Journal of Statistical Theory and Applications
Vol. **19(3)**, September 2020, *pp.* 383–390
DOI: https://doi.org/10.2991/jsta.d.200826.001; ISSN 1538-7887
https://www.atlantis-press.com/journals/jsta

Deriving Mixture Distributions Through Moment-Generating Functions

Subhash Bagui[1,*], Jia Liu[1], Shen Zhang[2]

[1] *Department of Mathematics and Statistics, The University of West Florida, Pensacola, FL, USA*
[2] *Department of Statistics, The University of Texas at San Antonio, San Antonio, TX, USA*

ARTICLE INFO

Article History
Received 10 Mar 2020
Accepted 14 Aug 2020

Keywords
Mixture distributions
Moment-generating functions
Characteristic functions
Hierarchical models
Over-dispersed models

Mathematics Subject Classification: 62E15; 62E10

ABSTRACT

This article aims to make use of moment-generating functions (mgfs) to derive the density of mixture distributions from hierarchical models. When the mgf of a mixture distribution doesn't exist, one can extend the approach to characteristic functions to derive the mixture density. This article uses a result given by E.R. Villa, L.A. Escobar, Am. Stat. 60 (2006), 75–80. The present work complements E.R. Villa, L.A. Escobar, Am. Stat. 60 (2006), 75–80 article with many new examples.

1. INTRODUCTION

A random variable X is said to have mixture distribution if it depends on a quantity that has a distribution. The mixture distribution arises from a hierarchical model (see [1], p. 165). A typical example of a hierarchical model is as follows: Consider a large number of eggs laid by an insect. The survival of each egg has a probability θ. If each egg's survival is independent, then we have a sequence of Bernoulli trials on egg's survival. Assume that the "large number" of eggs laid is a random variable and follows the Poisson distribution with parameter λ. Hence, if we let X is equal to the *number of survivors* and Y is equal to the *number of eggs laid*, then

$$X|Y \text{ follows Binomial } (Y, \theta) \text{ and } Y \text{ follows Poisson } (\lambda), \tag{1}$$

constitute a hierarchical model, where the notation $X|Y = y$ denotes the conditional distribution of X given $Y = y$ follows Binomial (y, θ), and the (marginal) distribution of Y is Poisson (λ). It turns out that the (marginal) distribution of X is Poisson $(\lambda\theta)$. Thus, the distribution of the number of survivors X, which is Poisson $(\lambda\theta)$, is a mixture distribution as it is a result of combining the distribution of $X|Y$ and the distribution of Y. In general, hierarchical models lead to mixture distributions.

Mixture models play an important role in the theory and practice. There are textbooks, monographs, and journal articles discussing history, theory, applications, and the usefulness of mixture models. Mixture models became popular as, among others: they (a) provide a simple device to include other variation and correlation in the model, (b) add model flexibility, and (c) allow modeling the data that arise in multi-stages. The literature shows several authors, namely Everitt and Hand [2], Titterington *et al.* [3], Böhning [4], McLachlan and Peel [5] have discussed mixture models and provided the statistical methodology and references on finite mixtures.

Lindsay [6] discussed the application of mixture models and its interrelation with other related fields, among others. In discussing mixture models, Casella and Berger [1] showed the derivation of mixture models from hierarchical models. Panjer and Willmot [7] consider applications of mixture models in actuarial sciences. Karlis and Xekalaki [8] derived results related to Poisson mixtures models with applications in various other fields. The mixture models of continuous and discrete types can be found in Johnson *et al.* [9,10] and Gelman *et al.* [11]. To fit plant quadrat data on the blue-green sedge, Skellam [12] used a mixture of binomial with varying sample sizes modeled with Poisson distributions. The Gamma mixture of Poisson r.v.'s yield negative-binomial, while Green and Yule [13] used this mixture distribution to model "*accident proneness*"; see Bagui and Mehra [14]. The research dated back to Pearson [15] shows modeling the mixing of different

*Corresponding author. Email: sbagui@uwf.edu

crab type with mixtures of two normal. The mixture distribution negative-binomial can arise in the distribution of the sum of N independent random variables, each having the same logarithmic distribution and N having a Poisson distribution; this mixture distribution was used in modeling biological spatial data; see Gurland [16], Bagui and Mehra [14].

