CONTRIBUTIONS TO STATISTICAL DISTRIBUTION THEORY WITH APPLICATIONS

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science and Engineering

2018

By
Xiao Jiang
School of Mathematics

Contents

Word Count: 999,999

List of Tables

List of Figures

Abstract

The whole thesis contains 10 chapters. Chapter 1 is the introductory chapter of my thesis and the main contributions are presented in Chapter 2 through to Chapter 9. Chapter 10 is the conclusion chapter. These chapters are motivated by applications to new and existing problems in finance, health care, sports and telecommunications.

In recent years, there has been a surge in applications of generalized hyperbolic distributions in finance. Chapter 2 provides a review of generalized hyperbolic and related distributions, including related programming packages. A real data application is presented which compares some of the distributions reviewed.

Chapter 3 and Chapter 4 derive conditions for stochastic, hazard rate, likelihood ratio, reversed hazard rate, increasing convex and mean residual life orderings of Pareto distributed variables and Weibull distributed variables, respectively. A real data application of the conditions is presented in each chapter.

Motivated by Lee and Cha [The American Statistician 69 (2015) 221-230], Chapter 5 introduces seven new families of discrete bivariate distributions. We reanalyze the football data in Lee and Cha (2015) and show that some of the newly proposed distributions provide better fits than the two families proposed by Lee and Cha (2015).

Chapter 6 derives the distribution of amplitude, its moments and the distribution of phase for thirty four flexible bivariate distributions. The results in part extend those given in Coluccia [IEEE Communications Letters, 17, 2013, 2364-2367].

Motivated by Schoenecker and Luginbuhl [IEEE Signal Processing Letters, 23, 2016, 644-647], Chapter 7 studies the characteristic function of products of two independent random variables. One follows the standard normal distribution and the other follows one of forty other continuous distributions. In this chapter, we give explicit expressions for the characteristic function of products, and some of the results are verified by simulations.

Cossette, Marceau and Perreault [Insurance: Mathematics and Economics, 64, 2015, 214-224] derived formulas for aggregation and capital allocation based on risks following two bivariate exponential distributions. Chapter 8 derives formulas for aggregation and capital allocation for thirty-three commonly known families of bivariate distributions. This collection of formulas could be a useful reference for financial risk management.

Chapter 9 derives expressions for the kth moment of the dependent random sum using copulas. It also extends Mao and Zhao[IMA Journal of Management Mathematics, 25, 2014, 421-433]'s results to the case where the components of the sum are not identically distributed. The practical usefulness of the results in terms of computational time and computational accuracy is demonstrated by simulation.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning

Copyright

Acknowledgements

Firstly, and most importantly, I would like to express my sincere gratitude to my supervisor, Dr. Saralees Nadarajah, for his great support, constant encouragement and guidance through my Ph.D. programme. I would never have imagined being fortunate enough to have such a good supervisor. I have learnt so much from him and he has inspired me to become a diligent researcher. I could not have achieved all I have without him, and I will be always indebted to him.

Secondly, I would also like to thank Dr. Christiana Charalambous for her great and continuous help along the way of my study and research. The finalised thesis can not be improved so much without her brilliant suggestions on correction.

Next, I would like to thank my dearest parents for their permanent love and unquestioning support for me. I am very proud of having them as my excellent mentors and friends. Thank them for always encouraging me through the years and for believing I could succeed. I would like to especially thank my wife, Lei Ni, for her love. It wouldn't have been the same to not have her presence in my life. I am very grateful for everything she has done for me, and will be forever in love with her.

Also, I would like to thank my colleagues, Dr. Stephen Chan, Dr. Jeffrey Chu, Rui Li, for all their valuable advice on my work and for sharing my enjoyments and sorrows over the past three years.

At last, I would like to thank the staff of the statistics group at the University of Manchester for their helpful support during this programme.

Chapter 1

Introduction

1.1 Aims

This thesis presents a combination of selected works which I have contributed to the field of distribution theory during my PhD studies over the past three years. The format of the main chapters presented are either papers which are published or papers which are currently under review for refereed journals.

The aims of each chapter are as follows:

- To provide a review of the univariate Generalized Hyperbolic distribution and its relatives;
- To study the smallest Pareto order statistics of two sets of independent but non-identical random variables with different scale and shape parameters;
- To study the smallest Weibull order statistics of two sets of independent but non-identical random variables with different scale and shape parameters;
- To construct new classes of discrete bivariate distributions to be applied to new and existing problems;
- To derive the distributions of amplitude and phase for bivariate distributions;
- To derive the characteristic function of products of two independent random variables;
- To derive aggregation and capital allocation formulas for bivariate distributions;
- To derive the moments of dependent random sums using copulas.

1.2 Motivations and contributions

Chapter 2: A review of Generalized Hyperbolic distributions In recent years, the Generalized Hyperbolic(GH) distribution has not only attracted increasing interest in academic research, but also been widely applied to a number of practical areas. Thus, we feel that it is necessary to provide a review of the GH distribution and its relatives. A review of related programming packages is also provided. We expect that this review can be a source of reference for academic researchers and further encourage a number of applications.

Chapter 3: Comparisons of smallest order statistics from Pareto distributions with different scale and shape parameters The work in this chapter is motivated by Torrado (2015). Smallest order statistics have been widely discussed and investigated in both academic research and industrial practice. Nowadays, Pareto distributions have been the most popular models applied in the area of finance and economics. We feel it is needed to produce and work on the smallest Pareto order statistics. In this chapter, we mainly consider three of the most popular Pareto distributions: the Pareto distribution of type I, the Pareto distribution of type II, and the Pareto distribution of type IV. A real data application is also presented as an example of the practical use of the smallest Pareto order statistics.

Chapter 4: Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters This chapter is closely related to the previous chapter. We use the same approach but apply it to the Weibull distributed random variables. In this chapter, we consider the standard Weibull distribution and a lower truncated Weibull distribution. The results presented are different from and are more general than those in Torrado (2015). A real data application is also presented as an example of the practical use of the smallest Weibull order statistics.

Chapter 5: New classes of discrete bivariate distributions with application to football data Dependent random quantities in a wide range of areas have been modelled through the application of bivariate/multivariate distributions. However, there has been relatively little research on the development of discrete bivariate/multivariate distributions. Motivated by Lee and Cha (2015), given three independent discrete random variables, e.g. U_1, U_2 and U_3, we apply possible combinations of common mathematical operators to these three variables and generate seven new classes of discrete bivariate distributions. A case study is also presented using the same data set as that in Lee and Cha (2015). It turns out that some of the newly proposed distributions provide better fits than models contributed by Lee and Cha (2015).

Chapter 6: Distributions of amplitude and phase for bivariate distribution Discussions on the distributions of the amplitude (R) and the phase (Θ) arise in many areas of the IEEE literature. In this chapter, we derive the marginal distribution of the amplitude (R), its moments and the marginal distribution of the phase (Θ) for a wide range of bivariate distributions. These include eleven bivariate normal distributions, eight bivariate t distributions, five bivariate Laplace distributions, two bivariate hyperbolic distributions, two bivariate Gumbel distributions, one bivariate logistic distribution and five other bivariate distributions. We expect that the details given in this chapter could be a useful reference for the IEEE community and encourage researchers to apply more non-normal distributions to real problems.

Chapter 7: On characteristic functions of products of two random variables Many variables in the real world can be subject to the normal distribution, so for a set of normally distributed random variables (U), it can be composed of its mean value (μ) plus the product of its standard deviation (σ) and a set of standard normal random variables (X), i.e. $U = \mu + \sigma \cdot X$. However, in the real world, the mean (μ) and the standard deviation (σ) are often themselves random variables. Under these circumstances, each variable (U_i) itself involves a product of two random variables, e.g. $\sigma_i \cdot X_i$. Motivated by Schoenecker and Luginbuhl (2016), in this chapter, we derive explicit expressions for the characteristic function of products of two random variables. One follows the standard normal distribution and the other follows one of forty-five other distributions. We expect that the details given in this chapter could be a source of reference for academic researchers and encourage further research on the theory of related functions.

Chapter 8: Aggregation and capital allocation formulas for bivariate distributions Results related to the sum of dependent risks have been of interest in both academic research and industrial practice. In recent years, several closed-form expressions for the distribution of aggregate risks, its Tail Value-at-Risk (TVaR) and TVaR based allocations have been widely developed. Motivated by Cossette et al. (2015), in this chapter, we derive the aggregation and capital allocation formulas for a comprehensive collection of bivariate distributions.

Chapter 9: The moments of dependent random sums using copulas The work in this chapter is motivated by Mao and Zhao (2013). We extend their work by giving a general form for the kth moment of dependent random sums using copulas. In terms of copulas, we also use a more flexible one proposed by Nadarajah (2015). Moreover, a further extension is briefly discussed considering both the claim amounts (X) and the time-intervals between occurrences (W) being independently but non-identically distributed. A simulation is preformed in order to illustrate the computational efficiency of the derived expressions in this chapter.

Chapter 2

Generalized hyperbolic distributions

2.1 Introduction

The Generalized Hyperbolic (GH) distribution was introduced by Barndorff-Nielsen (1977) in a study of Aeolian sand deposits. The GH distribution contains many known distributions as limiting or particular cases. These include the hyperbolic (Bagnold, 1941), normal inverse Gaussian (Blaesild, 1977), GH skew Student's t (Aas and Haf, 2006), variance gamma (Madan and Seneta, 1990), generalized inverse Gaussian (due to Georges Henri Halphen), Student's t (Gosset, 1908) and normal (de Moivre, 1738; Gauss, 1809) distributions. The GH distribution has attracted widespread theoretical interest in recent years. It has also been generalized by many authors in recent years.

The GH distribution is defined as a normal mean-variance mixture symmetric model with the generalized inverse Gaussian (GIG) distribution as its mixing distribution. It has a property called 'semi-heavy tail' because its log-density forms a hyperbola rather than a parabola. The tail decays slower than that of the Gaussian distribution, but lighter than the non-Gaussian distribution which exhibits extremely heavy tail behavior. So, compared with the Gaussian distribution, the GH distribution provides the probability of fitting the observations with skewness and heavy tails. As a matter of fact, in recent years, the GH distribution has been more and has become more widely used in modeling the distribution of skewed and/or heavy tailed observations from a variety of areas, especially in the area of finance.

Applications of the GH distribution include a number of areas. In recent years, however, the GH distribution has been the first choice to model a variety of financial data. Such applications have included: models for circulatory transit times in pharmacokinetics (Weiss, 1984); modeling neural activity (Iyengar and Liao, 1997); pricing options on dividend paying instruments (Weron, 1999); cloud/aerosol particle size distribution (Alexandrov and Lacis, 2000); models for Brazilian asset returns (Fajardo and Farias, 2002); European and Asian option pricing (Predota, 2005); analysis of the log-return series of the Chinese stock prices (Li and Wu, 2007); models for energy efficient neurons (Berger et al., 2011); models for roller-coaster failure rates (Gupta and Viles, 2011); models for daily returns of the PSI20 (Rege and de Menezes, 2012); value at risk estimation for the South African mining index (Huang et al., 2014); modeling electricity price returns (Nwobi, 2014); models for equity returns (Socgnia and Wilcox, 2014); empirical analysis of Bucharest stock exchange (Baciu, 2015); evaluating risk in gold prices (Chinhamu et al., 2015).

Because of the increasing interest in terms of methodology and applications, we feel it is timely to provide a review of the GH distribution and its relatives. We review in Section 2.2 the GH distribution and over twenty related distributions, including the GH skew Student's t, mixture of GH, geometric GH, generalized generalized inverse Gaussian, logarithmic generalized inverse Gaussian, Ψ_2 hypergeometric generalized inverse Gaussian, confluent hypergeometric generalized inverse Gaussian, Nakagami generalized inverse Gaussian, generalized Nakagami generalized inverse Gaussian, extended generalized inverse Gaussian, exponential reciprocal generalized inverse Gaussian, gamma generalized inverse Gaussian and exponentiated generalized inverse Gaussian distributions. For each distribution, we give expressions for the probability density function (pdf), the cumulative distribution function (cdf), moments and moment generating function (mgf). These details are not given for trivial particular cases.

A real data application comparing some of the reviewed distributions is discussed in Section 2.3 . Some known software for the GH and related distributions are summarized in Section 2.4.

In this chapter, we have reviewed only univariate GH and related distributions. A future work is to review bivariate, multivariate, complex variate and matrix variate GH distributions.

2.2 The collection

Here, we provide a list of known distributions related to the GH distribution. The list is by no means complete, but we believe we have covered the most important and popular classes of GH distributions.

2.2.1 Generalized hyperbolic distribution

Barndorff-Nielsen (1977, 1978) introduced the GH distribution. Its pdf is

$$f(x) = C\sqrt{\delta^2 + (x-\mu)^2}^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right) e^{\beta(x-\mu)}$$

for $-\infty < x < \infty$, $-\infty < \lambda < \infty$, $-\infty < \alpha < \infty$, $-\infty < \beta < \infty$, $\delta \geq 0$ and $-\infty < \mu < \infty$, where C is the normalizing constant given by

$$C = \sqrt{\alpha^2-\beta^2}^{\lambda} / \left[\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)\right]$$

and $K_v(\cdot)$ denotes the modified Bessel function of the third kind and with index v. Furthermore, $\delta \geq 0$ and $\mid \beta \mid < \alpha$ if $\lambda > 0$; $\delta > 0$ and $\mid \beta \mid < \alpha$ if $\lambda = 0$; $\delta > 0$ and $\mid \beta \mid \leq \alpha$ if $\lambda < 0$. β is an asymmetry parameter, δ is a scale parameter and μ is a location parameter.

Other parameterizations exist for the GH distribution: $\xi = \left(1+\delta\sqrt{\alpha^2-\beta^2}\right)^{-1/2}$ and $\chi = \xi\beta/\alpha$; $\eta = \delta\sqrt{\alpha^2-\beta^2}$ and $\rho = \beta/\alpha$.

