On the tail integral formulae for real-valued random variables

SARALEES NADARAJAH and IDIKA E. OKORIE

1. *Introduction*

The most important properties of any distribution are its moments. The first four moments, for example, can be used describe any data set fairly well. Moments can also be used for estimation.

So the calculation of moments needs efficient methods. The most traditional method has been to integrate the probability density function or the joint probability density function. But often distributions are more easily specified by (joint) cumulative distribution functions or (joint) survival functions. Since the early 2000s, several methods have been proposed to calculate moments based on (joint) cumulative distribution functions or (joint) survival functions. We mention Nadarajah and Mitov [1], Hong [2, 3] and Chakraborti *et al*. [4]. But all of these methods suppose that the random variables or the random vectors of interest are non-negative.

But certain results have been known for arbitrary random variables too. Hoeffding [5] expressed the covariance between two arbitrary random variables in terms of their joint survival function as

$$\mathrm{Cov}(X,Y) = \int_R \int_R \left[\Pr(X > x, Y > y) - \Pr(X > x)(\Pr(Y > y)\right] dx\, dy. \quad (1)$$

Mardia [6] extended (1) to $\mathrm{Cov}\left(X^r, Y^s\right)$, $r \geqslant 1$, $s \geqslant 1$ in terms of joint cumulative distribution functions. Block and Fang [7] extended (1) for multivariate random vectors involving joint cumulants.

If X is a non-negative continuous random variable then it is well known that

$$E(X) = \int_0^\infty \Pr(X > x)\, dx \quad (2)$$

and

$$E\left(X^n\right) = n \int_0^\infty x^{n-1} \Pr(X > x)\, dx. \quad (3)$$

Equation (2) follows by Fubini's theorem and

$$\int_0^\infty \Pr(X > x) dx = \int_0^\infty \int_0^\infty f(y) dy\, dx = \int_0^\infty \int_0^y f(y) dx\, dy = \int_0^\infty y f(y)\, dy,$$

where $f(\,)$ denotes the probability density function of X assumed to exist. Equation (3) follows by Fubini's theorem and

$$n \int_0^\infty x^{n-1} \Pr(X > x)\, dx = n \int_0^\infty x^{n-1} \int_x^\infty f(y)\, dy\, dx$$

$$= n \int_0^\infty \int_0^y x^{n-1} f(y)\, dx\, dy = \int_0^\infty y^n f(y)\, dy.$$

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

Lo [8] proved that (2) can be extended to the following, applicable for any integrable random variable X:

$$E(X) = \int_0^\infty \Pr(X > x)\,dx - \int_0^\infty \Pr(X > -x)\,dx. \tag{4}$$

But (4) has been known at least since Block and Fang [7].

In (4), the two integrals on the right-hand side can be regarded as the first moments of the right and left tails, respectively. Often, higher order moments of the tail are of interest. Examples of applications of high order tail moments include risk management (Berkowitz [9]), actuarial science (Kim [10]), asset pricing theory (Zambrano [11]) and estimation under serial dependence (Hoga [12]).

The aim of this Article is to extend (4) for moments of general order. Theorem 1 in Section 2 derives an expression for $E(X_n)$ in terms of tail moments when X is any integrable random variable. Theorem 2 in Section 2 derives an expression for $E(X^m Y^n)$ in terms of tail moments when (X, Y) is any integrable random vector. Theorem 3 in Section 2 derives an expression for $E\left(X_1^{m_1} X_2^{m_2} \dots X_p^{m_p}\right)$ in terms of tail moments when $(X_1, X_2, \dots, X_p)$ is any integrable random vector. The general multivariate case was not considered in Lo [8]. Some illustrations of the theorems are given in Section 3.

A future work is to consider more general functions of random variables or vectors than just powers. Some work in that direction is discussed in Section 4.

2. *Main results*

Theorem 1 extends Proposition 2.1 in Lo [8] for moments of general order. Theorem 2 extends Proposition 3.1 in Lo [8] for product moments of general order.

Theorem 1

Let X be an integrable random variable up to order n on the real line. Then

$$E(X_n) = n\int_0^\infty x^{n-1}\Pr(X > x)dx + n(-1)^n \int_0^\infty x^{n-1}\Pr(X < -x)dx \tag{5}$$

for $n \geqslant 1$.