2. MIXTURE MODEL

Consider a two-stage mixture model of type (1) where $X|Y \sim f_{X|Y}(x|y)$ and $Y \sim f_Y(y)$. It is customary to derive the mixture distribution $f_X(x)$ of the mixture X from the joint density of X and Y, $f_{X,Y}(x|y) = f_Y(y)f_{X|Y}(x|y)$, as

$$f_X(x) = \begin{cases} \int_{-\infty}^{\infty} f_Y(y)f_{X|Y}(x|y)dy \text{ if } Y \text{ is continuous} \\ \sum_y f_Y(y)f_{X|Y}(x|y) \text{ if } Y \text{ is dicrete} \end{cases} \tag{2}$$

Villa and Escobar [17] derived mixture distributions from the moment-generating function (mgf) of X, and they derived the mgf of X as $M_X(t) = E\left[M_{X|Y}(t)\right]$, where $M_{X|Y}(t)$ is the mgf of $X|Y$ and the expectation is over the distribution of the r.v. Y. From the known distribution of $X|Y$ with known mgf Villa and Escobar [17] rewrites $M_{X|Y}(t)$ as

$$M_{X|Y}(t) = a_1(t)e^{a_2(t)Y} \tag{3}$$

where $a_1(t)$ and $a_2(t)$ are functions of t and may also depend on the parameter of the distribution of $X|Y$. Then they arrived at the mgf of X as

$$M_X(t) = E\left[M_{X|Y}(t)\right] = a_1(t)M_Y\left[a_2(t)\right]. \tag{4}$$

In the above, we assumed that all mfg's exist. When mfg of $X|Y$ is not found from the known list, then the mgf of X can be computed directly as

$$M_X(t) = E\left(e^{tX}\right) = E\left[E\left(e^{tX}|Y\right)\right] \tag{5}$$

From the joint mgf of X and Y, $M_{X,Y}(t,s)$, one can also derive the mgf of X as

$$M_X(t) = E\left[M_{X|Y}(t)\right] = E\left[E\left(e^{tX+sY}|Y\right)\right] = E\left[e^{sY}E\left(e^{tX}|Y\right)\right] = E\left[e^{sY}M_{X|Y}(t)\right]. \tag{6}$$

Then by setting $s = 0$ in the above Eq. (6), one would get the mgf of X, $M_X(t)$.

The main goal of this article is to derive mixture distributions that complements the examples of Villa and Escobar [17] using the above mgf method.

3. EXAMPLES

There are situations where obtaining the mixture distributions by using mgf is much easier than getting it by the marginalization of the joint distribution. The examples considered here are the complements of the cases discussed by Villa and Escober [17].

3.1. The Binomial–Binomial Mixture

The mixture model of the Binomial mixture of Binomial random variables is

$$\left.\begin{array}{l} X|Y \sim \text{BIN}\left(Y, p_1\right) \\ Y \sim \text{BIN}\left(n, p_2\right) \end{array}\right\} \Rightarrow X \sim \text{BIN}\left(n, p_1p_2\right). \tag{7}$$

Proof. The mgf for $X|Y$ is

$$M_{X|Y}(t) = \left[q_1 + p_1e^t\right]^Y = e^{\ln(q_1+p_1e^t)Y} = a_1(t)e^{a_2(t)Y},$$

where $q_1 = 1 - p_1$, $a_1(t) = 1$, and $a_2(t) = \ln\left(q_1 + p_1e^t\right)$.

Now by Eq. (4),

$$\begin{aligned} M_X(t) &= a_1(t)M_Y\left[a_2(t)\right] = M_Y\left[\ln\left(q_1 + p_1e^t\right)\right] \\ &= \left[q_2 + p_2e^{\ln(q_1+p_1e^t)}\right]^n = \left[q_2 + p_2\left(q_1 + p_1e^t\right)\right]^n \\ &= \left[q + pe^t\right]^n, \text{ Binomial mgf}, \end{aligned} \tag{8}$$

where $p = p_1 p_2$ and $q = q_2 + p_2 q_1 = 1 - p_1 p_2 = 1 - p$.

Thus, it follows from (8) that $X \sim \text{BIN}\left(n, p_1 p_2\right)$, (see Appendix, Table A1, Villa and Escober [17]).