Because the GH distribution is a normal mean-variance mixture symmetric model with the GIG distribution as its mixing distribution, the random variable can be generated by the following algorithms:

1. simulate a random variable $Y \sim GIG(\lambda, \chi, \psi) = GIG(\lambda, \sigma^2, \alpha^2 - \beta^2)$;
2. simulate a standard normal random variable N;
3. return $X = \mu + \beta Y + \sqrt{Y}N$.

The mean and variance of the GH distribution are

$$E(X) = \mu + \frac{\delta\beta K_{\lambda+1}(\delta\gamma)}{\gamma K_{\lambda}(\delta\gamma)}$$

and

$$\mathrm{Var}(X) = \frac{\delta K_{\lambda+1}(\delta\gamma)}{\gamma K_{\lambda}(\delta\gamma)} + \frac{\beta^2\delta^2}{\gamma^2}\left[\frac{K_{\lambda+2}(\delta\gamma)}{K_{\lambda}(\delta\gamma)} - \frac{K^2_{\lambda+1}(\delta\gamma)}{K^2_{\lambda}(\delta\gamma)}\right],$$

respectively, where $\gamma = \sqrt{\alpha^2 - \beta^2}$. The mgf of the GH distribution is

$$M(t) = \frac{\gamma^\lambda e^{\mu x}}{[\alpha^2 - (\beta + x)^2]^{\lambda/2}} \frac{K_\lambda\left(\delta\sqrt{\alpha^2 - (\beta + x)^2}\right)}{K_\lambda(\delta\gamma)}.$$

Barndorff-Nielsen and Stelzer (2004) derived the nth central moment of the GH distribution as

$$E\left[(X-\mu)^n\right] = \frac{2^{\lceil \frac{n}{2} \rceil}\overline{\gamma}^\lambda \delta^{2\lceil \frac{n}{2} \rceil}\beta^m}{\sqrt{\pi}K_\lambda\left(\overline{\gamma}\right)\overline{\alpha}^{\lambda+\lceil \frac{n}{2} \rceil}} \sum_{k=0}^{\infty} \frac{2^k \overline{\beta}^{2k}\Gamma\left(k + \lceil n/2 \rceil + \frac{1}{2}\right)}{\overline{\alpha}^k(2k+m)!} K_{\lambda+k+\lceil n/2 \rceil}\left(\overline{\alpha}\right),$$

where μ is the mean, $\overline{\alpha} = \delta\alpha$, $\overline{\beta} = \delta\beta$, $\overline{\gamma} = \delta\gamma$, and $m = k \bmod 2$. Scott et al. (2008) derived the alternative form

$$E\left[(X-\mu)^n\right] = \sum_{\ell=\left[\frac{n+1}{2}\right]}^{n} \frac{n!\beta^{2\ell-n}}{(n-\ell)!(2\ell-n)!2^{k-\ell}} E\left(W^\ell\right),$$

where W is a generalized inverse Gaussian random variable with parameters λ, δ^2 and $\alpha^2 - \beta^2$.

2.2.2 Hyperbolic distribution

The hyperbolic (HP) distribution is the particular case of the GH distribution for $\lambda = 1$. Therefore, suppose Y is a random variable following the GIG distribution with parameters $(\lambda = 1, \chi, \psi)$, the hyperbolic distributed random variable, Z, can be generated as

$$(Z \mid Y) \sim N(\mu + \beta Y, Y),$$

where $-\infty < \beta < \infty$ and $-\infty < \mu < \infty$.

2.2.3 Log hyperbolic distribution

The log hyperbolic (LH) distribution has the specified density function proposed by Bhatia et al. (1988) in the study of liquid sprays. Its pdf is specified by

$$f(x) = \frac{\alpha}{x} \exp\left[-\alpha\sqrt{\delta^2 + (\ln x - \mu)^2} + \beta(\ln x - \mu)\right]$$

for $x > 0$. Its nth raw moment is

$$E\left(X^n\right) = e^{n\mu}\sqrt{\frac{\alpha^2-\beta^2}{\alpha^2-(\beta+n)^2}}\frac{K_1\left(\delta\sqrt{\alpha^2-(\beta+n)^2}\right)}{K_1\left(\delta\sqrt{\alpha^2-\beta^2}\right)}$$

provided $\alpha >\mid \beta + n \mid$.

The empirical studies in Bhatia et al. (1988) showed the LH distribution provided a superior fit to the observed data and suggested further application which required the detailed information about the droplet size variation throughout the spray. The LH distribution has been discussed in many applications on geophysics (Christiansen and Hartmann,1988; Wyrwoll and Smyth 1988; Sutherland and Lee, 1994; Bartholdy et al., 2007).

2.2.4 Normal inverse Gaussian distribution

The normal inverse Gaussian (NIG) distribution was introduced by Barndoff-Nielson (1997) in the studies of observations from turbulence and from finance. It is the particular case of the GH distribution for $\lambda = -1/2$. Suppose a random variable z follows the inverse Gaussian (IG) distribution with parameter γ and δ. Random variable Y is said to follow the NIG distribution if $Y \mid z \sim N(\mu + \beta z, z)$ for $-\infty < \beta < \infty$ and $-\infty < \mu < \infty$.

The NIG distribution can fit the distribution with skewness and heavy tails. Moreover, it converges to the Cauthy distribution, which exhibits extremely heavy tails, with α approaching zero. It can be more adequate to fit the distribution of considerably heavier tails than that of the GH distribution. Recently, many empirical studies have suggested that the NIG distribution provides an adequate fit to the financial data (Eriksson et al.,2009; Stentoft, 2008; Karlis and Lilleståöl, 2004; Karlis, 2002; Venter and de Jongh, 2002).

2.2.5 Variance gamma distribution

The variance gamma (hereafter VG) distribution is the particular case of the GH distribution for $\delta = 0$, which is obtained from the normal with mixing of scale parameter. It was first introduced by Madan and Seneta (1990) with an application to modeling asset returns and option pricing. Generalizations of the VG distribution are given in Zaks and Korolev (2013) and Krolev et al. (2015). The VG distribution has been widely applied to the area of finance. Such applications include: modeling returns for the Dow-Jones Industrial Average data (Hurst et al., 1997); modeling returns and pricing options on S&P's 500 Index (Madan, Carr and Chang, 1998); pricing Asian, lookback and barrier options (Avramidis and L'Ecuyer, 2006).

2.2.6 Generalized inverse Gaussian distribution

The generalized inverse Gaussian (hereafter GIG) distribution is the limiting case of the GH distribution for $\alpha\delta^2 \to \chi$, $\alpha - \beta = \frac{\psi}{2}$ and $\mu = 0$. Its pdf is

$$f(x) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda\left(\sqrt{\chi\psi}\right)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\psi x + \frac{\chi}{x}\right)\right\}.$$

Lemonte et al. (2011) introduced the reparameterization $\omega = \chi/2$ and $\eta = \psi/2$, so the pdf becomes

$$f(x) = Cx^{\lambda-1} \exp\left\{-\left(\eta x + \omega x^{-1}\right)\right\},$$

where the normalizing constant is $C = C(\lambda, \omega, \eta) = (\eta/\omega)^{\lambda/2} / \left\{2K_\lambda\left(2\sqrt{\eta\omega}\right)\right\}$. The corresponding cdf, mgf and nth moment are

$$F(x) = 1 - C\eta^{-\lambda}\Gamma\left(\lambda, \eta x; \eta\omega\right),$$

$$M(t) = \left(\frac{\eta}{\eta - t}\right)^{\lambda/2} \frac{K_\lambda\left(2\sqrt{(\eta - t)\omega}\right)}{K_\lambda\left(2\sqrt{\eta\omega}\right)}$$

and

$$E\left(X^n\right) = \left(\frac{\omega}{\eta}\right)^{n/2} \frac{K_{\lambda+n}\left(2\sqrt{\eta\omega}\right)}{K_\lambda\left(2\sqrt{\eta\omega}\right)},$$

respectively, where

$$\Gamma\left(a, x; b\right) = \int_x^\infty t^{a-1} e^{-t-\frac{b}{t}}\, dt.$$

The quantiles can be computed by the inserse of cumulative distribution function.

Since the GH distribution is a mixture with the GIG distribution as the mixing model, there has been a lot of research providing a comparison between the GH distribution and the GIG distribution. Some of the references have been listed in Section 2.1. A collection of particular or transformation cases of the GIG is also discussed in this chapter.

2.2.7 Student's t distribution

The Student's t distribution with ν degrees of freedom is the particular symmetric and heavy tailed case of the GH distribution for $\lambda = -\nu/2, \alpha = 0, \beta = 0$, and $\delta = \sqrt{\nu}$. The student's t distribution converges to the normal distribution as $\nu > 30$.

2.2.8 Normal distribution

The normal distribution is the particular symmetric case of the GH distribution for $\delta \to \infty$ and $\delta/\alpha \to \sigma^2$, which is consistent with the demonstration in 2.2.7. Moreover, the normal distribution is a particular case of $NIG(\mu, \alpha, \beta, \delta)$ distribution when $\beta = 0, \delta = \sigma^2 \cdot \alpha$ and $\alpha \to \infty$.

2.2.9 GH skew Student's t distribution

The GH skew Student's t (GHST) distribution is a limiting case of the GH distribution for $\lambda = -\nu/2$ and $\alpha \to | \beta |$. It is due to Aas and Haff (2006). Its pdf is

$$f(x) = \frac{\delta^v |\beta|^{(v+1)/2} K_{(v+1)/2}\left(\beta\sqrt{\delta^2 + (x-\mu)^2}\right) \exp\left[\beta(x-\mu)\right]}{2^{(v-1)/2}\Gamma(v/2)\sqrt{\pi}\left[\sqrt{\delta^2 + (x-\mu)^2}\right]^{(v+1)/2}}.$$

The mean and variance are

$$E(X) = \mu + \frac{\beta\delta^2}{v-2}$$

and

$$\text{Var}(X) = \frac{2\beta^2\delta^4}{(v-2)^2(v-4)} + \frac{\delta^2}{v-2},$$

respectively.

It is the only subclass, in our collection, of the GH family that has the property that one tail characterizes polynomial, and the other exhibits exponential behavior. Because of its special tail behavior, it has been used in measuring blood flow in the canine myocardium (Jones and Faddy, 2003); in portfolio optimization of the Dow Index (Hu and Kercheval, 2010); in forecasting expected shortfall for the S&P 500 Index and commercial stocks (Zhu and Galbraith, 2011); in modeling daily S&P 500 Index and TOPIX returns (Nakajima and Omori, 2012).

2.2.10 Mixture of GH distributions

Browne and McNicholas (2015) studied mixtures of GH distributions specified by the pdf

$$f(x) = C \sum_{i=1}^{N} w_i \sqrt{\delta_i^2 + (x-\mu_i)^2}^{\lambda_i - \frac{1}{2}} K_{\lambda_i - \frac{1}{2}} \left(\alpha_i \sqrt{\delta_i^2 + (x-\mu_i)^2} \right) e^{\beta_i (x-\mu_i)}$$

for $-\infty < x < \infty$, $-\infty < \lambda_i < \infty$, $-\infty < \alpha_i < \infty$, $-\infty < \beta_i < \infty$, $-\infty < \delta_i < \infty$ and $-\infty < \mu_i < \infty$, where C is the normalizing constant and w_i are weights. The expectation-maximization framework was used to estimate the parameters.

In real data analysis, Browne and McNicolas (2015) illustrated the mixture of GH distribution outperformed the mixture of skew-Gaussian distribution in clustering. Besides, the mixture of GH model was also suggested for other statistical learning analysis, such as semi-supervised classification, discriminant analysis, and density estimation.

2.2.11 Geometric GH distribution

Let $h(t) = 1/(1+t)$. Then the distribution specified by the characteristic function $h(-\ln\psi(t))$, where $\psi(t)$ is the characteristic function of the GH distribution, is referred to as the geometric GH distribution, see Klebanov and Rachev (2015) for details.

Since the Geometric distribution has been widely applied in modeling the distribution of the number of summands in random sums, the Geometric GH distribution is proposed for the further study on the distributions of sums of random number of random variables.

2.2.12 Generalized generalized inverse Gaussian distribution

The Generalized GIG (hereafter GGIG) distribution was introduced by Shakil et al. (2010) by solving the generalized Pearson differential equation. Its pdf is

$$f(x) = \frac{p}{2} \left(\frac{\alpha}{\beta} \right)^{v/2p} \frac{x^{v-1} \exp\left(-\alpha x^p - \beta x^{-p}\right)}{K_{v/p}\left(2\sqrt{\alpha\beta}\right)}$$

for $x > 0$, $\alpha > 0$, $\beta > 0$, $-\infty < v < \infty$ and $p > 0$.

The GGIG distribution has the characteristic of skewness to the right and exhibits most of the properties of skewed distributions. Shakil et al.(2010) applied the proposed GGIG distribution to a real problem in forestry in comparison with the performaces of the gamma, log-normal and inverse Gaussian distributions,respectively. The result showed the GGIG distribution provided the best fit. Many of our reviewed distributions in this chapter are also the particular cases of the GGIG distribution, which includes: the hyperbolic distribution for $p = 1$ and $v = 1$; the GIG distribution for $p = 1$; the reciprocal inverse Gaussian distribution for $p = 1$ and $v = \frac{1}{2}$; the hyperbola distribution for $p = 1$ and $v = 0$; the generalized gamma distribution for $\alpha = (1/\mu)^p$, $\beta = 0$ and $v = pk$. The corresponding cdf and nth moment are

$$F(x) = \frac{1}{2\left(\alpha\beta\right)^{v/2p} K_{v/p}\left(2\sqrt{\alpha\beta}\right)} \sum_{k=0}^{\infty} \frac{(-1)^k \left(\alpha\beta\right)^k}{k!} \gamma\left(\frac{v - pk}{p}, \alpha x^p\right)$$

and

$$E\left(X^n\right) = \left(\frac{\beta}{\alpha}\right)^{n/2p} \frac{K_{(n+v)/p}\left(2\sqrt{\alpha\beta}\right)}{K_{v/p}\left(2\sqrt{\alpha\beta}\right)},$$

respectively, where $\gamma(a, x)$ denotes the incomplete gamma function. The quantiles can be computed by the inverse of cumulative distribution function.

