1t}, \ldots, \epsilon_{Nt})'\}$ generated by a CCC can be written as

$$\boldsymbol{\epsilon}_t = \boldsymbol{D}_t \boldsymbol{z}_t, \tag{9}$$

where $\{\boldsymbol{z}_t\} = \{(z_{1t}, \ldots, z_{Nt})'\}$ is an iid series of innovations with zero mean and covariance matrix $\boldsymbol{R} = [\rho_{ij}]_{i,j=1,\ldots,N}$ such that $\rho_{ii} = 1$, $i = 1, \ldots, N$. Furthermore, $\boldsymbol{D}_t = \mathrm{diag}(h_{1t}, \ldots, h_{Nt})$, where h_{it} is asset i's conditional standard deviation, $i = 1, \ldots, N$.

In principle, any suitable volatility model can be used to describe the dynamics of the conditional standard deviations h_{it}. However, as observed by Pelletier (2006), closed–form multi–step covariance matrix forecasts can be calculated when the volatility dynamics are specified as an absolute value GARCH (AVGARCH) process,[1] the simplest

[1] Use of the AVGARCH model of Taylor (1986) instead of Bollerslev's (1986) specification in terms of the variances and squared lagged shocks also appears to improve the fit in particular for stock returns (e.g., Giot and Laurent, 2003; and Lejeune, 2009).

form of which is[2]

$$\begin{aligned} h_{it} &= \omega_i + \alpha_i|\epsilon_{i,t-1}| + \beta_i h_{i,t-1} = \omega_i + c_{i,t-1}h_{i,t-1}, \\ \omega_i &> 0, \quad \alpha_i, \beta_i \geq 0, \quad i = 1, \ldots, N, \end{aligned}$$

where $c_{it} = \alpha_i|z_{it}| + \beta_i$. With E_t denoting an expectation conditional on the information up to time t, the τ–step conditional covariance between assets i and j is

$$\begin{aligned} h_{ij,t}(\tau) &:= \mathrm{E}_t(\epsilon_{i,t+\tau}\epsilon_{j,t+\tau}) \\ &= \mathrm{E}(z_{i,t+\tau}z_{j,t+\tau})\,\mathrm{E}_t(h_{i,t+\tau}h_{j,t+\tau}) \\ &= \rho_{ij}\,\mathrm{E}_t(h_{i,t+\tau}h_{j,t+\tau}) \\ &= \rho_{ij}\,\mathrm{E}_t[(\omega_i + c_{i,t+\tau-1}h_{i,t+\tau-1})(\omega_j + c_{j,t+\tau-1}h_{j,t+\tau-1})] \\ &= \rho_{ij}(\omega_i\omega_j + \omega_i c_j h_{jt}(\tau-1) + \omega_j c_i h_{it}(\tau-1)) + c_{ij}h_{ij,t}(\tau-1), \end{aligned} \quad (10)$$

where $c_i = \mathrm{E}(c_{it}) = \alpha_i\,\mathrm{E}(|z_{it}|) + \beta_i$, $h_{it}(\tau) = \mathrm{E}_t(h_{i,t+\tau}) = \omega_i(1-c_i^{\tau-1})/(1-c_i) + c_i^{\tau-1}h_{i,t+1}$, and $c_{ij} = \mathrm{E}(c_{it}c_{jt}) = \alpha_i\alpha_j\,\mathrm{E}\,|z_{it}z_{jt}| + \alpha_i\beta_j\,\mathrm{E}\,|z_{it}| + \alpha_j\beta_i\,\mathrm{E}\,|z_{jt}| + \beta_i\beta_j$, which involves $\mathrm{E}\,|z_{it}z_{jt}|$. Since $h_{i,t+1}$ and $h_{j,t+1}$ are determined by the information up to time t, recursion (10) provides closed–form multi–step conditional covariances.

Note that we don't have to assume multivariate normality of $\boldsymbol{z}_t$ in (9) to explicitly calculate $c_{ij} = \mathrm{E}(c_{it}c_{jt})$ in (10). In fact, due to the leptokurtic shape of most asset return distributions, it is typically more suitable to assume a thicker tailed distribution, such as the Student's t distribution. The t distribution belongs to the class of *normal variance mixtures* (e.g., McNeil et al., 2015, Sec. 6.2), which can be written as $\boldsymbol{z}_t = \sqrt{u_t}\boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}, \boldsymbol{R})$, and $\{u_t\}$ is a non–negative scalar-valued iid random sequence independent of $\{\boldsymbol{\eta}_t\}$.[3] Result (3) then clearly continues to hold for this class of distributions provided that u_t has finite expectation (scaled to unity).