3.2. The Negative-Binomial–Binomial Mixture

The mixture model of the Negative-binomial mixture of Binomial random variables is

$$\left.\begin{array}{l} X|Y \sim \text{BIN}\left(Y, p_1\right) \\ Y \sim \text{NEGBIN}\left(\alpha, p_2\right) \end{array}\right\} \Rightarrow X \sim \text{NEGBIN}(\alpha, p), \tag{9}$$

where $p = p_2 / \left(p_2 + p_1 p_2\right)$ and $q_2 = 1 - p_2$.

Proof. The mgf for $X|Y$ is

$$M_{X|Y}(t) = \left[q_1 + p_1 e^t\right]^Y = e^{\ln\left(q_1 + p_1 e^t\right) Y} = a_1(t) e^{a_2(t) Y},$$

where $q_1 = 1 - p_1$, $a_1(t) = 1$, and $a_2(t) = \ln\left(q_1 + p_1 e^t\right)$.

Now by Eq. (4),

$$\begin{aligned} M_X(t) &= a_1(t) M_Y\left[a_2(t)\right] = M_Y\left[\ln\left(q_1 + p_1 e^t\right)\right] \\ &= \left[\frac{p_2}{1 - \left(1 - p_2\right) e^{\ln\left(q_1 + p_1 e^t\right)}}\right]^\alpha = \left[\frac{p_2}{1 - \left(1 - p_2\right)\left(q_1 + p_1 e^t\right)}\right]^\alpha, \text{ negative-binomial mgf,} \\ &= \left[p / \left\{1 - (1 - p) e^t\right\}\right]^\alpha, \end{aligned} \tag{10}$$

where $p = p_2 / (p_2 + p_1 p_2)$. Thus, it follows from (10) that $X \sim \text{NEGBIN}(\alpha, p)$, (see Appendix, Table A2, Villa and Escober [17]).

3.3. The Exponential–Exponential Mixture

The mixture model of the exponential mixture of shifted exponential random variables is

$$\left.\begin{array}{l} X|Y \sim \lambda e^{-\lambda(x - Y)},\ x \geq Y \\ Y \sim (1 + \lambda) e^{-(1 + \lambda) y}, y \geq 0 \end{array}\right\} \Rightarrow X \sim \lambda(1 + \lambda) e^{-\lambda x}\left(1 - e^{-x}\right),\ x \geq 0. \tag{11}$$

Proof. The mgf for $X|Y$ is given by

$$M_{X|Y}(t) = E_{X|Y}\left[e^{tX}\right] = \lambda \int_Y^\infty e^{tx} e^{-\lambda(x - Y)} dx = \lambda e^{\lambda Y} \int_Y^\infty e^{-(\lambda - t) x} dx = \frac{\lambda e^{tY}}{(\lambda - t)},\ t < \lambda, \tag{12}$$

where $a_1(t) = \frac{\lambda}{\lambda - t}$ and $a_2(t) = t$.

Now by Eq. (4),

$$M_X(t) = \frac{\lambda}{\lambda - t} M_Y(t) = \frac{\lambda(1 + \lambda)}{(\lambda - t)(1 + \lambda - t)}. \tag{13}$$

The above mgf is the mgf of the density $f_X(x) = \lambda(1 + \lambda) e^{-\lambda x}\left(1 - e^{-x}\right), x \geq 0$. It follows from (13) that $X \sim \lambda(1 + \lambda) e^{-\lambda x}\left(1 - e^{-x}\right),\ x \geq 0$.

3.3.1. *A specific exponential–exponential mixture*

The mixture model of a specific exponential mixture of shifted exponential random variables is

$$\left.\begin{array}{l} X|Y \sim e^{-(x - Y)},\ x \geq Y \\ Y \sim e^{-y},\ y \geq 0 \end{array}\right\} \Rightarrow X \sim \text{Gamma}(2, 1),\ x \geq 0. \tag{14}$$

Proof. The mgf for $X|Y$ is given by

$$M_{X|Y}(t) = E_{X|Y}\left[e^{tX}\right] = \frac{e^{tY}}{1-t}, \tag{15}$$

where $a_1(t) = \frac{1}{1-t}$ and $a_2(t) = t$. Now by Eq. (4),

$$M_Y(t) = \frac{1}{1-t}M_y(t) = \frac{1}{1-t} \times \frac{1}{1-t} = \frac{1}{(1-t)^2}. \tag{16}$$

Eq. (16) confirms that $X \sim \text{Gamma}(2, 1)$.