2.2.13 Logarithmic generalized inverse Gaussian distribution

Knowing that the shape of the GIG density function is rather similar to that of the log-normal distribution. In order to approaching 'normality', Sichel et al. (1997) introduced the logarithmic generalized inverse Gaussian (hereafter LGIG) distribution. Its pdf is

$$f(x) = \frac{1}{2K_a(b)} \exp\left(ax - b\cosh x\right)$$

for $-\infty < x < \infty$, $-\infty < a < \infty$ and $b > 0$. The mgf is

$$M(t) = \frac{K_{a+t}(b)}{K_a(b)}.$$

It is suggested for the application on modeling the distribution of automobile insurance claims due to disaster damage, or on modeling the logarithmic distribution of single ore values. Such ores can be as oil, gas and gold, whose formation has a hydrothermal character.

2.2.14 Ψ_2 hypergeometric generalized inverse Gaussian distribution

Inspired by the gamma, Weibull and inverse Gaussian distributions, which are all the combinations of the power function and the exponential function, Saboor et al. (2014) introduced Ψ_2 hypergeometric generalized inverse Gaussian distribution (hereafter PHGIG). Its pdf is

$$f(x) = \frac{1}{\Gamma^{(a,b,p_1,p_2,\alpha)}(m)} x^{m-1} \exp\left(-p_1 x - \frac{p_2}{x}\right) \Psi_2\left(a; \frac{b}{2}, \frac{b+1}{2}; -\frac{\alpha}{x}, -x\right)$$

for $x > 0$, $-\infty < m < \infty$, $p_1 > 0$, $p_2 > 0$, $m \neq 0$, $a \neq 0$ and $b \neq 0$, where

$$\Psi_2(a; b; c; x, y) = \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(a)_{j+m} x^j y^m}{(b)_j (c)_m j! m!}$$

and

$$\Gamma^{(a,b,p_1,p_2,\alpha)}(m) = 2\left(\frac{p_2}{p_1}\right)^{m/2} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(a)_{j+k}\left(-\alpha\sqrt{p_1/p_2}\right)^j (p_2/p_1)^{k/2}}{(b/2)_j \left(\frac{b+1}{2}\right)_k j! k!} K_{j-k-m}\left(2\sqrt{p_1 p_2}\right),$$

where $(\theta)_k = \theta(\theta+1)\cdots(\theta+k-1)$ denotes the ascending factorial. The mgf and the nth moment are

$$M(t) = \frac{2p_2^{m/2}(p_1 - t)^{m/2}}{\Gamma^{(a,b,p_1,p_2,\alpha)}(m)} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(a)_{j+k}\left[-\alpha\sqrt{\frac{p_1-t}{p_2}}\right]^j \left[\frac{p_2}{p_1-t}\right]^{k/2}}{(b/2)_j \left(\frac{b+1}{2}\right)_k j! k!} K_{j-k-m}\left(2\sqrt{(p_1-t)\, p_2}\right)$$

and

$$E\left(X^n\right) = \frac{\Gamma^{(a,b,p_1,p_2,\alpha)}(m+n)}{\Gamma^{(a,b,p_1,p_2,\alpha)}(m)},$$

respectively.

A real case study on modeling the distribution of 100 observed breaking stress of carbon fibres (in *Gba*) showed that the PHGIG distribution provided an adequate fit. Further applications are suggested in physical and biological series.

2.2.15 Confluent hypergeometric generalized inverse Gaussian distribution

Saboor et al. (2015) introduced another case of the combination of the power function and the exponential function, which is the confluent hypergeometric generalized inverse Gaussian distribution (hereafter CHGIG). Its pdf is

$$f(x) = C\left(a, b, p_1, p_2, \alpha\right) x^{m-1} \exp\left(-p_1 x - \frac{p_2}{x}\right) {}_1F_1(\lambda; b; -\alpha x)$$

for $x > 0$, $m > 0$, $p_1 > 0$, $p_2 > 0$, $\alpha \geq 0$, $\lambda \neq 0$ and $b \neq 0$, where

$$ {}_1F_1(a;b;x) = \sum_{j=0}^{\infty} \frac{(a)_j x^j}{(b)_j j!} $$

denotes the confluent hypergeometric function. The first two moments are

$$ E(X) = \frac{C\left(m+1, a, b, p_1, p_2, \alpha\right)}{C\left(m, a, b, p_1, p_2, \alpha\right)} $$

and

$$ E\left(X^2\right) = \frac{C\left(m+2, a, b, p_1, p_2, \alpha\right)}{C\left(m, a, b, p_1, p_2, \alpha\right)}. $$

The mgf is

$$ \begin{aligned} M\left(t\right) = {} & \frac{\Gamma(b)}{C\Gamma(\lambda)} \sum_{i=1}^{\infty} \frac{\left(p_1 - t\right)^{i-m} p_2^i}{i!} G_{3,3}^{2,2}\left(\frac{\alpha}{p_1 - t} \middle| \begin{array}{c} 1-m, 1-\lambda, -m+i+1 \\ 0, 1-m, 1-b, -m-i \end{array} \right) \\ & - \frac{\Gamma(b)}{C\Gamma(\lambda)} \sum_{i=1}^{\infty} \frac{\left(p_1 - t\right)^{i} p_2^{i+m}}{i!} G_{2,4}^{2,2}\left(p_2\alpha \middle| \begin{array}{c} 1-m, 1-\lambda \\ 0, 1-m, 1-b, -m-i \end{array} \right) \end{aligned} $$

provided that $p_1 > t$ and $p_2 > 0$, where $C = C\left(m, a, b, p_1, p_2, \alpha\right)$ and

$$ G_{p,q}^{m,n}\left(z \middle| \begin{array}{c} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{array} \right) = \frac{1}{2\pi \mathrm{i}} \int_L \frac{\prod_{j=1}^{m} \Gamma\left(b_j + s\right) \prod_{j=1}^{n} \Gamma\left(1 - a_j - s\right)}{\prod_{j=n+1}^{p} \Gamma\left(a_j + s\right) \prod_{j=m+1}^{q} \Gamma\left(1 - b_j - s\right)} z^{-s} ds $$

denotes the Meijer G function, $\mathrm{i} = \sqrt{-1}$ is the complex unit and L denotes an integration path, see Section 9.3 in Gradshteyn and Ryzhik (2000) for a description of this path.

The CHGIG distribution is shown to outperform the gamma, ESGamma, generalized gamma and the generalized inverse Gaussian distributions in modeling the 'Ball-bearing' data set. Details of the data set can be found at `http://homepage.tudelft.nl/n9d04/occ/510/oc_510.html`.

2.2.16 Nakagami generalized inverse Gaussian distribution

Both Nakagami distribution and generalized Nakagami distribution have shown their excellent performances on fitting the distribution of ultrasound echo envelope compared with the previously widely used Rayleigh model in Shankar (2001). In order to better describe ultrasonic tissues' heavy tail behavior, and motivated by the success of Nakagami-inverse Gaussian distribution (NIGD) (Agrawal, 2006), Agrawal and Karmeshu (2006) introduced the Nakagami generalized inverse Gaussian (hereafter NGIG) distribution. Its pdf is

$$f(x) = \frac{2m^m x^{2m-1}}{\theta^\gamma \Gamma(m) K_\gamma(\lambda/\theta)} \left[\frac{\theta^2\left(\lambda + 2mx^2\right)}{\lambda}\right]^{(\gamma-m)/2} K_{\gamma-m}\left(\frac{1}{\theta}\sqrt{\lambda\left(\lambda+2mx^2\right)}\right)$$

for $x > 0$, $\theta > 0$ and $\lambda > 0$. Its nth raw moment is

$$E\left(X^n\right) = \left(\frac{\theta}{m}\right)^{n/2} \frac{\Gamma\left(m+\frac{n}{2}\right)}{\Gamma(m)} \frac{K_{\gamma+\frac{n}{2}}(\lambda/\theta)}{K_\gamma(\lambda/\theta)}$$

for $n \geq 1$.

2.2.17 Generalized Nakagami generalized inverse Gaussian distribution

Following the previous work contributed by Agrawal and Karmeshu (2006), Gupta and Karmeshu (2015) introduced the generalized Nakagami generalized inverse Gaussian (hereafter GNGIG) distribution. Its pdf is

$$f(x) = \frac{2vm^m x^{2mv-1}}{\theta^\eta \Gamma(m) K_\eta(\lambda/\theta)} \left[\frac{\theta^2\left(\lambda + 2mx^{2v}\right)}{\lambda}\right]^{(\eta-m)/2} K_{\eta-m}\left(\frac{1}{\theta}\sqrt{\lambda\left(\lambda+2mx^{2v}\right)}\right)$$

for $x > 0$, $\lambda \geq 0$, $\theta > 0$ and $-\infty < \eta < \infty$. Its nth raw moment is

$$E\left(X^n\right) = \left(\frac{\theta}{m}\right)^{\frac{n}{2v}} \frac{\Gamma\left(m+\frac{n}{2v}\right)}{\Gamma(m)} \frac{K_{\eta+\frac{n}{2v}}(\lambda/\theta)}{K_\eta(\lambda/\theta)}$$

for $n \geq 1$.

2.2.18 Extended generalized inverse Gaussian distribution

The extended generalized inverse Gaussian (hereafter, EXGIG) distribution is generated from the generalized inverse Gaussian distribution as a power transformation in Jϕrgensen (1982). Thus, if $X \sim GIG(\lambda, \delta, \gamma)$, $Y = X^{1/\theta}$ with $\theta > 0$, then $Y \sim EGIG(\lambda, \delta, \gamma, \theta)$. The pdf of the EXGIG distribution due to Silva et al. (2006) is

$$f(x) = \left(\frac{\gamma}{\delta}\right)^\lambda \frac{\theta}{K_\lambda(\gamma\delta)} x^{\lambda\theta-1} \exp\left[-\frac{1}{2}\left(\delta^2 x^{-\theta} + \gamma^2 x^\theta\right)\right]$$

for $x > 0$, $-\infty < \lambda < \infty$, $\delta > 0$, $\gamma > 0$ and $\theta > 0$. Its nth raw moment is

$$E(X^n) = \left(\frac{\gamma}{\delta}\right)^{n/\theta} \frac{K_{\lambda+\frac{n}{2}}(\gamma\delta)}{K_\lambda(\gamma\delta)}.$$

The EXGIG distribution has many well-known particular cases, such as generalized gamma distribution, Weibull distribution, truncated normal distribution. Details can be found in Appendix A, Silva et al. (2006).

2.2.19 Exponential reciprocal generalized inverse Gaussian distribution

Gómez-Déniz et al. (2015) introduced the exponential reciprocal generalized inverse Gaussian (hereafter, ERGIG) distribution. Suppose $Y\ GIG(\gamma, \delta, \nu)$, and $Z = 1/Y$, which follows the reciprocal of the inverse Gaussian distribution (hereafter, RGIG) (γ, δ, ν). X is said to follow the exponential-reciprocal generalized inverse Gaussian distribution (hereafter, ERGIG) $(\gamma, \delta,)$, if $f(x) \propto \exp(x/z)$. Its pdf is

$$f(x) = \frac{\delta}{\gamma}\left(\frac{\gamma}{\sqrt{2x+\gamma^2}}\right)^{\nu+1} \frac{K_{\nu+1}\left(\delta\sqrt{2x+\gamma^2}\right)}{K_\nu(\delta\gamma)}$$

for $x > 0$, $-\infty < \nu < \infty$, $\gamma > 0$ and $\delta > 0$. The corresponding cdf and first two moments are

$$F(x) = \left(\frac{\gamma}{\delta}\right)^{\nu} \frac{K_\nu\left(\gamma\sqrt{2x+\gamma^2}\right)}{K_\nu(\delta\gamma)}$$

and

$$E(X) = \frac{\gamma}{\delta}\frac{K_{\nu-1}(\delta\gamma)}{K_\nu(\delta\gamma)},$$

$$E\left(X^2\right) = \frac{\gamma^2}{\delta^2}\frac{K_{\nu-2}(\delta\gamma)}{K_\nu(\delta\gamma)},$$

respectively. The quantiles can be computed by the inverse of cumulative distribution function.

2.2.20 Gamma generalized inverse Gaussian distribution

Gómez-Déniz et al. (2013) introduced a new class of distribution, the gamma generalized inverse Gaussian (hereafter GAGIG) distribution, which is the mixture of gamma distribution and the generalized inverse Gaussian distribution. Its pdf is

$$f(x) = \frac{(x\mu)^a}{x\Gamma(a)} \left(\frac{\psi}{\psi + 2\mu^2 x} \right)^{\frac{a+\lambda}{2}} \frac{K_{a+\lambda}\left(\psi/\mu_1\right)}{K_\lambda\left(\psi/\mu\right)}$$

for $x > 0$, $-\infty < \lambda < \infty$, $a > 0$, $\mu > 0$ and $\psi > 0$, where $\mu_1 = \mu\sqrt{\psi/\left(\psi + 2\mu^2 x\right)}$. Its first two moments are

$$E(X) = \frac{a}{\mu} \frac{K_{\lambda-1}\left(\psi/\mu\right)}{K_\lambda\left(\psi/\mu\right)}$$

and

$$E\left(X^2\right) = \frac{a^2}{\mu^2} \frac{K_{\lambda-2}\left(\psi/\mu\right)}{K_\lambda\left(\psi/\mu\right)} + \frac{a}{\mu^2} \frac{K_{\lambda-2}\left(\psi/\mu\right)}{K_\lambda\left(\psi/\mu\right)}.$$

Because of the possibility of measuring risk heterogeneity, the GAGIG distribution is considered competitive to the log-normal, gamma, inverse Gaussian, Frechet, Weibull and other widely used distributions in modeling the insurance claim amount caused by the natural disaster.

