

Discussion of "The Log-Gamma Distribution and Non-Normal Error"

Elias S. W. Shiu^{1a}

¹ University of Iowa

Keywords: Log-gamma distribution, cumulant, T.N. Thiele, linear exponential family

Variance

Vol. 16, Issue 1, 2023

A key goal of the discussion is to show a relatively simple way to derive the cumulant formulas in Appendix B of the paper "The Log-Gamma Distribution and Non-Normal Error." It also suggests alternative derivation for various results in the paper. The concept of cumulants was introduced by T. N. Thiele. The discussion also points out some important actuarial roles played by Mr. Thiele.

Address for Correspondence: elias-shiu@uiowa.edu

There are many interesting historical facts in this paper. I would like to add the following. The concept of cumulants, which plays a prominent role in the paper, was introduced in 1899 by Thorvald Nicolai Thiele (1838-1910) under the name half-invariants, several decades before the rediscovery by the eminent statistician R. A. Fisher. Thiele was a Danish actuary, astronomer, mathematician, and statistician. He was a cofounder of the Danish insurance company Hafnia and was its Mathematical Director (actuary). Also, he was the founding president of the Danish Actuarial Society, a corresponding member of the British Institute of Actuaries, and a member of the Board of the Permanent Committee of the International Congresses of Actuaries. The net premium reserve differential equation,

$$\frac{d}{dt} {}_tV = P + {}_tV\delta - (b - {}_tV)\mu_{x+t},$$

in the theory of life contingencies was due to him, although it was not published until after he died. More information about Thiele can be found in the *Encyclopedia of Actuarial Science* article by Norberg (2004) or by looking up the website *MacTutor History of Mathematics Archive* (Robertson and O'Connor, n.d.).

APPENDIX B

Let me present a relatively simple way to derive the formulas in Appendix B. The cumulant-generating function of a random variable X is

$$\Psi_X(t) := \ln [M_X(t)].$$

Because $\kappa_n(X)$, the n th cumulant of X , is the n th derivative of $\Psi_X(t)$ at 0, $\kappa_n(X)$ is the coefficient of $\frac{t^n}{n!}$ in the Maclaurin

series of $\Psi_X(t)$. It turns out that it is not necessary to do "tedious" differentiations to determine the Maclaurin series because we have the logarithm series

$$\ln(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \dots$$

Observe that, with $\mu := E[X]$,

$$M_X(t) = E[e^{tX}] = e^{t\mu} E[e^{t(X-\mu)}],$$

and that, with $\mu_j := E[(X-\mu)^j]$, $j = 2, 3, 4, \dots$,

$$\begin{aligned} E[e^{t(X-\mu)}] &= E\left[\sum_{j \geq 0} \frac{[t(X-\mu)]^j}{j!}\right] \\ &= \sum_{j \geq 0} E\left[\frac{[t(X-\mu)]^j}{j!}\right] \\ &= 1 + 0 + \sum_{j \geq 2} \frac{\mu_j t^j}{j!}. \end{aligned}$$

Hence, by using the logarithm series and some school algebra,

$$\begin{aligned} \Psi_X(t) &= \mu t + \ln\left(1 + \sum_{j \geq 2} \frac{\mu_j t^j}{j!}\right) \\ &= \mu t + \sum_{j \geq 2} \frac{\mu_j t^j}{j!} - \frac{1}{2} \left(\sum_{j \geq 2} \frac{\mu_j t^j}{j!}\right)^2 \pm \dots \\ &= \mu t + \frac{\mu_2 t^2}{2!} + \frac{\mu_3 t^3}{3!} + \frac{(\mu_4 - 3\mu_2^2) t^4}{4!} + \frac{(\mu_5 - 10\mu_2\mu_3) t^5}{5!} + \dots \end{aligned}$$

SECTION 2

Let me derive the cumulant formulas on the right column of page 175 in a more general context. From the left column on page 175, one can see that the probability density func-

^a Dr. Shiu is Professor of Actuarial Science at the University of Iowa. From 1976 to 1991, he was a professor of actuarial science at the University of Manitoba and a consultant for the Great-West Life Assurance Company, Winnipeg, Canada. He joined the University of Iowa in 1992 as Principal Financial Group Professor. He is an Honorary Editor for the journal *Insurance: Mathematics and Economics* and was a Co-Editor for the *North American Actuarial Journal*. His paper "Option Pricing by Esscher Transforms" is the most cited paper ever published in the *Transactions of the Society of Actuaries*.