3.3.2. The exponential–normal mixture

The mixture model of the exponential mixture of normal random variables is

$$\left.\begin{array}{l} X|Y \sim \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-Y)^2}, \ -\infty < x < \infty \\ Y \sim e^{-y}, \ y \geq 0 \end{array}\right\} \Rightarrow X \sim \text{Convolution of } Y \text{ and an independent } Z \sim N(0,1), \tag{17}$$

Proof. The mgf for $X|Y$ is given by

$$M_{X|Y}(t) = E_{X|Y}\left[e^{tX}\right] = e^{Yt+t^2/2}, \tag{18}$$

where $a_1(t) = e^{t^2/2}$ and $a_2(t) = t$. Now by Eq. (4),

$$M_X(t) = e^{t^2/2}M_Y(t) = \frac{e^{t^2/2}}{(1-t)}, \tag{19}$$

which is the mgf of the convolution of Y and an independent $Z \sim N(0, 1)$. Thus, X follows the distribution of the convolution of Y and an independent $Z \sim N(0, 1)$.

3.4. The Poisson–Chi-square Mixture

The mixture model of the Poisson mixture of Chi-square random variables is

$$\left.\begin{array}{l} X|Y \sim \chi^2_{n+2Y}, \ x \geq Y \\ Y \sim \text{POI}(\lambda) \end{array}\right\} \Rightarrow X \sim \chi^2_{n;2\lambda}. \tag{20}$$

Proof. The mgf for $X|Y$ is given by

$$M_{X|Y}(t) = E_{X|Y}\left[e^{tX}\right] = (1-2t)^{-(n+2Y)/2} = (1-2t)^{-n/2}e^{-\ln(1-2t)Y}, \tag{21}$$

where $a_1(t) = (1-2t)^{-n/2}$ and $a_2(t) = -\ln(1-2t)$. Now by Eq. (4),

$$\begin{aligned} M_X(t) &= (1-2t)^{-n/2}M_Y\left[-\ln(1-2t)\right] \\ &= (1-2t)^{-n/2}e^{\lambda\left(e^{-\ln(1-2t)}-1\right)} \\ &= (1-2t)^{-n/2}e^{2t\lambda/(1-2t)}, \end{aligned} \tag{22}$$

which is an mgf of a noncentral chi-square distribution that has a noncentrality parameter 2λ. It follows from (22) that X has the noncentral chi-square distribution with the noncentrality parameter 2λ.

3.5. The Geometric–Gamma Mixture

The mixture model of the Geometric mixture of Gamma random variables is

$$\left.\begin{aligned} X|Y &\sim \text{Gamma}(Y,\beta) \\ Y &\sim \text{Geometric}(\theta) \end{aligned}\right\} \Rightarrow X \sim \text{EXP}(\beta/\theta). \tag{23}$$

Proof. The mgf for $X|Y$ is given by

$$M_{X|Y}(t) = E_{X|Y}\left[e^{tX}\right] = (1-\beta t)^{-Y} = e^{-Y\ln(1-\beta t)}, \tag{24}$$

where $a_1(t) = 1$ and $a_2(t) = -\ln(1-\beta t)$. Now by the Eq. (4),

$$\begin{aligned} M_X(t) = M_Y\left[\ln(1-\beta t)^{-1}\right] &= \frac{\theta e^{\ln(1-\beta t)^{-1}}}{1-(1-\theta)e^{\ln(1-\beta t)^{-1}}} \\ &= \frac{\theta(1-\beta t)^{-1}}{1-(1-\theta)(1-\beta t)^{-1}} = \frac{\theta}{\theta-\beta t} = \frac{1}{1-(\beta/\theta)t}, \end{aligned} \tag{25}$$

which is an mgf of an exponential distribution with shape parameter β/θ. Thus, $X \sim EXP(\beta/\theta)$.

4. EXTENSION TO MIXTURES THAT DO NOT HAVE AN mgf

When mgfs do not exist for the mixture distribution, one uses the characteristic function (cf) for the mixture distributions.

The cf of an r.v. X is defined by $\phi_X(t) = E\left(e^{itX}\right)$, where $t \in R$ and $i = \sqrt{-1}$. The conditional characteristic of $X|Y$ is denoted and defined by $\phi_{X|Y}(t) = E\left(e^{itX}|Y\right)$.