2.2.21 Exponentiated generalized inverse Gaussian distribution

Lemonte and Cordeiro (2011) introduced the exponentiated generalized inverse Gaussian (EGIG) distribution. Let G denote the cdf of a GIG distribution, that is

$$G(x) = 1 - C\eta^{-\lambda}\Gamma\left(\lambda, \eta x; \eta w\right)$$

for $x > 0$ with C as defined in Section 2.2.6. The EGIG distribution is defined by the cdf

$$F(x) = G^\beta(x)$$

for $x > 0$ and $\beta > 0$. The corresponding pdf is

$$f(x) = \beta G^{\beta-1}(x) g(x)$$

for $x > 0$, where g denotes the pdf of the GIG distribution. Therefore, the GIG distribution is a particular case of the EGIG distribution when $\beta = 1$. If $w = 0$, the EGIG distribution is referred to as the exponentiated gamma distribution (hereafter EGamma) (Nadarajah and Kotz 2006). It is obvious that the EGIG distribution can be applied to a wide range of scenarios compared with the

GIG distribution. The nth moment and the mgf of the EGIG distribution are

$$E\left(X^n\right) = \left(\frac{w}{\eta}\right)^{n/2} \sum_{j=0}^{\infty} p_j \frac{K_{\lambda+n+j}\left(2\sqrt{\eta w}\right)}{K_{\lambda+j}\left(2\sqrt{\eta w}\right)}$$

and

$$M(t) = \left(\frac{\eta}{\eta - t}\right)^{\lambda/2} \sum_{j=0}^{\infty} p_j \left(\frac{\eta}{\eta - t}\right)^{j/2} \frac{K_{\lambda+j}\left(2\sqrt{(\eta - t)w}\right)}{K_{\lambda+j}\left(2\sqrt{\eta w}\right)},$$

respectively, where

$$p_j = \beta v_j C(\lambda, w, \eta)/C(\lambda + j, w, \eta)$$

and v_j are determined by

$$[1 - F(x)]^{\beta - 1} = \sum_{j=0}^{\infty} v_j x^j.$$

The quantiles can be computed by the inverse of cumulative distribution function.

2.3 Real data application

Here, we compare the performances of some of the GH related distributions in Section 2.2 using a real data set. The data we use are S&P / IFC (Standard & Poor's / International Finance Corporation) global daily price indices in United States dollars for South Africa. The data cover the period from the 1st of January 1996 to the 31st of October 2008. The data were obtained from the database Datastream. Following common practice, daily log returns were computed as first order differences of logarithms of daily price indices.

We fitted the following distributions: the five-parameter GH distribution, the four-parameter HP distribution, the four-parameter NIG distribution, the four-parameter VG distribution, the three-parameter GIG distribution, the four-parameter GHST distribution, the four-parameter GGIG distribution, the five-parameter PHGIG distribution, the five-parameter CHGIG distribution, the five-parameter GNGIG distribution, the four-parameter EGIG distribution and the four-parameter GAGIG distribution. Each distribution was fitted by the method of maximum likelihood. Table 2.1 gives the log-likelihood values, values of the Akaike Information Criterion (AIC), values of the

Bayesian Information Criterion (BIC) and p-values based on the Kolmogorov-Smirnov statistic. The AIC is due to Akaike (1974). The BIC is due to Schwarz (1978). The smaller the values of these criteria the better the fit. For more discussion on these criteria, see Burnham and Anderson (2004) and Fang (2011).

We can see that the five-parameter GH distribution gives the smallest AIC, smallest BIC and the largest p-value, which is greater than 0.05. That is to say, we can not reject the Null hypothesis of the Kolmogorov test at the 5% level of significance. The Null hypothesis is that the sample is drawn from the reference distribution. So, the GH distribution gives the best fit to the data. The GH distribution provides significantly better fits than all of its limiting and particular cases, including the HP, NIG, VG, GIG and GHST distributions. The GGIG distribution gives the largest AIC and the smallest p-value. The PHGIG distribution gives the largest BIC. So, these two distributions may be thought to give the worst fits. At the five percent level of significance, the GH, CHGIG and GNGIG distributions give adequate fits.

2.4 Computer software

Software for the GH and related distributions are widely available. Some software available from the R package (R Development Core Team, 2016) are:

- the package ghyp due to David Luethi and Wolfgang Breymann. According to the authors, the package provides "Detailed functionality for working with the univariate and multivariate Generalized Hyperbolic distribution and its special cases (Hyperbolic (hyp), Normal Inverse Gaussian (NIG), Variance Gamma (VG), skewed Student-t and Gaussian distributions). Especially, it contains fitting procedures, an AIC-based model selection routine, and functions for the computation of density, quantile, probability, random variates, expected shortfall and some portfolio optimization and plotting routines as well as the likelihood ratio test. In addition, it contains the Generalized Inverse Gaussian distribution".
- the package GeneralizedHyperbolic due to David Scott. According to the author, the package provides "functions for the hyperbolic and related distributions. Density, distribution and quantile functions and random number generation are provided for the hyperbolic distribution, the generalized hyperbolic distribution, the generalized inverse Gaussian distribution and the skew-Laplace distribution. Additional functionality is provided for the hyperbolic distribution, normal inverse Gaussian distribution and generalized inverse Gaussian distribution, including

fitting of these distributions to data. Linear models with hyperbolic errors may be fitted using hyperblmFit".

- the package HyperbolicDist due to David Scott. The package provides a collection of functions for working with the hyperbolic and related distributions.

  For the hyperbolic distribution functions, the package provides "the density function, distribution function, quantiles, random number generation and fitting the hyperbolic distribution to data (hyperbFit). The function hyperbChangePars will interchange parameter values between different parameterisations. The mean, variance, skewness, kurtosis and mode of a given hyperbolic distribution are given by hyperbMean, hyperbVar, hyperbSkew, hyperbKurt, and hyperbMode respectively. For assessing the fit of the hyperbolic distribution to a set of data, the log-histogram is useful. Q-Q and P-P plots are also provided for assessing the fit of a hyperbolic distribution. A Cramer-von-Mises test of the goodness of fit of data to a hyperbolic distribution is given by hyperbCvMTest. S3 print, plot and summary methods are provided for the output of hyperbFit".

  For the generalized hyperbolic distribution functions, the package provides "the density function, distribution function, quantiles, and for random number generation. The function ghypChangePars will interchange parameter values between different parameterisations. The mean, variance, and mode of a given generalized hyperbolic distribution are given by ghypMean, ghypVar, ghypSkew, ghypKurt, and ghypMode respectively. Q-Q and P-P plots are also provided for assessing the fit of a generalized hyperbolic distribution".

  For the generalized inverse Gaussian distribution functions, the package provides "the density function, distribution function, quantiles, and for random number generation. The function gigChangePars will interchange parameter values between different parameterisations. The mean, variance, skewness, kurtosis and mode of a given generalized inverse Gaussian distribution are given by gigMean, gigVar, gigSkew, gigKurt and gigMode respectively. Q-Q and P-P plots are also provided for assessing the fit of a generalized inverse Gaussian distribution".

- the package frmqa due to Thanh T. Tran. According to the author, the package comprises of "R and C++ functions which deal with issues relating to financial risk management and quantitative analysis by applying uni- and multi-variate generalized hyperbolic and related distributions. These issues are approached from both directions: analytical (i.e., deriving and programming analytic formulae) and numerical. Note that the latter appears to be the only approach currently used in practice because of the intractability of these families of distributions, which is caused by the presence of the modified Bessel function of the third kind

BesselK $K_\lambda(z)$, in their density functions. In this package, the naming of special functions and their related functions (e.g., BesselK and incomplete BesselK functions) follows the convention used in the R package gsl".

- the package VarianceGamma due to David Scott and Christine Yang Dong. According to the authors, the package provides "functions for the variance gamma distributions. Density, distribution and quantile functions. Functions for random number generation and fitting of the variance gamma to data. Also, functions for computing moments of the variance gamma distribution of any order about any location. In addition, there are functions for checking the validity of parameters and to interchange different sets of parameterizatons for the variance gamma distribution".

- the package SkewHyperbolic due to David Scott. According to the author, the package provides "a collection of functions for working with the skew hyperbolic Student t-distribution. Functions are provided for the density function (dskewhyp), distribution function (pskewhyp), quantiles (qskewhyp) and random number generation (rskewhyp). There are functions that fit the distribution to data (skewhypFit). The mean, variance, skewness, kurtosis and mode can be found using the functions skewhypMean, skewhypVar, skewhypSkew, skewhypKurt and skewhypMode respectively, and there is also a function to calculate moments of any order skewhypMom. To assess goodness of fit, there are functions to generate a Q-Q plot (qqskewhyp) and a P-P plot (ppskewhyp). S3 methods print, plot and summary are provided for the output of skwewhypFit".

- the package GIGrvg due to Josef Leydold and Wolfgang Hormann. According to the authors, the package provides "Generator and density function for the Generalized Inverse Gaussian (GIG) distribution".

- the package SkewHyperbolic due to David Scott and Fiona Grimson. According to the authors, the package provides "functions for the density function, distribution function, quantiles and random number generation for the skew hyperbolic t-distribution. There are also functions that fit the distribution to data. There are functions for the mean, variance, skewness, kurtosis and mode of a given distribution and to calculate moments of any order about any centre. To assess goodness of fit, there are functions to generate a Q-Q plot, a P-P plot and a tail plot".

- the package MixGHD due to Cristina Tortora, Ryan P. Browne, Brian C. Franczak, and Paul D. McNicholas. According to the authors, the package carries out "model-based clustering, classification and discriminant analysis using five different models. The models are all based on the generalized hyperbolic distribution. The first model 'MGHD' is the classical mixture of generalized hyperbolic distributions. The 'MGHFA' is the mixture of generalized hyperbolic factor analyzers for high dimensional data sets. The 'MSGHD', mixture of multiple scaled generalized hyperbolic distributions. The 'cMSGHD' is a 'MSGHD' with convex contour plots. The 'MCGHD', mixture of coalesced generalized hyperbolic distributions is a new more flexible model".

- the package QRM due to Bernhard Pfaff, Marius Hofert, Alexander McNeil, and Scott Ulmann. Among other roles, the package "computes values of density and random numbers for uni- and multivariate Generalized Hyperbolic distribution in new QRM parameterization (chi, psi, gamma) and in standard parametrization (alpha, beta, delta)" and "updates estimates of mixing parameters in EM estimation of generalized hyperbolic".
- the package QRMlib due to Alexander McNeil and Scott Ulmann. Among other roles, the package "computes values of density and random numbers for uni- and multivariate Generalized Hyperbolic distribution in new QRM parameterization (chi, psi, gamma)" and "calculates moments of univariate generalized inverse Gaussian (GIG) distribution".

Distribution	$\ln L$	AIC	BIC	p-value
GH	227.2	-444.4	-413.8	0.086
HP	220.1	-432.2	-407.7	0.039
NIG	219.9	-431.8	-407.3	0.031
VG	219.7	-431.4	-406.9	0.030
GIG	218.9	-431.8	-413.4	0.031
GHST	220.3	-432.6	-408.1	0.040
GGIG	219.0	-430.0	-405.5	0.001
PHGIG	221.7	-433.4	-402.8	0.044
CHGIG	221.8	-433.6	-403.0	0.064
GNGIG	222.8	-435.6	-405.0	0.073
GAGIG	219.2	-430.4	-405.9	0.012
EGIG	219.1	-430.2	-405.7	0.003

Table 2.1: Fitted distributions.

Chapter 3

Smallest Pareto order statistics

3.1 Introduction

Let $X_1, X_2, \ldots, X_n$ be independent but non-identical random variables from one population and let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$. Let $Y_1, Y_2, \ldots, Y_n$ be independent but non-identical random variables from another population and let $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. There has been little work on the stochastic comparison of $X_{1:n}$ and $Y_{1:n}$. The work that we are aware of are: Khaledi and Kochar (2006), Fang and Tang (2014), Li and Li (2015) and Torrado (2015). Balakrishnan and Zhao (2013) provide an excellent review of other known work.

Beyond statistical theory, smallest order statistics arise in many applied areas. Some recent applications have included: empirical studies of price dispersion on the Internet (Warin and Leiter, 2012); utility maximization frameworks for fair and efficient multicasting in multicarrier wireless cellular networks (Liu et al., 2013); degradation pattern prediction of a polymer electrolyte membrane fuel cell stack (Bae et al., 2014).

The aim of this chapter is to study orderings between $X_{1:n}$ and $Y_{1:n}$ when $X_1, X_2, \ldots, X_n$ are independent but non-identical Pareto random variables from one population and $Y_1, Y_2, \ldots, Y_n$ are independent but non-identical Pareto random variables from another population.

Pareto distributions are the most popular models in finance, economics and related areas. Pareto distributions are known as the mixtures of exponential distributions with Gamma mixing weights. They are commonly used to model random variables like income, risk and prices as they often exhibit heavy tails. Areas where Pareto random variables arise include: measures of the sales quantities in analysing the competition between Amazon and BarnesandNoble(Chevalier and Goolsbee, 2003); Ruin theory and reinsurance pricing (Morales, 2005; Cai et. al., 2017; Grahovac, 2018); the VaR and CVaR-based optimization in insurance contract (Asimit et al., 2017); bias reduction in estimation for automobile insurance portfolio (Beirlant, 2018). Often comparison of the smallest Pareto order statistics has practical appeal in the area of insurance, as the insurer would like the claim to be as small as possible. Suppose $X_1, X_2, \ldots, X_n$ are insurance claims made over a fixed period by a population with one characteristic (e.g., females) and $Y_1, Y_2, \ldots, Y_n$ are insurance claims made over the same fixed period by a population with another characteristic (e.g., males). Then $X_{1:n}$ and $Y_{1:n}$ are the smallest insurance claims made and it makes sense to compare them to see which characteristic is associated with smaller claims. We shall return to this example in Section 3.5.

We consider three of the most popular Pareto distributions: the Pareto distribution of type I, the Pareto distribution of type II and the Pareto distribution of type IV. We also briefly discuss results for the Feller Pareto and discrete Pareto distributions.

A Pareto random variable of type I with shape parameter a and scale parameter b denoted by $PI(a,b)$ has the probability density, cumulative distribution and hazard rate functions specified by

$$f(x) = \frac{a}{b}\left(\frac{x}{b}\right)^{-a-1},$$

$$F(x) = 1 - \left(\frac{x}{b}\right)^{-a}$$

and

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{a}{x},$$

respectively, for $x > b > 0$ and $a > 0$.