Proof:

Write $X = X_+ - X_-$, where $X_+ = \max(X,0)$ and $X_- = \max(-X,0)$. Then

$$E(X^n) = E\left[(X_+ - X_-)^n\right] = \sum_{m=0}^{n} \binom{n}{m} (-1)^m E\left(X_+^{n-m} X_-^m\right).$$

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

Since $X_+X_- = 0$,

$$\begin{aligned} E(X^n) &= E(X_+^n) + (-1)^n E(X_-^n) \\ &= n\int_0^\infty x^{n-1} \Pr(X_+ > x)\,dx + n(-1)^n \int_0^\infty x^{n-1} \Pr(X_- > x)\,dx \\ &= n\int_0^\infty x^{n-1} \Pr(X > x)\,dx + n(-1)^n \int_0^\infty x^{n-1} \Pr(X < -x)\,dx. \end{aligned}$$

The proof is complete.

Equation (4) is the particular case of (5) for $n = 1$.

Theorem 2

Let (X, Y) be an integrable random vector up to order (m, n). Then

$$\begin{aligned} E(X^mY^n) &= mn\int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X > x, Y > y)\,dx\,dy \\ &\quad + mn(-1)^{m+n} \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X < -x, Y < -y)\,dx\,dy \\ &\quad + mn(-1)^n \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X > x, Y < -y)\,dx\,dy \\ &\quad + mn(-1)^m \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X < -x, Y > y)\,dx\,dy \end{aligned} \tag{6}$$

for $m \geqslant 1$ and $n \geqslant 1$.

Proof:

Write $X = X_+ - X_-$ and $Y = Y_+ - Y_-$, where $X_+ = \max(X, 0)$, $X_- = \max(-X, 0)$, $Y_+ = \max(Y, 0)$ and $Y_- = \max(-Y, 0)$. Then

$$\begin{aligned} E(X^mY^n) &= E\left[(X_+ - X_-)^m (Y_+ - Y_-)^n\right] \\ &= \sum_{p=0}^{m} \sum_{q=0}^{n} \binom{m}{p}\binom{n}{q} (-1)^{p+q} E\left(X_+^{m-p} X_-^p Y_+^{n-q} Y_-^q\right). \end{aligned}$$

Since $X_+X_- = 0$ and $Y_+Y_- = 0$,

$$\begin{aligned} E(X^mY^n) &= E(X_+^mY_+^n) + (-1)^{m+n}E(X_-^mY_-^n) + (-1)^n E(X_+^mY_-^n) + (-1)^m E(X_-^mY_+^n) \\ &= mn\int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X_+ > x, Y_+ > y)\,dx\,dy \\ &\quad + mn(-1)^{m+n} \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X_- > x, Y_- > y)\,dx\,dy \\ &\quad + mn(-1)^n \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X_+ > x, Y_- > y)\,dx\,dy \\ &\quad + mn(-1)^m \int_0^\infty \int_0^\infty x^{m-1}y^{n-1} \Pr(X_- > x, Y_+ > y)\,dx\,dy \end{aligned}$$

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

$$= mn\int_0^\infty\int_0^\infty x^{m-1}y^{n-1}\Pr(X > x, Y > y)dx\,dy$$

$$+ mn(-1)^{m+n}\int_0^\infty\int_0^\infty x^{m-1}y^{n-1}\Pr(X < -x, Y < -y)dx\,dy$$

$$+ mn(-1)^{n}\int_0^\infty\int_0^\infty x^{m-1}y^{n-1}\Pr(X > x, Y < -y)dx\,dy$$

$$+ mn(-1)^{m}\int_0^\infty\int_0^\infty x^{m-1}y^{n-1}\Pr(X < -x, Y > y)dx\,dy.$$

The proof is complete.

Theorem 3 extends Theorems 1 and 2 for multivariate random vectors. Its proof is similar to that of Theorem 2. Lo [8] considered only the bivariate case.

Theorem 3

Let $(X_1, X_2, \ldots, X_p)$ be an integrable random vector up to order $(m_1, m_2, \ldots, m_p)$. Let $A_{j,0} = \{X_j > x_j\}$ and $A_{j,1} = \{X_j < -x_j\}$ for $j = 1, 2, \ldots, p$. Then

$$E\left(X_1^{m_1}X_2^{m_2}\ldots X_p^{m_p}\right) = \left(\prod_{i=1}^{p} m_i\right)\int_0^\infty\int_0^\infty\ldots\int_0^\infty\left(\prod_{i=1}^{p} x_i^{m_i-1}\right)$$

$$\cdot\sum_{i_1=0}^{1}\sum_{i_2=0}^{1}\ldots\sum_{i_p=0}^{1}(-1)^{i_1+i_2+\ldots+i_p}\Pr\left(\bigcap_{j=1}^{p}A_{j,i_j}\right)\left(\prod_{i=1}^{p}dx_i\right)$$

for $m_1 \geqslant 1, m_2 \geqslant 1, \ldots, m_p \geqslant 1$.