The Chi-square–Normal Mixture

$$\left.\begin{aligned} X|Y &\sim N(0,1/Y) \\ Y &\sim \chi_1^2 \end{aligned}\right\} \Rightarrow X \sim \text{Cauchy}(0,1). \tag{26}$$

Proof. The cf of $X|Y$ is

$$\phi_{X|Y}(t) = E\left(e^{itX}|Y\right) = e^{-t^2/2Y}. \tag{27}$$

The cf of X can be obtained from the Eq. (27) as

$$\phi_X(t) = E\left[\phi_{X|Y}(t)\right] = E\left[e^{-t^2/2Y}\right] = \int_0^\infty e^{-t^2/2y}\frac{1}{\sqrt{2\pi}}y^{1/2-1}e^{-y/2}dy. \tag{28}$$

Now with the transformation $|t|z = y$ and simplifying (28), we have

$$\phi_X(t) = \sqrt{\frac{2|t|}{\pi}} \times \frac{1}{2}\int_0^\infty z^{1/2-1}e^{-|t|/2(z+1/z)}dz = \sqrt{\frac{2|t|}{\pi}} \times K_{1/2}\left(|t|\right), \tag{29}$$

where, $K_u(v)$ is the modified Bessel function of third kind defined by

$$K_u(v) = \frac{1}{2}\int_0^\infty z^{u-1}e^{-\frac{v}{2}(z+1/z)}dz, \ -\infty < u < \infty. \tag{30}$$

It should be noted that asymptotic form for the modified Bessel function of the third kind is

$$K_\alpha(z) \sim \sqrt{\frac{\pi}{2z}}e^{-z}\left(1 + \frac{4\alpha^2-1}{8z} + \frac{\left(4\alpha^2-1\right)\left(4\alpha^2-9\right)}{2!\,(8z)^2} + \frac{\left(4\alpha^2-1\right)\left(4\alpha^2-9\right)\left(4\alpha^2-25\right)}{3!\,(8z)^3} + \cdots\right). \tag{31}$$

Therefore, by Eq. (31), we have

$$K_{1/2}(v) = \sqrt{\frac{\pi}{2z}}e^{-z}. \tag{32}$$

Now by Eqs. (29) and (32), we obtain the cf X as

$$\phi_X(t) = \sqrt{\frac{2|t|}{\pi}} \times \sqrt{\frac{\pi}{2|t|}} e^{-|t|} = e^{-|t|}, \tag{33}$$

which is the cf of the standard Cauchy distribution. Thus, we conclude that $X \sim \text{Cauchy}(0, 1)$.

Remarks.. The F distribution arises from the mixture of chi-square and Gamma distribution and it has no mgf. In this case, one may derive the cf of this mixture distribution. The t-distribution arises from the mixture of chi-square and normal distribution and it has no mgf, Hurst [18]. Similarly, Pareto distribution is a mixture of distributions and has no mgf. In all these cases, a cf can be used in the derivation of the mixture distributions.

5. CONCLUDING REMARKS

Mixture models play vital roles in statistics. They are used in modeling actuarial applications, biological spatial data, "accident proneness," plant data on sedge Carex flacca, and applied in many other areas of statistics. Because of the high importance of mixture distributions, students should be exposed to mixture distributions as soon as they have familiarity with conditional expectations. In the current textbooks, mixture distributions are derived from the joint distribution that originated from hierarchical mixture models as a marginal distribution.

This article finds mixture distributions using mgf method. The derivation of mixture distribution using mgfs is, in general, more straightforward and shorter than an origin of the marginal density of mixture random variable from a joint density. It is because, in the present method, one relies on mgfs that have already been derived or available. However, there are examples where the derivation of the marginal density of the mixture r.v. from a joint density is much simpler.

On the other hand, there are two difficulties in the mgf methods. First, one cannot get $a_1(t)$ and $a_2(t)$ as given in Eq. (3), for all mfs $M_{X|Y}(t)$ of the conditional distributions of $X|Y$. Second, sometimes it is hard to map the derived mgf $M_X(t)$ with a distribution. It requires good knowledge with familiarization of various distributions and corresponding mgfs. The mixtures that do not have an mfg, one can extend the mgfs method to cfs method.