A Pareto random variable of type II with shape parameter a and scale parameter b denoted by $PII(a,b)$ has the probability density, cumulative distribution and hazard rate functions specified by

$$f(x) = \frac{a}{b}\left(1+\frac{x}{b}\right)^{-a-1},$$

$$F(x) = 1 - \left(1+\frac{x}{b}\right)^{-a}$$

and

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{a}{x+b},$$

respectively, for $x > 0$, $a > 0$ and $b > 0$.

A Pareto random variable of type IV with shape parameters a, b and scale parameter c denoted by $PIV(a, b, c)$ has the probability density, cumulative distribution and hazard rate functions specified by

$$f(x) = \frac{a}{bc^{1/b}} x^{\frac{1}{b}-1} \left[1 + \left(\frac{x}{c}\right)^{1/b}\right]^{-a-1},$$

$$F(x) = 1 - \left[1 + \left(\frac{x}{c}\right)^{1/b}\right]^{-a}$$

and

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{ax^{\frac{1}{b}-1}}{b\left(c^{\frac{1}{b}} + x^{\frac{1}{b}}\right)},$$

respectively, for $x > 0$, $a > 0$, $b > 0$ and $c > 0$.

Let $\overline{F}_i$ and f_i denote the survival and probability density functions of X_i. Let $\overline{G}_i$ and g_i denote the survival and probability density functions of Y_i. Then the survival and probability density functions of $X_{1:n}$ are

$$\overline{F}_{1:n}(x) = \prod_{i=1}^{n} \overline{F}_i(x)$$

and

$$f_{1:n}(x) = \left[\prod_{i=1}^{n} \overline{F}_i(x)\right] \left[\sum_{i=1}^{n} \frac{f_i(x)}{\overline{F}_i(x)}\right],$$

respectively. Similarly, the survival and probability density functions of $Y_{1:n}$ are

$$\overline{G}_{1:n}(x) = \prod_{i=1}^{n} \overline{G}_i(x)$$

and

$$g_{1:n}(x) = \left[\prod_{i=1}^{n} \overline{G}_i(x)\right] \left[\sum_{i=1}^{n} \frac{g_i(x)}{\overline{G}_i(x)}\right],$$

respectively.

The orderings that we consider are: $X_{1:n}$ is smaller than $Y_{1:n}$ in stochastic order denoted by $X_{1:n} \leq_{st} Y_{1:n}$ if $\overline{F}_{1:n}(x) \leq \overline{G}_{1:n}(x)$ for all x; $X_{1:n}$ is smaller than $Y_{1:n}$ in hazard rate order denoted by $X_{1:n} \leq_{hr} Y_{1:n}$ if $f_{1:n}(x)/\overline{F}_{1:n}(x) \geq g_{1:n}(x)/\overline{G}_{1:n}(x)$ for all x; $X_{1:n}$ is smaller than $Y_{1:n}$ in reversed hazard rate order denoted by $X_{1:n} \leq_{rh} Y_{1:n}$ if $f_{1:n}(x)/\left[1 - \overline{F}_{1:n}(x)\right] \leq g_{1:n}(x)/\left[1 - \overline{G}_{1:n}(x)\right]$ for all x; $X_{1:n}$ is smaller than $Y_{1:n}$ in likelihood ratio order denoted by $X_{1:n} \leq_{lr} Y_{1:n}$ if $g_{1:n}(x)/f_{1:n}(x)$ is increasing in x for all x for which the ratio is well defined; $X_{1:n}$ is smaller than $Y_{1:n}$ in increasing convex order denoted by $X_{1:n} \leq_{icx} Y_{1:n}$ if $\int_t^\infty \overline{F}_{1:n}(x)dx \leq \int_t^\infty \overline{G}_{1:n}(x)dx$ for all t; $X_{1:n}$ is smaller than $Y_{1:n}$ in mean residual life order denoted by $X_{1:n} \leq_{mrl} Y_{1:n}$ if $\int_t^\infty x f_{1:n}(x)dx/\overline{F}_{1:n}(t) \leq \int_t^\infty x g_{1:n}(x)dx/\overline{G}_{1:n}(t)$ for all $t > 0$. Further details of these orderings can be found in Shaked and Shanthikumar (2007).

In this chapter, we derive if and only if conditions for stochastic, hazard rate, likelihood ratio, reversed hazard rate, increasing convex and mean residual life orderings of $X_{1:n}$ and $Y_{1:n}$. Section 3.2 gives the results when $X_{1:n}$ is a minimum of independent Pareto type I random variables and $Y_{1:n}$ is a minimum of another set of independent Pareto type I random variables with different shape and scale parameters. Section 3.3 gives the results when $X_{1:n}$ is a minimum of independent Pareto type II random variables and $Y_{1:n}$ is a minimum of another set of independent Pareto type II random variables with different shape and scale parameters. Section 3.4 gives the results when $X_{1:n}$ is a minimum of independent Pareto type IV random variables and $Y_{1:n}$ is a minimum of another set of independent Pareto type IV random variables with different shape and scale parameters. Section 3.5 presents an application to insurance claims data. Section 3.6 gives further results on stochastic orderings, discussion and future work.

3.2 Results for the smallest of Pareto type I random variables

Let $X_i \sim PI(a_i, b_i)$, $i = 1, 2, \ldots, n$ be independent Pareto type I random variables. Let $Y_i \sim PI(c_i, d_i)$, $i = 1, 2, \ldots, n$ be independent Pareto type I random variables. Let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$

and $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. Note that

$$\overline{F}_{1:n}(x) = \prod_{i=1}^{n} \left(\frac{x}{b_i}\right)^{-a_i} \tag{3.1}$$

and

$$f_{1:n}(x) = \left(\sum_{i=1}^{n} \frac{a_i}{x}\right) \left[\prod_{i=1}^{n} \left(\frac{x}{b_i}\right)^{-a_i}\right] \tag{3.2}$$

for $x > \max(b_1, b_2, \ldots, b_n)$, $b_i > 0$ and $a_i > 0$. Similarly,

$$\overline{G}_{1:n}(x) = \prod_{i=1}^{n} \left(\frac{x}{d_i}\right)^{-c_i} \tag{3.3}$$

and

$$g_{1:n}(x) = \left(\sum_{i=1}^{n} \frac{c_i}{x}\right) \left[\prod_{i=1}^{n} \left(\frac{x}{d_i}\right)^{-c_i}\right] \tag{3.4}$$

for $x > \max(d_1, d_2, \ldots, d_n) > 0$, $d_i > 0$ and $c_i > 0$.

As noted, $X_{1:n}$ and $Y_{1:n}$ are defined over different domains. Any comparison of these random variables should be over a common domain. In our results, we take the common domain of $X_{1:n}$ and $Y_{1:n}$ as (e, ∞), where $e = \max(\max(b_1, b_2, \ldots, b_n), \max(d_1, d_2, \ldots, d_n))$. Throughout this section, we let $a = a_1 + a_2 + \cdots + a_n$, $c = c_1 + c_2 + \cdots + c_n$, $\alpha = \prod_{i=1}^{n} b_i^{a_i}$, $\beta = \prod_{i=1}^{n} d_i^{c_i}$, $x_0 = (\beta/\alpha)^{1/(c-a)}$ and $y_0 = \{\beta(a-1)/[\alpha(c-1)]\}^{1/(c-a)}$.

Theorem 3.2.1 gives an if and only if condition for stochastic ordering of $X_{1:n}$ and $Y_{1:n}$. Theorem 3.2.2 gives an if and only if condition for hazard rate ordering. Theorem 3.2.3 gives an if and only if condition for likelihood ratio ordering. Theorem 3.2.4 gives an if and only if condition for reversed hazard rate ordering. Theorem 3.2.5 gives an if and only if condition for increasing convex ordering. Theorem 3.2.6 gives an if and only if condition for mean residual life ordering.

Theorem 3.2.1. *$X_{1:n} \leq_{st} Y_{1:n}$ if and only if either $c = a$ and $\beta/\alpha \geq 1$ or $c < a$ and $x_0 \leq e$.*

Proof: By (3.1) and (3.3),

$$\frac{\overline{F}_{1:n}(x)}{\overline{G}_{1:n}(x)} = x^{c-a} \frac{\alpha}{\beta} \leq 1$$

for all $x \geq e$ if and only if either $c = a$ and

$$\frac{\beta}{\alpha} \geq 1$$

or $c < a$ and $x_0 \leq e \leq x$. The result follows. □

Theorem 3.2.2. *$X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $a \geq c$.*

Proof: By (3.1) to (3.4),

$$\frac{f_{1:n}(x)/\overline{F}_{1:n}(x)}{g_{1:n}(x)/\overline{G}_{1:n}(x)} = \frac{a}{c}.$$

The result follows. □

Theorem 3.2.3. *$X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $a > c$.*

Proof: By (3.2) and (3.4),

$$\frac{f_{1:n}(x)}{g_{1:n}(x)} = x^{c-a}\frac{a\alpha}{c\beta}.$$

The result follows. □

Theorem 3.2.4. *$X_{1:n} \leq_{rh} Y_{1:n}$ if and only if either $c < a$ and $h(e) \leq 0$ or $c = a$ and $\alpha \leq \beta$, where $h(x) = a\alpha x^c - c\beta x^a + (c-a)\alpha\beta$.*

Proof: By (3.1) to (3.4),

$$\frac{f_{1:n}(x)/\left[1-\overline{F}_{1:n}(x)\right]}{g_{1:n}(x)/\left[1-\overline{G}_{1:n}(x)\right]} = \frac{a\alpha x^{c-a}\left(1-\beta x^{-c}\right)}{c\beta\left(1-\alpha x^{-a}\right)} \leq 1.$$

Let $h(x) = a\alpha x^c - c\beta x^a + (c-a)\alpha\beta$. The above inequality function holds if and only if $h(x) \leq 0$ for all $x \geq e$.

If $c = a$, the condition can be simplified to $\alpha \leq \beta$.

If $c < a$, we take the first and the second derivative of $h(x)$ and obtain $h^{'}(x) = ac\alpha x^{c-1} - ac\beta x^{a-1}$, $h^{''}(x) = acx^{a-2}\left[\alpha(c-1)x^{c-a} - \beta(a-1)\right]$. Note that $h^{'}(x_0) = 0$ and $h^{''}(x_0) = acx_0^{a-2}\beta(c-a) < 0$ when $c < a$. So, the $h^{'}(x_0) = 0$ is the maximum value of $h^{'}(x)$. Thus, $h(x)$ is a non-increasing function. Because $e \leq x$ for all x, once $h(e) \leq 0$, $h(x) \leq 0$ for all x, thus $X_{1:n} \leq_{rh} Y_{1:n}$. □

Theorem 3.2.5. *Suppose $a > 1$ and $c > 1$. Then $X_{1:n} \leq_{icx} Y_{1:n}$ if and only if either $c = a$ and $\beta \geq \alpha$ or $c < a$ and $y_0 \leq e$.*

Proof: By (3.1) and (3.3),

$$\int_t^{\infty} \overline{F}_{1:n}(x)dx = \frac{\alpha}{a-1} t^{1-a}$$

and

$$\int_t^{\infty} \overline{G}_{1:n}(x)dx = \frac{\beta}{c-1} t^{1-c}.$$

Note that $\frac{\alpha}{a-1} t^{1-a} \leq \frac{\beta}{c-1} t^{1-c}$ if and only if either $c = a$ and $\beta \geq \alpha$ or $c < a$ and $y_0 \leq e$. □

Theorem 3.2.6. *Suppose $a > 1$ and $c > 1$. Then $X_{1:n} \leq_{mrl} Y_{1:n}$ if and only if $c \leq a$.*

Proof: By (3.1) to (3.4),

$$\frac{\int_t^{\infty} x f_{1:n}(x)dx}{\overline{F}_{1:n}(t)} = \frac{at}{a-1}$$

and

$$\frac{\int_t^{\infty} x g_{1:n}(x)dx}{\overline{G}_{1:n}(t)} = \frac{ct}{c-1}.$$

Note that $at/(a-1) \leq ct/(c-1)$ if and only if $c \leq a$. □

We can observe the following from Theorems 3.2.1 to 3.2.6: i) the hazard rate and likelihood ratio orderings are equivalent; ii) if $a > 1$ and $c > 1$ then the hazard rate, likelihood ratio and mean residual life orderings are equivalent; iii) the stochastic ordering implies hazard rate and likelihood ratio orderings; iv) the reversed hazard rate ordering implies hazard rate and likelihood ratio orderings; v) if $a > 1$ and $c > 1$ then increasing convex ordering implies mean residual life ordering.

3.3 Results for the smallest of Pareto type II random variables

Note that the Pareto distribution with type II is also known as the Lomax distribution. Let $X_i \sim PII\,(a_i, b_i)$, $i = 1, 2, \ldots, n$ be independent Pareto type II random variables. Let $Y_i \sim PII\,(c_i, d_i)$, $i = 1, 2, \ldots, n$ be independent Pareto type II random variables. Let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$

and $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. Note that

$$\overline{F}_{1:n}(x) = \prod_{i=1}^{n} \left(1 + \frac{x}{b_i}\right)^{-a_i}, \tag{3.5}$$

$$f_{1:n}(x) = \left(\sum_{i=1}^{n} \frac{a_i}{x + b_i}\right) \left[\prod_{i=1}^{n} \left(1 + \frac{x}{b_i}\right)^{-a_i}\right], \tag{3.6}$$

$$\overline{G}_{1:n}(x) = \prod_{i=1}^{n} \left(1 + \frac{x}{d_i}\right)^{-c_i} \tag{3.7}$$

and

$$g_{1:n}(x) = \left(\sum_{i=1}^{n} \frac{c_i}{x + d_i}\right) \left[\prod_{i=1}^{n} \left(1 + \frac{x}{d_i}\right)^{-c_i}\right] \tag{3.8}$$

for $x > 0$, $a_i > 0$, $b_i > 0$, $c_i > 0$ and $d_i > 0$.