tion (pdf) of the log-gamma random variable Y has the form

$$f_Y(u) = e^{\alpha u + g(u) + h(\alpha)},$$

which means that Y is a member of the *linear exponential family*, a concept that can be found in the actuarial textbook Klugman, Panjer, and Willmot (2019). Because a pdf integrates to 1, we have

$$\int_{-\infty}^{\infty} e^{\alpha u + g(u)} du = e^{-h(\alpha)}.$$

Thus, the moment-generating function of Y is

$$\begin{aligned} M_Y(t) &= \int_{-\infty}^{\infty} e^{(t+\alpha)u + g(u) + h(\alpha)} du \\ &= e^{-h(t+\alpha) + h(\alpha)}, \end{aligned}$$

which implies that the cumulant-generating function is

$$\begin{aligned} \Psi_Y(t) &= h(\alpha) - h(t + \alpha) \\ &= h(\alpha) - \sum_{j \geq 0} \frac{h^{(j)}(\alpha) t^j}{j!} \\ &= - \sum_{j \geq 1} \frac{h^{(j)}(\alpha) t^j}{j!}. \end{aligned}$$

Hence for the linear exponential family, we have for $j = 1, 2, 3, \dots$,

$$\kappa_j = -h^{(j)}(\alpha);$$

also,

$$\kappa_{j+1} = \frac{\partial}{\partial \alpha} \kappa_j.$$

The actuarial textbook Kaas et al. (2008, 300) calls $-h(\alpha)$ the *cumulant function*. For the random variable Y in Section 2,

$$-h(\alpha) = \ln \Gamma(\alpha) + \alpha \ln \theta.$$

SECTION 4

There seems no need for the moment-generating function calculation in the right column of page 177. Here,

$$Y = \gamma \ln X_1 - \gamma \ln X_2 + C.$$

With X_1 and X_2 being independent random variables,

$$\begin{aligned} \Psi_Y(t) &= \Psi_{\gamma \ln X_1}(t) + \Psi_{-\gamma \ln X_2}(t) + Ct \\ &= \Psi_{\ln X_1}(\gamma t) + \Psi_{\ln X_2}(-\gamma t) + Ct. \end{aligned}$$

Because the cumulants of a random variable can be obtained from the coefficients of the Maclaurin series of the cumulant-generating function, it follows from the last equation that

$$\kappa_1(Y) = \gamma \kappa_1(\ln X_1) - \gamma \kappa_1(\ln X_2) + C,$$

$$\kappa_j(Y) = \gamma^j \kappa_j(\ln X_1) + (-\gamma)^j \kappa_j(\ln X_2), \quad j = 2, 3, 4, \dots$$

In Section 4, X_1 is assumed to be a *Gamma*($\alpha, 1$) random variable, and X_2 a *Gamma*($\beta, 1$) random variable. Hence,

$$\kappa_j(\ln X_1) = \frac{d^j}{d\alpha^j} \ln \Gamma(\alpha)$$

and

$$\kappa_j(\ln X_2) = \frac{d^j}{d\beta^j} \ln \Gamma(\beta).$$

SECTION 7

It is stated in the Conclusion paragraph of the paper that “loss distributions ... are non-negative, positively skewed, and right-tailed.” Then it is claimed that “most loss distributions are transformations of the gamma distribution” which I find confusing. It may be useful for me to state the following result, whose proof can be found in a short discussion by Ko and Ng (2007). The distribution of any positive random variable can be arbitrarily closely approximated by a weighted average of Erlang distributions, where the weights are positive. An Erlang distribution is a gamma distribution with the shape parameter α being a positive integer.

Also, Ko and Ng (2007) show that the distribution of any positive random variable can be arbitrarily closely approximated by a weighted average of exponential distributions, but some of the weights may be negative. There are other approximation results given in this two-page discussion by Ko and Ng (2007).

Submitted: July 01, 2021 EDT. Accepted: October 04, 2021 EDT. Published: February 13, 2023 EDT.

REFERENCES

Kaas, Rob, Marc Goovaerts, Jan Dhaene, and Michel Denuit. 2008. *Modern Actuarial Risk Theory: Using R*. 2nd ed. New York: Springer. <https://doi.org/10.1007/978-3-540-70998-5>.

Klugman, S.A., H.H. Panjer, and G.E. Willmot. 2019. *Loss Models: From Data to Decisions*. 5th ed. New York: Wiley.

Ko, B., and A.C.Y. Ng. 2007. “Discussion on ‘Stochastic Life Annuities’ by D. Dufresne.” *North American Actuarial Journal* 11 (3): 170–71.

Norberg, R. 2004. “Thiele, T. N. (1838-1910).” In *Encyclopedia of Actuarial Science*, 1671–72. New York: Wiley.

Robertson, E., and J. O’Connor. n.d. “MacTutor History of Mathematics Archive.” University of St. Andrews. <https://mathshistory.st-andrews.ac.uk/>.