As pointed out by [17], the idea of using mgf method for mixture distribution can be introduced in senior mathematical statistics courses at the level of Wackerly *et al.* [19] and Larsen and Marx [20] for students who are exposed to conditional expectations and mgfs. This article is directed to first-year graduate students in the mathematical statistics course at the level of Casella and Berger [1]. The mgf technique is underexposed in the current textbooks. From a pedagogical standpoint, mgf techniques can be used as a useful tool to derive mixture distributions. For mixtures that do not have an mgf, students with a background in complex analysis may use the cfs to extend the approach. Finally, the techniques learned here can be profitably used in the study of Bayes' procedures.

CONFLICTS OF INTEREST

Authors have no conflict of interest to declare.

AUTHORS' CONTRIBUTIONS

All authors contributed equally.

ACKNOWLEDGMENTS

The authors would like to thank the Editor and referees for their careful reading of the paper.

REFERENCES

1. G. Casella, R.L. Berger, Statistical Inference, second ed., Duxbury, Pacific Grove, CA, USA, 2002.
2. B.S. Everitt, D.J. Hand, Finite Mixture Distribution, Chapman and Hall, New York, NY, USA, 1981.
3. D.M. Titterington, A.F.M. Smith, U.E. Makov, Statistical Analysis of Finite Mixture Distributions, Wiley, New York, NY, USA, 1985.
4. D. Böhning, Computer-assisted Analysis of Mixtures, Marcel Dekker, New York, NY, USA, 1999.
5. G.J. McLachlan, D. Peel, Finite Mixture Models, Wiley, New York, NY, USA, 2000.
6. B.G. Lindsay, Mixture Models: Theory, Geometry, and Applications, NSF-CBMS Regional Conference Series in Probability and Statistics, vol. 5, Institute of Mathematical Statistics, Hayward, CA, USA, 1995.
7. H.H. Panjer, G.E. Willmot, Insurance Risk Models, Society of Actuaries, Schaumburg, IL, USA, 1992.

8. D. Karlis, E. Xekalaki, Int. Stat. Rev. 73 (2005), 35–58.
9. L.J. Johnson, S. Kotz, A.W. Kemp, Univariate Discrete Distributions, Wiley, New York, NY, USA, 1992.
10. L.J. Johnson, S. Kotz, N. Balakrishnan, Continuous Univariate Distributions, Wiley, New York, NY, USA, 1994.
11. A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data Analysis, second ed., Chapman & Hall/CRC, New York, NY, USA, 2003.
12. J.G. Skellam, Biometrika. 39 (1952), 346–362.
13. M. Greenwood, G.U. Yule, J. R. Stat. Soc. A. 83 (1920), 255–279.
14. S.C. Bagui, K.L. Mehra, Am. J. Math. Stat. 9 (2019), 44–50.
15. K. Pearson, Philos. Trans. R. Soc. Lond. A. 185 (1894), 71–110.
16. J. Gurland, Am. J. Public Health. 49 (1959), 1388–1389.
17. E.R. Villa, L.A. Escobar, Am. Stat. 60 (2006), 75–80.
18. S. Hurst, Financial Mathematics Research Report No. FMRR 006-95, The Australian National University, Canberra, Australia, 1995. http://www.maths.anu.edu.au/research.reports/srr/95/
19. D.D. Wackerly, W. Mendenhall III, R.L. Scheaffer, Mathematical Statistics with Applications, sixth ed., Duxbury, Pacific Grove, CA, USA, 2002.
20. R.J. Larsen, M.L. Marx, An Introduction to Mathematical Statistics and its Applications, sixth ed., Pearson, Boston, MA, USA, 2018.

APPENDIX

Table A.1 Examples of other mixture distributions that have moment-generating functions (mgfs) [17].

Conditional	Mixing	Marginal
$X \mid Y \sim f(x \mid Y)$	$Y \sim g(y)$	$X \sim f_X(x)$
BIN (Y, p)	POI (λ)	POI $(p\lambda)$
POI (Y)	GAM (α, β)	NEGBIN $(\alpha, 1/(1+\beta))$
NOR (Y, σ^2)	NOR (μ, τ^2)	NOR$(\mu, \sigma^2 + \tau^2)$
POI (Y)	ϕ POI (α)	Neyman-A (λ, ϕ)
POI (Y)	GIG (γ, χ, ψ)	SICHEL $\left(\gamma, \frac{2}{\psi+2}, \sqrt{\chi(\psi+2)}\right)$
LEV (Y, σ)	σ SEV $(\xi, 1)$	LOGIS$(\mu, \sigma), \mu = \xi\sigma$
$\sum_1^Y X_i, X_i \sim^{\text{iid}}$ LOGSER (p)	POI (λ)	NEGBIN $\left(-\frac{\lambda}{\ln p}, p\right)$

Table A.2 List of probability density functions or probability mass functions and corresponding moment-generating functions (mgfs) [17].