3. *Examples*

In this section, we give two examples which illustrate the use of Theorems 1 and 2. In both examples, the distribution is more easily specified by the cumulative distribution function.

If X is a random variable such that $\Pr(X < -x) = \Pr(X > x)$ for all x then it is said to have a *symmetric* distribution. By Theorem 1, the nth moment of any symmetric distribution can be written as

$$E\left(X^n\right) = n\left[1 + (-1)^n\right]\int_0^\infty x^{n-1}\Pr(X > x)\,dx, \tag{7}$$

which is zero if n is odd. The even order moments are the same as those of $|X|$.

For example, consider the symmetric Weibull distribution specified by

$$\Pr(X < -x) = \Pr(X > x) = \tfrac{1}{2}\exp\left(-|x|^\alpha\right)$$

for $\alpha > 0$ and all x. By Wikipedia [13],

$$E\left(|X|^n\right) = \frac{n}{\alpha}\Gamma\left(\frac{n}{\alpha}\right), \tag{8}$$

where $\Gamma(\,)$ denotes the gamma function.

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

If (X, Y) is a random vector such that $\Pr(X > x, Y > y) = \Pr(X < -x, Y < -y) = \Pr(X > x, Y < -y) = \Pr(X < -x, Y > y)$ for all x and y then it is said to have a *symmetric* bivariate distribution. By Theorem 2, the product moment of any symmetric bivariate distribution can be written as

$$E(X^m Y^n) = mn\left[1+(-1)^m+(-1)^n+(-1)^{m+n}\right]\int_0^\infty\int_0^\infty x^{m-1}y^{n-1}\Pr(X>x,Y>y)dx\,dy, \quad (9)$$

which is zero if m is odd or n is odd. The other product moments are the same as those of $(|X|, |Y|)$.

For example, consider the symmetric bivariate Pareto distribution specified by

$$\begin{aligned}\Pr(X > x, Y > y) &= \Pr(X < -x, Y < -y)\\ &= \Pr(X > x, Y < -y)\\ &= \Pr(X < -x, Y > y)\\ &= \frac{1}{4}(1 + x + y)^{-a}\end{aligned}$$

for $a > 0$ and all x and y. By Section 3.1 in Sarabia *et al*. [14], the product moment of $(|X|, |Y|)$ can be written as

$$E(|X|^m |Y|^n) = mnB(a - n, n)B(a - m - n, m) \quad (10)$$

provided that $a > m + n$.

4. *Discussion*

There have been several proofs for (1). One simple proof is as follows. Let (X_1, Y_1) and (X_2, Y_2) be independent and identical copies of (X, Y). We can write

$$X_1 - X_2 - = \int (I_{\{X_1 \geqslant u\}} - I_{\{X_2 \geqslant u\}})\,du,$$

where I_A denotes the indicator function. Then the left-hand side of (1) can be written as

$$\begin{aligned}\mathrm{Cov}(X, Y) &= E(XY) - E(X)E(Y)\\ &= \tfrac{1}{2}E[(X_1 - X_2)(Y_1 - Y_2)]\\ &= \tfrac{1}{2}E\int\int (I_{\{X_1 \geqslant u\}} - I_{\{X_2 \geqslant u\}})(I_{\{Y_1 \geqslant v\}} - I_{\{Y_2 \geqslant v\}})\,du\,dv\\ &= \int\int [1-\Pr(X\leqslant u)-\Pr(Y\leqslant v)+\Pr(X\leqslant u, Y\leqslant v)-\Pr(X>u)\Pr(Y>v)]\,du\,dv,\end{aligned}$$

where the last step follows by Fubini's theorem. Hence we have the right-hand side of (1).