Theorem 3.3.1 gives an if and only if condition for stochastic ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.2 gives an if and only if condition for hazard rate ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.3 gives an if and only if condition for likelihood ratio ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.4 gives an if and only if condition for reversed hazard rate ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.5 gives an if and only if condition for increasing convex ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.6 gives an if and only if condition for mean residual life ordering when $b_i = d_i =$ constant for all i. Theorem 3.3.7 gives a condition for stochastic ordering when $a_i = c_i =$ constant for all i. Theorem 3.3.8 gives a condition for hazard rate ordering when $a_i = c_i =$ constant for all i.

Theorem 3.3.1. *Suppose $b_i = d_i = c$ for all i. Then $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (3.5) and (3.7),

$$\frac{\overline{F}_{1:n}(x)}{\overline{G}_{1:n}(x)} = \left(\frac{c}{x + c}\right)^{\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} c_i}.$$

The result follows. □

Theorem 3.3.2. *Suppose $b_i = d_i = c$ for all i. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (3.5) to (3.8),

$$\frac{f_{1:n}(x)/\overline{F}_{1:n}(x)}{g_{1:n}(x)/\overline{G}_{1:n}(x)} = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} c_i}.$$

The result follows. □

Theorem 3.3.3. *Suppose $b_i = d_i = c$ for all i. Then $X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i > \sum_{i=1}^{n} c_i$.*

Proof: By (4.2) and (4.4),

$$\frac{f_{1:n}(x)}{g_{1:n}(x)} = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} c_i} \left(\frac{c}{x+c} \right)^{\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} c_i}.$$

The result follows. □

Theorem 3.3.4. *Suppose $b_i = d_i = c$ for all i. Then $X_{1:n} \leq_{rh} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (3.5) to (3.8),

$$\frac{f_{1:n}(x)/\left[1-\overline{F}_{1:n}(x)\right]}{g_{1:n}(x)/\left[1-\overline{G}_{1:n}(x)\right]} = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} c_i} \frac{\left(\frac{x}{c}+1\right)^{\sum_{i=1}^{n} c_i} - 1}{\left(\frac{x}{c}+1\right)^{\sum_{i=1}^{n} a_i} - 1}.$$

This can be rewritten as

$$\frac{f_{1:n}(x)/\left[1-\overline{F}_{1:n}(x)\right]}{g_{1:n}(x)/\left[1-\overline{G}_{1:n}(x)\right]} = \frac{h\left(\sum_{i=1}^{n} c_i\right)}{h\left(\sum_{i=1}^{n} a_i\right)}$$

for

$$h(\alpha) = \left[\left(\frac{x}{c}+1\right)^{\alpha} - 1\right]/\alpha.$$

Note that we can write

$$h^{'}(\alpha) = -\alpha^{-2} q(\alpha),$$

where

$$q(\alpha) = \left(\frac{x}{c}+1\right)^{\alpha} - 1 - \alpha\left(\frac{x}{c}+1\right)^{\alpha}\log\left(\frac{x}{c}+1\right).$$

Note also that

$$q^{'}(\alpha) = -\alpha\left(\frac{x}{c}+1\right)^{\alpha}\log^2\left(\frac{x}{c}+1\right) < 0.$$

Since $q(0) = 0$, this implies $q(\alpha) < 0$ for all $\alpha > 0$, so $h^{'}(\alpha) > 0$ for all $\alpha > 0$. Hence the result. □

Theorem 3.3.5. *Suppose $b_i = d_i = c$ for all i, $\sum_{i=1}^{n} a_i > 1$ and $\sum_{i=1}^{n} c_i > 1$. Then $X_{1:n} \leq_{icx} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (3.5) and (3.7),

$$\int_t^{\infty} \overline{F}_{1:n}(x)dx = \frac{c}{\sum_{i=1}^{n} a_i - 1}\left(1+\frac{t}{c}\right)^{1-\sum_{i=1}^{n} a_i}$$

and

$$\int_t^{\infty} \overline{G}_{1:n}(x)dx = \frac{c}{\sum_{i=1}^{n} c_i - 1}\left(1+\frac{t}{c}\right)^{1-\sum_{i=1}^{n} c_i}.$$

The ordering holds if and only if

$$\left(1+\frac{t}{c}\right)^{\sum_{i=1}^{n} c_i - \sum_{i=1}^{n} a_i} \leq \frac{\sum_{i=1}^{n} a_i - 1}{\sum_{i=1}^{n} c_i - 1}$$

for all $t > 0$. This holds if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$. □

Theorem 3.3.6. *Suppose $b_i = d_i = c$ for all i, $\sum_{i=1}^{n} a_i > 1$ and $\sum_{i=1}^{n} c_i > 1$. Then $X_{1:n} \leq_{mrl} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (3.5) to (3.8),

$$\frac{\int_t^\infty x f_{1:n}(x)dx}{\overline{F}_{1:n}(t)} = \frac{t\sum_{i=1}^n a_i + c}{\sum_{i=1}^n a_i - 1}$$

and

$$\frac{\int_t^\infty x g_{1:n}(x)dx}{\overline{G}_{1:n}(t)} = \frac{t\sum_{i=1}^n c_i + c}{\sum_{i=1}^n c_i - 1}.$$

The ordering holds if and only if

$$c\left(\sum_{i=1}^n c_i - \sum_{i=1}^n a_i\right) \le t\left(\sum_{i=1}^n a_i - \sum_{i=1}^n c_i\right)$$

for all $t > 0$. This holds if and only if $\sum_{i=1}^n a_i \ge \sum_{i=1}^n c_i$. □

Theorem 3.3.7. *Suppose $a_i = c_i = c$ for all i. Then $X_{1:n} \le_{st} Y_{1:n}$ if $b_i \le d_i$ for all i.*

Proof: By (3.5) and (3.7),

$$\frac{\overline{F}_{1:n}(x)}{\overline{G}_{1:n}(x)} = \prod_{i=1}^n \left(\frac{b_i}{x+b_i}\frac{x+d_i}{d_i}\right)^c.$$

The result follows. □

Theorem 3.3.8. *Suppose $a_i = c_i = c$ for all i. Then $X_{1:n} \le_{hr} Y_{1:n}$ if $b_i \le d_i$ for all i.*

Proof: By (3.5) to (3.8),

$$\frac{f_{1:n}(x)/\overline{F}_{1:n}(x)}{g_{1:n}(x)/\overline{G}_{1:n}(x)} = \frac{\sum_{i=1}^n \frac{1}{x+b_i}}{\sum_{i=1}^n \frac{1}{x+d_i}}.$$

The result follows. □

We can observe the following from Theorems 3.3.1 to 3.3.8: i) if $b_i = d_i$ for all i then the stochastic, hazard rate, likelihood ratio and reversed hazard rate orderings are equivalent; ii) if

$b_i = d_i = c$ for all i, $\sum_{i=1}^{n} a_i > 1$ and $\sum_{i=1}^{n} c_i > 1$ then the stochastic, hazard rate, likelihood ratio, reversed hazard rate, mean residual life and increasing convex orderings are equivalent; iii) if $a_i = c_i$ for all i then the stochastic and hazard rate orderings are equivalent.

3.4 Results for the smallest of Pareto type IV random variables

Let $X_i \sim PIV(\alpha, \gamma, \sigma)$, $i = 1, 2, \ldots, n$ be independent Pareto type IV random variables. Let $Y_i \sim PIV(a, b, c)$, $i = 1, 2, \ldots, n$ be independent Pareto type IV random variables. Let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$ and $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. Note that

$$\overline{F}_{1:n}(x) = \left[1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-n\alpha}, \tag{3.9}$$

$$f_{1:n}(x) = \frac{n\alpha}{\gamma\sigma^{\frac{1}{\gamma}}} \left[1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-n\alpha-1} x^{\frac{1}{\gamma}-1}, \tag{3.10}$$

$$\overline{G}_{1:n}(x) = \left[1 + \left(\frac{x}{c}\right)^{\frac{1}{b}}\right]^{-na} \tag{3.11}$$

and

$$g_{1:n}(x) = \frac{na}{bc^{\frac{1}{b}}} \left[1 + \left(\frac{x}{c}\right)^{\frac{1}{b}}\right]^{-na-1} x^{\frac{1}{b}-1} \tag{3.12}$$

for $x > 0$, $a > 0$, $b > 0$, $c > 0$, $\alpha > 0$, $\gamma > 0$ and $\sigma > 0$.

Theorem 3.4.1 gives an if and only if condition for stochastic ordering when $a = \alpha$. Theorem 3.4.2 gives an if and only if condition for hazard rate ordering when $a = \alpha$ and $b = \gamma$. Theorem 3.4.3 gives an if and only if condition for likelihood ratio ordering when $a = \alpha$ and $b = \gamma$.

Theorem 3.4.1. *Suppose $a = \alpha$. Then $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $\gamma = b$ and $c \geq \sigma$.*

Proof: By (3.9) and (3.11),

$$\overline{F}_{1:n}(x) \leq \overline{G}_{1:n}(x)$$

for all $x > 0$ if and only if

$$\left[1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-na} \leq \left[1 + \left(\frac{x}{c}\right)^{\frac{1}{b}}\right]^{-na}$$

for all $x > 0$ if and only if

$$\left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}} \geq \left(\frac{x}{c}\right)^{\frac{1}{b}}$$

for all $x > 0$ if and only if $\gamma = b$ and $c \geq \sigma$. $\square$

Theorem 3.4.2. *Suppose $a = \alpha$ and $b = \gamma$. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $c \geq \sigma$.*

Proof: By (3.9) to (3.12),

$$\frac{f_{1:n}(x)}{\overline{F}_{1:n}(x)} \geq \frac{g_{1:n}(x)}{\overline{G}_{1:n}(x)}$$

for all $x > 0$ if and only if

$$\frac{1}{\sigma^{\frac{1}{\gamma}}}\left[1 + \left(\frac{x}{\sigma}\right)^{\frac{1}{\gamma}}\right]^{-1} \geq \frac{1}{c^{\frac{1}{\gamma}}}\left[1 + \left(\frac{x}{c}\right)^{\frac{1}{\gamma}}\right]^{-1}$$

for all $x > 0$ if and only if

$$\sigma^{\frac{1}{\gamma}} + x^{\frac{1}{\gamma}} \leq c^{\frac{1}{\gamma}} + x^{\frac{1}{\gamma}}$$

for all $x > 0$ if and only if $c \geq \sigma$. $\square$

Theorem 3.4.3. *Suppose $a = \alpha$ and $b = \gamma$. Then $X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $c \geq \sigma$.*

Proof: By (3.14) and (3.16), we can write

$$\frac{g_{1:n}(x)}{f_{1:n}(x)} = \left(\frac{\sigma}{c}\right)^{-n\alpha/\gamma}\left(1 + \frac{c^{1/\gamma} - \sigma^{1/\gamma}}{\sigma^{1/\gamma} + x^{1/\gamma}}\right)^{-n\alpha-1}.$$

This is an increasing function of x for all $x > 0$ if and only if $c \geq \sigma$. $\square$

3.5 Real data application

Here, we illustrate the results of Sections 3.2 and 3.3 using a real data set. The data is on automobile insurance claims from a large midwestern (US) property and casualty insurer for private passenger automobile insurance. The data has several variables, but the ones we use here are: i) the amount paid to settle and close a claim in United States dollars; ii) gender (male or female); iii) state (coded as a number between 1 and 12). More details of the data can be found in the R contributed package insuranceData (R Development Core Team, 2016).

The insurer would like the claim to be as small as possible. We ask the questions: is the minimum claim smaller for female operators than for male operators? Is it the other way around? To answer these questions, we take $X_1, X_2, \ldots, X_{12}$ as denoting the claims for female operators for the twelve states. We take $Y_1, Y_2, \ldots, Y_{12}$ as denoting the claims for male operators for the twelve states. The aim is to see how $X_{1:12}$ and $Y_{1:12}$ are ordered.

It is reasonable to assume that $X_1, X_2, \ldots, X_{12}$ are independent random variables since they correspond to different states. Likewise for $Y_1, Y_2, \ldots, Y_{12}$. We suppose X_i has the Pareto type I distribution with parameters (a_i, b_i) and Y_i has the Pareto type I distribution with parameters (c_i, d_i) for $i = 1, 2, \ldots, 12$. The maximum likelihood estimates of these parameters are shown in Table 3.1. Also given in this table are p-values based on the Kolmogorov test for goodness of fit of the Pareto type I distribution. These p-values show that the Pareto type I distribution provides reasonable fit. However, because the parameters are estimated instead of taking true values, the Kolmogorov test may be generating a conservative p-value. To address this issue, we can use Fisher's method. It assumes that $-2\sum_{i=1}^{k} \ln(p_i) \sim \chi^2_{2k}$, where p_i is the p-value for the i^{th} hypothesis test. Here, in our application, $k = 24$, $-2\sum_{i=1}^{k} \ln(p_i) = 80.218$. The statistic is only slightly smaller than the critical value 81.075 ($= \chi^2_{48}$), which indicates that we fail to reject all of the null hypotheses. That is, the Pareto type I distribution provides resonable fit. An alternative method for goodness-of-fit can be the Stephen's half-sample method (Stephens et. al., 1978). It mainly estimates the unkown parameters using half of the sample and test the goodness-of-fit based on the difference between the estimated cumulative distribution function and the empirical distribution function. As further evidence, probability plots of the fits are shown in Figures 3.1 and 3.2. These results are expected since Pareto distributions are popular models for insurance claims.