A major drawback of the CCC is the assumption of constant conditional correlations, which is often rejected in applications to financial markets. To overcome this drawback, Pelletier (2006) allows the conditional correlation to change according to a Markov–switching process. The resulting regime–switching model for dynamic correlations (RSDC) still admits the calculation of closed–form covariance matrix forecasts (Pelletier, 2006; Haas, 2010), and it appears to perform well in empirical applications (e.g., Pelletier, 2006; Giamouridis and Vrontos, 2007; Haas, 2010; and Charlot et al., 2016).

References

Azzalini, A. (1985) "A Class of Distributions which Includes the Normal Ones" *Scandinavian Journal of Statistics* **12**, 171-178.

[2] More flexible GARCH specifications (e.g., allowing for an asymmetric response of volatility to positive and negative shocks) can be considered, such as the class of AVGARCH models investigated in He and Teräsvirta (1999) and Ling and McAleer (2002). An extension of the CCC to allow for feedback between the conditional volatilities is considered in Jeantheau (1998), Ling and McAleer (2003), He and Teräsvirta (2004), Conrad and Karanasos (2010), and Francq and Zakoïan (2012).

[3] E.g., the multivariate t distribution is obtained when u_t has an inverse gamma distribution.

Bollerslev, T. (1986) “Generalized Autoregressive Conditional Heteroskedasticity” *Journal of Econometrics* **31**, 307-327.

Bollerslev, T. (1990) “Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model” *Review of Economics and Statistics* **72**, 498-505.

Charlot, P., O. Darné and Z. Moussa (2016) “Commodity Returns Co-movements: Fundamentals or ‘Style’ Effect?” *Journal of International Money and Finance* **68**, 130-160.

Conrad, C. and M. Karanasos (2010) “Negative Volatility Spillovers in the Unrestricted ECCC-GARCH Model” *Econometric Theory* **26**, 838-862.

Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton University Press: Princeton.

Francq, C. and J.-M. Zakoïan (2012) “QML Estimation of a Class of Multivariate Asymmetric GARCH Models” *Econometric Theory* **28**, 179-206.

Giamouridis, D. and I. D. Vrontos (2007) “Hedge Fund Portfolio Construction: A Comparison of Static and Dynamic Approaches” *Journal of Banking & Finance* **31**, 199-217.

Giot, P. and S. Laurent (2003) “Value–at–Risk for Long and Short Trading Positions” *Journal of Applied Econometrics* **18**, 641-664.

Haas, M. (2010) “Covariance Forecasts and Long-run Correlations in a Markov-switching Model for Dynamic Correlations” *Finance Research Letters* **7**, 86-97.

He, C. and T. Teräsvirta (1999) “Properties of Moments of a Family of GARCH Processes” *Journal of Econometrics* **92**, 173-192.

He, C. and T. Teräsvirta (2004) “An Extended Constant Conditional Correlation GARCH Model and its Fourth-Moment Structure” *Econometric Theory* **20**, 904-926.

Helmert, F. R. (1876) ”Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers directer Beobachtungen gleicher Genauigkeit” *Astronomische Nachrichten. Expedition auf der Königlichen Sternwarte bei Kiel* **88**, 113-132.

Jeantheau, T. (1998) “Strong Consistency of Estimators for Multivariate ARCH Models” *Econometric Theory* **14**, 70-86.

Kamat, A. R. (1953) “Incomplete and Absolute Moments of the Multivariate Normal Distribution with Some Applications" *Biometrika* **40**, 20-34.

Lejeune, B. (2009) "A Diagnostic M-test for Distributional Specification of Parametric Conditional Heteroscedasticity Models for Financial Data" *Journal of Empirical Finance* **16**, 507-523.

Li, W. V. and A. Wei (2009) "Gaussian Integrals Involving Absolute Value Functions" in *IMS Collections, High Dimensional Probability V: The Luminy Volume* by C. Houdré, V. Koltchinskii, D. M. Mason and M. Peligrad, Eds., Springer: Berlin, 43-59.