As already mentioned in Section 1, (1) has been extended by several authors. Theorem 1 of Cuadras [15] extended (1) to $\mathrm{Cov}(f(X), g(Y))$ when f, g are of bounded variation, $E(|f(X)g(Y)|) < \infty$, $E(|f(X)|) < \infty$ and

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

$E(|g(Y)|) < \infty$. Cuadras showed that

$$\begin{aligned}&\mathrm{Cov}\left(f(X),\, g(Y)\right)\\ &= \int_R \int_R \left[\Pr(X > x, Y > y) - \Pr(X > x)\Pr(Y > y)\right] df(x)\, dg(y). \qquad (11)\end{aligned}$$

Equation (11) may not hold if f and g are discontinuous or not sufficiently continuous functions. Theorem 2.4 in Lo [15] considered the case that f and g are non-decreasing real functions. In this case, the theorem showed that the right-hand side of (11) can be expressed as a sum of four double integrals taking the form of (11) with respect to left-continuous and right-continuous parts of f, g. Theorem 3.1 and Corollary 2.7 in Lo [15] showed that a double integral of the form of (11) holds if: (i) both f and g are right-continuous; (ii) f is right-continuous and g is left-continuous; (iii) f is left-continuous and g is right-continuous; (iv) both f and g are left-continuous; (v) both f and g are continuous; (vi) joint cumulative distribution function of (X, Y) is continuous; (vii) both f and g are absolutely continuous functions. Remark 3.2 in Lo [16] considered the case that f is non-decreasing and g is absolutely continuous. In this case, Lo[16] showed that the right-hand side of (11) can be expressed as a sum of two double integrals taking the form of (11) with respect to g and left-continuous and right-continuous parts of f.

All of these papers involve the covariance between functions of two random variables. A future work is to see if this can be extended for the multivariate case, i.e. dependence between functions of three or more random variables.

Acknowledgement

The authors would like to thank the Editor and the referee for careful reading and comments which greatly improved the paper.

References

1. S. Nadarajah and K. Mitov, Product moments of multivariate random vectors, *Communications in statistics – theory and methods* **32** (2003) pp. 47-60.
2. L. Hong, A remark on the alternative expectation formula, *The American Statistician* **66** (2012) pp. 232-233.
3. L. Hong, Another remark on the alternative expectation formula, *The American Statistician* **69** (2015) pp. 157-159.
4. S. Chakraborti, F. Jardim and E. Epprecht, (2018), Higher order moments using the survival function: the alternative expectation, *The American Statistician* **73** (2019) pp. 191-194.
5. W. Hoeffding, Masstabinvariante Korrelationtheorie, *Schriften Math. Inst. Univ. Berlin* **5** (1940) pp. 181-233.

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press

6. K. V. Mardia, Some contributions to contingency-type bivariate distributions, *Biometrika* **54** (1967) pp. 235-249.
7. H. W. Block and Z. Fang, A multivariate extension of Hoeffding's lemma, *Annals of Probability* **16** (1988) pp. 1803-1820.
8. A. Lo, Demystifying the integrated tail probability expectation formula, *The American Statistician* **73** (2019) pp. 367-374.
9. J. Berkowitz, Testing density forecasts, with applications to risk management, *Journal of Business and Economic Statistics* **19** (2001) pp. 465-474.
10. J. H. T. Kim, Conditional tail moments of the exponential family and its related distributions, *North American Actuarial Journal* **14** (2010) pp. 198-216.
11. J. A. Zambrano, Higher-order tail moments in asset-pricing theory in *Handbook of global financial markets,* World Scientific (2017).
12. Y. Hoga, Extreme conditional tail moment estimation under serial dependence, *Journal of Financial Econometrics* **17** (2019) pp. 587-615.
13. https://en.m.wikipedia.org/wiki/Weibull_distribution#Shannon_entropy
14. J. M. Sarabia, E. Gomez-Deniz, F. Prieto and V. Jorda, Risk aggregation in multivariate dependent Pareto distributions, *Insurance: Mathematics and Economics* **71** (2016) pp. 154-163.
15. C. M. Cuadras, On the covariance between functions, *Journal of Multivariate Analysis* **81** (2002) pp. 19-27.
16. A. Lo, Functional generalizations of Hoeffding's covariance lemma and a formula for Kendall's tau, *Statistics and Probability Letters* **122** (2017) pp. 218-226.

10.1017/mag.2022.122 © The Authors, 2022
Published by Cambridge University Press on behalf of The Mathematical Association

SARALEES NADARAJAH
Department of Mathematics, University of Manchester, Manchester M13 9PL
e-mail: *mbbsssn2@manchester.ac.uk*
IDIKA E. OKORIE
Department of Mathematics, Khalifa University, P.O. Box 127788, Abu Dhabi, UAE
email: *idika.okorie@ku.ac.ae*

https://doi.org/10.1017/mag.2022.122 Published online by Cambridge University Press