Given the parameter estimates in Table 3.1, we have $\widehat{e} = \max\left(\widehat{b_1}, \widehat{b_2}, \ldots, \widehat{b_{12}}, \widehat{d_1}, \widehat{d_2}, \ldots, \widehat{d_{12}}\right) = 129.18$, $\widehat{a} = \widehat{a_1} + \widehat{a_2} + \cdots + \widehat{a_{12}} = 3.890338$, $\widehat{c} = \widehat{c_1} + \widehat{c_2} + \cdots + \widehat{c_{12}} = 3.746099$, $\widehat{\alpha} = \prod_{i=1}^{n} \widehat{b_i}^{\widehat{a_i}} = 2950607$, $\widehat{\beta} = \prod_{i=1}^{n} \widehat{d_i}^{\widehat{c_i}} = 1560400$, $\widehat{x_0} = \left(\widehat{\alpha}/\widehat{\beta}\right)^{1/(\widehat{a}-\widehat{c})} = 82.82813$, $\widehat{y_0} = \left\{\widehat{\beta}\,(\widehat{a}-1) / \left[\widehat{\alpha}\,(\widehat{c}-1)\right]\right\}^{1/(\widehat{c}-\widehat{a})} = 58.08202$ and $\widehat{h(\widehat{e})} = -2.548012 \times 10^{13}$. The standard deviations of the maximum likelihood estimates are presented in Table 3.1. By the definition of the Pareto distribution of type I, the scale parameter (i.e. b, d) should be smaller than or equal to the minimum of the fitted data. For the purpose of obtaining the maximum likelihood estimation, we let the scale parameter equal to the minimumof the corresponding dataset directly. We use the Bootstrap method to approximate the standard deviation of the estimated scale parameter (i.e. $\widehat{b}$ or $\widehat{d}$) of each dataset. However, the Bootstrap performs poorly when the sample size is small. The standard errors of $\widehat{e}$, $\widehat{a}$, $\widehat{c}$, $\widehat{x_0}$, $\widehat{y_0}$ and $\widehat{h(\widehat{e})}$ were computed as 11.9, 0.122, 0.042, 5.65, 5.43 and 1.276×10^6, respectively. The corresponding confidence intervals are $(117.28, 141.08)$, $(3.769, 4.012)$, $(3.704, 3.788)$, $(77.178, 88.478)$, $(52.652, 63.512)$, $(-2.5480121 \times 10^{13}, -2.5480118 \times 10^{13})$.

After accounting for the standard errors and calculating the confidence interval for each of the estimate, conclusions can be drawn as $\widehat{c} < \widehat{a}$ and $\widehat{x_0} < \widehat{e}$. So, Theorem 3.2.1 shows that the minimum claim for female operators is stochastically less than the minimum claim for male operators. Since $\widehat{c} < \widehat{a}$, Theorem 3.2.2 shows that the minimum claim for female operators is less than the minimum claim for male operators with respect to hazard rate order. Since $\widehat{c} < \widehat{a}$, Theorem 3.2.3 shows that the minimum claim for female operators is less than the minimum claim for male operators with respect to likelihood ratio order. After accounting for the standard errors and calculating the confidence interval for each of the estimate, conclusions can be drawn as $\widehat{c} < \widehat{a}$ and $\widehat{h(\widehat{e})} < 0$. So, Theorem 3.2.4 shows that the minimum claim for female operators is less than the minimum claim for male operators with respect to reversed hazard rate order. After accounting for the standard errors and calculating the confidence interval for each of the estimate, conclusions can be drawn as $\widehat{a} > 1$, $\widehat{c} > 1$, $\widehat{c} < \widehat{a}$ and $\widehat{y_0} < \widehat{e}$. So, Theorem 3.2.5 shows that the minimum claim for female operators is less than the minimum claim for male operators with respect to increasing convex order. Since $\widehat{a} > 1$, $\widehat{c} > 1$ and $\widehat{c} < \widehat{a}$, Theorem 3.2.6 shows that the minimum claim for female operators is less than the minimum claim for male operators with respect to mean residual life order.

The established orderings between $X_{1:12}$ and $Y_{1:12}$ are confirmed by the plots of the ratio of survival functions, the ratio of probability density functions, the ratio of hazard rate functions, the ratio of reversed hazard rate functions, the ratio of cumulative survival functions and the ratio of mean residual life functions shown in Figure 3.3.

We also fitted the Pareto type II distribution to the data on $X_1, X_2, \ldots, X_{12}$ and $Y_1, Y_2, \ldots, Y_{12}$. The results in Section 3.3 are however limited to the cases $b_i = d_i =$ constant for all i or $a_i = c_i =$ constant for all i. A likelihood ratio test of $b_i = d_i =$ constant for all i or $a_i = c_i =$ constant for all i based on the fitted estimates was not accepted.

Gender	State	Sample size	Parameter estimates(s.d.)	p-value for Kolmogorov test
F	1	51	$\widehat{a_1} = 0.2897983$ (0.041), $\widehat{b_1} = 25$ (8.58)	0.143
F	2	435	$\widehat{a_2} = 0.2161304$ (0.010), $\widehat{b_2} = 10$ (8.81)	0.250
F	3	134	$\widehat{a_3} = 0.4179904$ (0.036), $\widehat{b_3} = 91$ (8.32)	0.207
F	4	235	$\widehat{a_4} = 0.2673118$ (0.017), $\widehat{b_4} = 21$ (1.49)	0.284
F	5	239	$\widehat{a_5} = 0.3327483$ (0.021), $\widehat{b_5} = 61$ (9.90)	0.235
F	6	115	$\widehat{a_6} = 0.3266533$ (0.030), $\widehat{b_6} = 50$ (9.75)	0.130
F	7	117	$\widehat{a_7} = 0.4726428$ (0.044), $\widehat{b_7} = 125$ (9.60)	0.069
F	8	94	$\widehat{a_8} = 0.2464126$ (0.025), $\widehat{b_8} = 25$ (21.02)	0.183
F	9	76	$\widehat{a_9} = 0.3926721$ (0.045), $\widehat{b_9} = 73$ (10.96)	0.282
F	10	46	$\widehat{a_{10}} = 0.4126978$ (0.061), $\widehat{b_{10}} = 75$ (20.17)	0.112
F	11	820	$\widehat{a_{11}} = 0.2487056$ (0.009), $\widehat{b_{11}} = 18$ (2.29)	0.063
F	12	215	$\widehat{a_{12}} = 0.2665744$ (0.018), $\widehat{b_{12}} = 25$ (2.40)	0.143
M	1	115	$\widehat{c_1} = 0.2765855$ (0.026), $\widehat{d_1} = 26$ (14.46)	0.243
M	2	687	$\widehat{c_2} = 0.2949511$ (0.011), $\widehat{d_2} = 35$ (0.80)	0.390
M	3	214	$\widehat{c_3} = 0.3333928$ (0.023), $\widehat{d_3} = 45$ (4.71)	0.192
M	4	431	$\widehat{c_4} = 0.2891588$ (0.014), $\widehat{d_4} = 32$ (2.51)	0.272
M	5	383	$\widehat{c_5} = 0.3371954$ (0.017), $\widehat{d_5} = 65$ (4.06)	0.284
M	6	154	$\widehat{c_6} = 0.2761947$ (0.022), $\widehat{d_6} = 26$ (8.19)	0.184
M	7	159	$\widehat{c_7} = 0.3135779$ (0.025), $\widehat{d_7} = 50$ (20.3)	0.251
M	8	153	$\widehat{c_8} = 0.4271482$ (0.035), $\widehat{d_8} = 129$ (11.86)	0.182
M	9	132	$\widehat{c_9} = 0.3177517$ (0.028), $\widehat{d_9} = 55$ (1.61)	0.286
M	10	123	$\widehat{c_{10}} = 0.3510437$ (0.032), $\widehat{d_{10}} = 60$ (24.61)	0.114
M	11	1360	$\widehat{c_{11}} = 0.2145939$ (0.006), $\widehat{d_{11}} = 10$ (1.97)	0.267
M	12	276	$\widehat{c_{12}} = 0.3145054$ (0.019), $\widehat{d_{12}} = 50$ (5.18)	0.160

Table 3.1: Sample sizes, parameter estimates(standard deviation), and p-values based on the Kolmogorov test.

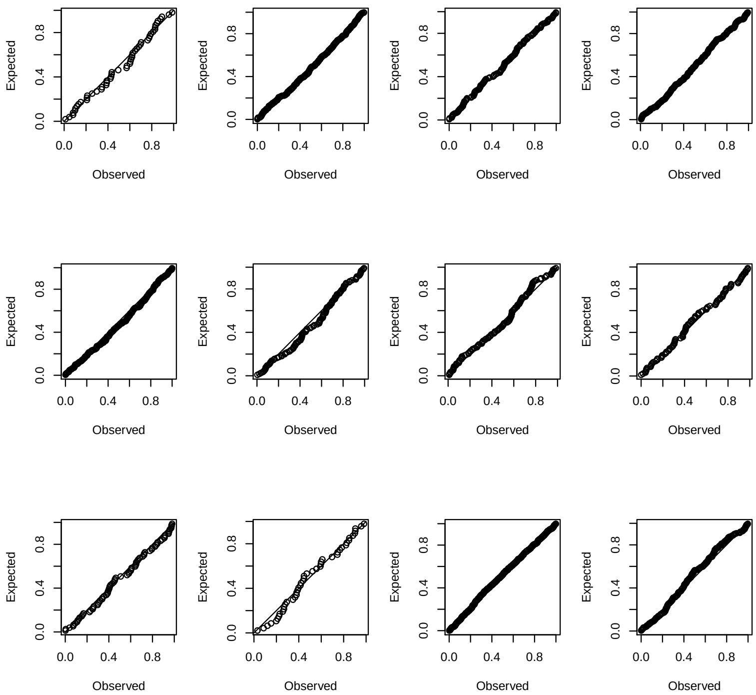

Figure 3.1: Probability plots of the fits of the Pareto type I distribution for claims made by female operators from the 12 states.

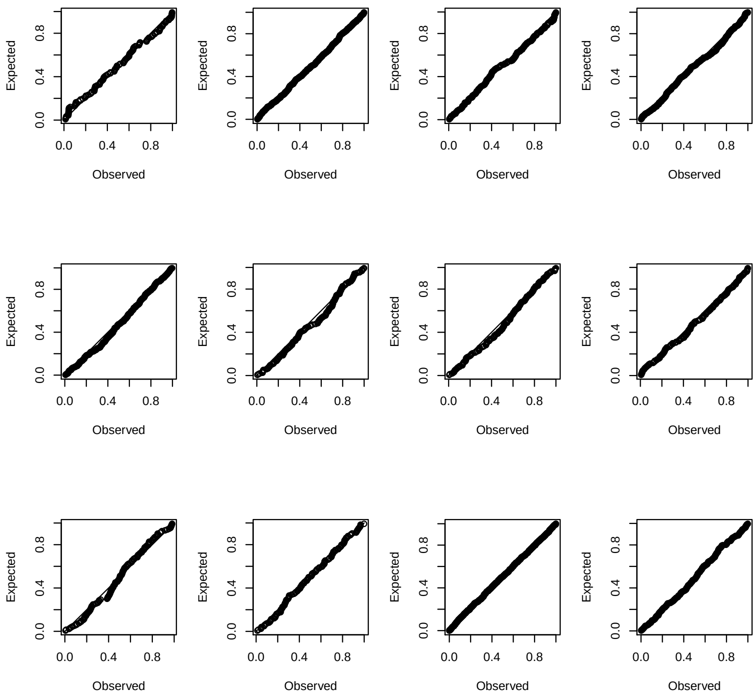

Figure 3.2: Probability plots of the fits of the Pareto type I distribution for claims made by male operators from the 12 states.

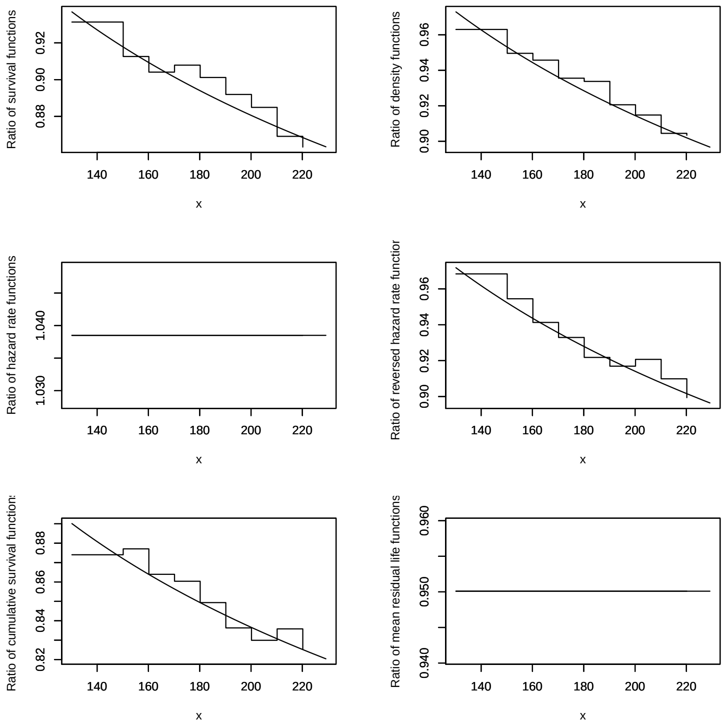

Figure 3.3: Plots of $\widehat{\overline{F}_{1:12}}(x)/\widehat{\overline{G}_{1:12}}(x)$ (top left), $\widehat{f_{1:12}}(x)/\widehat{g_{1:12}}(x)$ (top right), $\left[\widehat{f_{1:12}}(x)/\widehat{\overline{F}_{1:12}}(x)\right] / \left[\widehat{g_{1:12}}(x)/\widehat{\overline{F}_{1:12}}(x)\right]$ (middle left), $\left\{\widehat{f_{1:12}}(x)/\left[1-\widehat{\overline{F}_{1:12}}(x)\right]\right\} / \left\{\widehat{g_{1:12}}(x)/\left[1-\widehat{\overline{G}_{1:12}}(x)\right]\right\}$ (middle right), $\int_t^\infty \overline{F}_{1:12}(x)dx / \int_t^\infty \overline{G}_{1:12}(x)dx$ (bottom left) and $\left[\int_t^\infty x f_{1:12}(x)dx/\overline{F}_{1:12}(t)\right] / \left[\int_t^\infty x g_{1:12}(x)dx/\overline{G}_{1:12}(t)\right]$ (bottom right). Also shown are the empirical versions of these ratios.