Distribution	Pdf/pmf$f(x)$	mgf
BIN (n, p)	$\binom{n}{k} p^k (1-p)^{n-k}, k = 0, 1, \cdots, n$	$\left(1 - p + pe^t\right)^n$
CHISQ(n) –$\left(\chi_n^2\right)$	$\frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, x \geq 0$	$\frac{1}{(-2t)^{n/2}}, t < \frac{1}{2}$
CHISQ –noncentral $\left(\chi_{n;\lambda}^2\right)$	$\frac{1}{2} e^{-(x+\lambda)/2} \left(\frac{x}{\lambda}\right)^{k/4-1/2} I_{k/2-1}\left(\sqrt{\lambda x}\right)$	$\exp\left(\frac{\lambda t}{1-2t}\right)(1-2t)^{-k/2}, t < 1/2$
EXP (λ)	$\lambda e^{-\lambda x}, x \geq 0$	$\frac{\lambda}{\lambda - t}, t < \lambda$
GAM (α, β)	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, x \geq 0$	$\frac{1}{(1-\beta t)^\alpha}, t < 1/\beta$
GEO (p)	$p(1-p)^{k-1}, k = 1, 2, \cdots$	$\frac{p\exp(t)}{1-(1-p)\exp(t)}, t < -\ln(1-p)$
GIG (γ, χ, ψ)	$\frac{(\psi/\chi)^{\gamma/2} x^{\gamma-1}}{2K_\gamma\left(\sqrt{\chi\psi}\right)} \exp\left[-\frac{1}{2}\left(\frac{\chi}{x} + \psi x\right)\right], x > 0$	$\frac{K_\gamma\left[\sqrt{\chi(\psi-2t)}\right]}{(1-2t/\psi)^{\gamma/2} K_\gamma\left(\sqrt{\chi\psi}\right)}, t < \psi/2$
LEV (μ, σ)	$\frac{1}{\sigma}\phi_{\text{lev}}\left(\frac{x-\mu}{\sigma}\right), x \in R$	$\exp(\mu t)\Gamma(1-\sigma t), t < 1/\sigma$
LOGIS (μ, σ)	$\frac{1}{\sigma}\phi_{\text{logis}}\left(\frac{x-\mu}{\sigma}\right), x \in R$	$\exp(\mu t)\Gamma(1-\sigma t)\Gamma(1+\sigma t), \lvert t \rvert < \frac{1}{\sigma}$
LOGSER (p)	$\frac{-(1-p)^x}{x\ln p}, x = 1, 2, \cdots$	$\frac{\ln[1-(1-p)\exp(t)]}{\ln p}, t - \ln(1-p)$
NEGBIN (α, p)	$\binom{\alpha+x-1}{x} p^\alpha (1-p)^x, x = 0, 1, \cdots$	$\left[\frac{p}{1-(1-p)\exp(t)}\right]^\alpha, t < -\ln(1-p)$
NOR (μ, σ^2)	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, x \in R$	$\exp\left(\mu t + \sigma^2 t^2/2\right)$
SEV (μ, σ)	$\frac{1}{\sigma}\phi_{\text{logis}}\left(\frac{x-\mu}{\sigma}\right), x \in R$	$\exp(\mu t)\Gamma(1+\sigma t), t > -1/\sigma$
SICHEL (γ, θ, α)	$\frac{(1-\sigma)^{\gamma/2}}{K_\psi\left(\alpha\sqrt{1-\sigma}\right)} \frac{(\theta\alpha/2)^x}{x!} K_{x+\sigma}(\alpha), x \in N$	$\left[\frac{1-\theta}{1-\theta\exp(t)}\right]^{\gamma/2} \frac{K_\gamma\left[\alpha\sqrt{1-\theta\exp(t)}\right]}{K_\gamma\left(\alpha\sqrt{1-\theta}\right)}, t < -\ln\theta$