Ling, S. and M. McAleer (2002) "Stationarity and the Existence of Moments of a Family of GARCH Processes" *Journal of Econometrics* **106**, 109-117.

Ling, S. and M. McAleer (2003) "Asymptotic Theory for a Vector ARMA-GARCH Model" *Econometric Theory* **19**, 280-310.

McNeil, A. J., R. Frey and P. Embrechts (2015) *Quantitative Risk Management* (2nd ed), Princeton University Press: Princeton.

Nabeya, S. (1951) "Absolute Moments in 2-dimensional Normal Distribution" *Annals of the Institute of Statistical Mathematics* **3**, 2-6.

Owen, D. B. (1956) "Tables for Computing Bivariate Normal Probabilities" *Annals of Mathematical Statistics* **27**, 1075-1090.

Pelletier, D. (2006) "Regime Switching for Dynamic Correlations" *Journal of Econometrics* **131**, 445-473.

Sibuya, M. (1959) "Bivariate Extreme Statistics, I" *Annals of the Institute of Statistical Mathematics* **11**, 195-210.

Taylor, S. J. (1986) *Modelling Financial Time Series*, John Wiley & Sons: Chichester.

Wellner, J. A. and R. T. Smythe (2002) "Computing the Covariance of Two Brownian Area Integrals" *Statistica Neerlandica* **56**, 101–109.

Appendix: Computation of $\Pr(X > 0, Y > 0)$

A detailed derivation of (6) along the lines of Owen (1956) is presented. For alternative derivations, see Cramér (1946, p. 290) and McNeil et al. (2015, p. 253).

Define

$$\begin{aligned}
\overline{\Phi}(x, y; \rho) &= \int_x^\infty \int_y^\infty \phi(u, v; \rho) dv du \\
&= \int_x^\infty \phi(u) \int_y^\infty \frac{1}{\sqrt{1-\rho^2}} \phi\left(\frac{v - \rho u}{\sqrt{1-\rho^2}}\right) dv du \\
&= \int_x^\infty \phi(u) \left(1 - \Phi\left(\frac{y - \rho u}{\sqrt{1-\rho^2}}\right)\right) du \\
&= \int_x^\infty \phi(u) \Phi\left(\frac{\rho u - y}{\sqrt{1-\rho^2}}\right) du, \qquad (11)
\end{aligned}$$

where $\Phi(z) = 1 - \Phi(-z)$ was used in the last line. The derivative with respect to ρ is (cf. Sibuya, 1959)

$$\begin{aligned}
\frac{d\overline{\Phi}(x, y; \rho)}{d\rho} &= \int_x^\infty \phi(u) \phi\left(\frac{y - \rho u}{\sqrt{1-\rho^2}}\right) \frac{u - \rho y}{(1-\rho^2)^{3/2}} du \\
&= \frac{1}{\sqrt{1-\rho^2}} \int_x^\infty \phi(y) \phi\left(\frac{u - \rho y}{\sqrt{1-\rho^2}}\right) \frac{u - \rho y}{\sqrt{1-\rho^2}} \frac{du}{\sqrt{1-\rho^2}} \\
&= \frac{\phi(y)}{\sqrt{1-\rho^2}} \int_{(x-\rho y)/\sqrt{1-\rho^2}}^\infty z \phi(z) dz \qquad (12) \\
&= \frac{1}{\sqrt{1-\rho^2}} \phi(y) \phi\left(\frac{x - \rho y}{\sqrt{1-\rho^2}}\right) \qquad (13) \\
&= \phi(x, y; \rho),
\end{aligned}$$

where $\phi'(z) = -z\phi(z)$ was used to go from (12) to (13). Therefore (cf. Owen, 1956),

$$\begin{aligned}
\Pr(X > 0, Y > 0) &= \overline{\Phi}(0, 0; \rho) \\
&= \int_0^\rho \phi(0, 0; \xi) d\xi + \overline{\Phi}(0, 0; 0) \\
&= \frac{1}{2\pi} \int_0^\rho \frac{d\xi}{\sqrt{1-\xi^2}} + \frac{1}{4} \\
&= \frac{\arcsin \rho}{2\pi} + \frac{1}{4}.
\end{aligned}$$