3.6 Discussion and future work

Sections 3.2 to 3.4 have only considered orderings between minimums of Pareto random variables. One may be also interested in the orderings between maximums of Pareto random variables. We now give results on orderings between $X_{n:n}$ and $Y_{n:n}$ when $X_1, X_2, \ldots, X_n$ are independent $PI(a, b)$ random variables and $Y_1, Y_2, \ldots, Y_n$ are independent $PI(c, d)$ random variables. Note that

$$F_{n:n}(x) = \left[1 - \left(\frac{x}{b}\right)^{-a}\right]^n, \tag{3.13}$$

$$f_{n:n}(x) = \left[1 - \left(\frac{x}{b}\right)^{-a}\right]^{n-1} \frac{nax^{-a-1}}{b^{-a}}, \tag{3.14}$$

$$G_{n:n}(y) = \left[1 - \left(\frac{y}{d}\right)^{-c}\right]^n \tag{3.15}$$

and

$$g_{n:n}(y) = \left[1 - \left(\frac{y}{d}\right)^{-c}\right]^{n-1} \frac{ncy^{-c-1}}{d^{-c}} \tag{3.16}$$

for $x > b > 0$, $y > d > 0$, $a > 0$ and $c > 0$. Let $e = \max(b, d)$.

Theorem 3.6.1 gives an if and only if condition for stochastic ordering of $X_{n:n}$ and $Y_{n:n}$. Theorem 3.6.2 gives a condition for reversed hazard rate ordering. Theorem 3.6.3 gives a condition for likelihood ratio ordering.

Theorem 3.6.1. *Suppose $d \geq b$, $X_{n:n} \leq_{st} Y_{n:n}$ if and only if $e^{c-a} \leq d^c b^{-a}$ and $c \leq a$.*

Proof: By (3.13) and (3.15),

$$\overline{F}_{n:n}(x) \leq \overline{G}_{n:n}(x)$$

for all $x \geq e$ if and only if

$$1 - \left[1 - \left(\frac{x}{b}\right)^{-a}\right]^n \leq 1 - \left[1 - \left(\frac{x}{d}\right)^{-c}\right]^n$$

for all $x \geq e$ if and only if

$$\left(\frac{x}{b}\right)^{-a} \leq \left(\frac{x}{d}\right)^{-c}$$

for all $x \geq e$ if and only if

$$x^{c-a} \leq b^{-a} d^{c}$$

for all $x \geq e$. Hence, the result. □

Theorem 3.6.2. *Suppose* $b \geq d$, $X_{n:n} \leq_{rh} Y_{n:n}$ *if* $c \leq a$, $(c-a)(e/d)^c + a \leq 0$ *and* $(e^c d^{-c} - 1) / (e^a b^{-a} - 1) \leq c/a$.

Proof: By (3.13) to (3.16),

$$\frac{f_{n:n}(x)}{F_{n:n}(x)} \leq \frac{g_{n:n}(x)}{G_{n:n}(x)}$$

for all $x \geq e$ if and only if

$$\frac{nax^{-a-1}}{b^{-a} - x^{-a}} \leq \frac{ncx^{-c-1}}{d^{-c} - x^{-c}}$$

for all $x \geq e$ if and only if

$$\frac{x^c d^{-c} - 1}{x^a b^{-a} - 1} \leq \frac{c}{a}$$

for all $x \geq e$. Let

$$\omega(x) = \frac{x^c d^{-c} - 1}{x^a b^{-a} - 1}.$$

Its first derivative can be written as

$$\omega^{'}(x) = \frac{b^{-a} x^{a-1}\left[(c-a)d^{-c}x^{c} + a\right] - cd^{-c}x^{c-1}}{\left(x^a b^{-a} - 1\right)^2}.$$

Note that $\omega^{'}(x) \leq 0$ for all $x \geq e$ if $(c-a)d^{-c}x^c + a \leq 0$ for all $x \geq e$. The latter holds if $c \leq a$ and $(c-a)(e/d)^c + a \leq 0$. Hence, the result. □

Theorem 3.6.3. $X_{n:n} \leq_{lr} Y_{n:n}$ *if* $c < a$ *and* $ad^{-c}e^{c-a} < cb^{-a}$.

Proof: By (3.14) and (3.16),

$$\frac{f_{n:n}(x)}{g_{n:n}(x)} = \frac{ab^{na}}{cd^{nc}} x^{c-a} \left(\frac{b^{-a} - x^{-a}}{d^{-c} - x^{-c}} \right)^{n-1}.$$

Let

$$\omega(x) = \frac{b^{-a} - x^{-a}}{d^{-c} - x^{-c}}.$$

Its first derivative can be written as

$$\omega^{'}(x) = \frac{(c-a)x^{-a-c-1} + x^{-c-1}\left(ad^{-c}x^{c-a} - cb^{-a}\right)}{\left(d^{-c} - x^{-c}\right)^2}.$$

Note that $\omega^{'}(x) < 0$ for all $x \geq e$ if $c < a$ and $ad^{-c}x^{c-a} < cb^{-a}$ for all $x \geq e$. Hence, the result. □

The Pareto density function is non-log-concave, however its cumulative distribution function is concave. An ordering related to log concavity is the proportional likelihood ordering due to Ramos Romero and Sordo Diaz (2001). We write $X_{1:n} \leq_{plr} Y_{1:n}$ if $g_{1:n}(\lambda x)/f_{1:n}(x)$ is an increasing function of x for all possible x and $0 < \lambda < 1$.

Suppose $X_i \sim PI\left(a_i, b_i\right)$, $i = 1, 2, \ldots, n$ are independent Pareto type I random variables and $Y_i \sim PI\left(c_i, d_i\right)$, $i = 1, 2, \ldots, n$ are also independent Pareto type I random variables. Theorem 3.6.4 gives an if and only if condition for proportional likelihood ratio ordering of $X_{1:n}$ and $Y_{1:n}$.

Theorem 3.6.4. *$X_{1:n} \leq_{plr} Y_{1:n}$ if and only if $\sum_{i=1}^{n} a_i > \sum_{i=1}^{n} c_i$.*

Proof: By (3.2) and (3.4),

$$\frac{f_{1:n}(x)}{g_{1:n}(\lambda x)} = x^{\sum_{i=1}^{n} c_i - \sum_{i=1}^{n} a_i} \frac{\lambda^{1+\sum_{i=1}^{n} c_i} \left(\sum_{i=1}^{n} a_i\right) \prod_{i=1}^{n} b_i^{a_i}}{\left(\sum_{i=1}^{n} c_i\right) \prod_{i=1}^{n} d_i^{c_i}}.$$

The result follows. □

Suppose now $X_i \sim PII\left(a_i, b_i\right)$, $i = 1, 2, \ldots, n$ are independent Pareto type II random variables and $Y_i \sim PII\left(c_i, d_i\right)$, $i = 1, 2, \ldots, n$ are also independent Pareto type II random variables. Theorem 3.6.5 gives a condition for proportional likelihood ratio ordering of $X_{1:n}$ and $Y_{1:n}$.

Theorem 3.6.5. *Suppose $b_i = d_i = c$ for all i. Then $X_{1:n} \leq_{plr} Y_{1:n}$ if $\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{n} c_i$.*

Proof: By (4.2) and (4.4), we can write

$$\frac{g_{1:n}(\lambda x)}{f_{1:n}(x)} = \frac{c^{\sum_{i=1}^n c_i - \sum_{i=1}^n a_i} \left(\sum_{i=1}^n c_i \right)}{\sum_{i=1}^n a_i} \omega(x),$$

where

$$\omega(x) = (\lambda x + c)^{-1 - \sum_{i=1}^n c_i} (x + c)^{1 + \sum_{i=1}^n a_i}.$$

The derivative of $\log \omega(x)$ is

$$\frac{d \log \omega(x)}{dx} = -\frac{\lambda}{\lambda x + c} \left(1 + \sum_{i=1}^n c_i \right) + \frac{1}{x + c} \left(1 + \sum_{i=1}^n a_i \right).$$

This derivative is positive if and only if

$$\left(\sum_{i=1}^n a_i - \sum_{i=1}^n c_i \right) \lambda x > c(\lambda - 1) + c \left(\lambda \sum_{i=1}^n c_i - \sum_{i=1}^n a_i \right),$$

which holds since $\lambda < 1$ and $\sum_{i=1}^n a_i \geq \sum_{i=1}^n c_i$. $\square$

An interesting study is orderings between $X_{1:n}$ and $Y_{1:n}$ when X_i and Y_i are Feller Pareto random variables (Johnson et al., 1994, equation (20.4)) with

$$f(x) = \frac{a x^{a-1}}{b^a B(p, q)} \left[1 + \left(\frac{x}{b} \right)^a \right]^{-q-1} \left\{ 1 - \left[1 + \left(\frac{x}{b} \right)^a \right]^{-1} \right\}^{p-1},$$

$$F(x) = I_{1 - \left[1 + \left(\frac{x}{b} \right)^a \right]^{-1}} (p, q)$$

and

$$\overline{F}(x) = I_{\left[1 + \left(\frac{x}{b} \right)^a \right]^{-1}} (q, p)$$

for $x > 0$, $a > 0$, $b > 0$, $p > 0$ and $q > 0$, where $B(p, q)$ and $I_x(p, q)$ denote the beta function and incomplete beta function ratio defined by

$$B(p, q) = \int_0^1 t^{p-1} (1 - t)^{q-1} dt$$

and

$$I_x(p,q) = \frac{1}{B(p,q)} \int_0^x t^{p-1}(1-t)^{q-1} dt,$$

respectively, for $0 < x < 1$, $p > 0$ and $q > 0$.

The study will require finding conditions like the following holding

$$\left\{ I_{\left[1+\left(\frac{x}{b}\right)^a\right]^{-1}}(q,p) \right\}^n \leq \left\{ I_{\left[1+\left(\frac{x}{d}\right)^c\right]^{-1}}(s,r) \right\}^n$$

for all $x > 0$. Hence, properties of the incomplete beta function ratio will be needed for the study.

However, Feller Pareto random variables can be represented as the ratio of two independent gamma random variables. It may be easier to study orderings of Feller Pareto random variables using this fact. We suggest this as a possible future work.

Another interesting study is orderings between $X_{1:n}$ and $Y_{1:n}$ when X_i and Y_i are discrete Pareto random variables. There are several discrete Pareto distributions. One of the simplest discrete Pareto distribution is the zeta distribution with

$$f(x) = \left(\sum_{m=1}^{\infty} m^{-a} \right)^{-1} x^{-a},$$

$$F(x) = \left(\sum_{m=1}^{\infty} m^{-a} \right)^{-1} \left(\sum_{m=1}^{x} m^{-a} \right)$$

and

$$\overline{F}(x) = \left(\sum_{m=1}^{\infty} m^{-a} \right)^{-1} \left(\sum_{m=x+1}^{\infty} m^{-a} \right)$$

for $x = 1, 2, \ldots$ and $a > 1$. The study will require finding conditions like the following holding

$$\left(\sum_{m=1}^{\infty} m^{-a} \right)^{-n} \left(\sum_{m=x+1}^{\infty} m^{-a} \right)^{n} \leq \left(\sum_{m=1}^{\infty} m^{-b} \right)^{-n} \left(\sum_{m=x+1}^{\infty} m^{-b} \right)^{n}$$

for all $x = 1, 2, \ldots$. We suggest this as another possible future work.

Chapter 4

Smallest Weibull order statistics

4.1 Introduction

Let $X_1, X_2, \ldots, X_n$ be independent but non-identical random variables from one population and let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$. Let $Y_1, Y_2, \ldots, Y_n$ be independent but non-identical random variables from another population and let $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. There has been little work on the stochastic comparison of $X_{1:n}$ and $Y_{1:n}$. The work that we are aware of are: Khaledi and Kochar (2006), Fang and Tang (2014), Li and Li (2015) and Torrado (2015). Balakrishnan and Zhao (2013) provide an excellent review of other known work.

The aim of this chapter is to study orderings between $X_{1:n}$ and $Y_{1:n}$ when $X_1, X_2, \ldots, X_n$ are independent but non-identical Weibull random variables from one population and $Y_1, Y_2, \ldots, Y_n$ are independent but non-identical Weibull random variables from another population. The results here are different from and are more general than those in Torrado (2015).

Weibull distributions are the most popular models for lifetime data (Murthy et al., 2003). Often comparison of the smallest order statistics has practical appeal in lifetime modeling. Suppose $X_1, X_2, \ldots, X_n$ are survival times of patients from a population with one characteristic (e.g., females) and $Y_1, Y_2, \ldots, Y_n$ are survival times of patients from a population with another characteristic (e.g., males). Then $X_{1:n}$ and $Y_{1:n}$ are the smallest survival times and it makes sense to compare them to see which characteristic is associated with smaller survival times. We shall return to this example in Section 4.5.

We consider two Weibull distributions: the standard Weibull distribution and a lower truncated

Weibull distribution.

A Weibull random variable with shape parameter α and scale parameter λ denoted by $W(\alpha,\lambda)$ has the probability density, cumulative distribution and hazard rate functions specified by

$$f(x) = \alpha\lambda^{\alpha}x^{\alpha-1}e^{-(\lambda x)^{\alpha}},$$

$$F(x) = 1 - e^{-(\lambda x)^{\alpha}},$$

and

$$h(x) = \alpha\lambda(\lambda x)^{\alpha-1},$$

respectively, for $x > 0$, $\alpha > 0$ and $\lambda > 0$.

A lower truncated Weibull random variable with shape parameter α and scale parameter λ denoted by $LTW(\alpha,\lambda)$ has the probability density, cumulative distribution and hazard rate functions specified by

$$f(x) = \alpha\lambda^{\alpha}x^{\alpha-1}e^{1-(\lambda x)^{\alpha}},$$

$$F(x) = 1 - e^{1-(\lambda x)^{\alpha}},$$

and

$$h(x) = \alpha\lambda(\lambda x)^{\alpha-1},$$

respectively, for $x > 1/\lambda$, $\alpha > 0$ and $\lambda > 0$.

Let $\overline{F}_i$ and f_i denote the survival and probability density functions of X_i. Let $\overline{G}_i$ and g_i denote the survival and probability density functions of Y_i. Then the survival and probability density functions of $X_{1:n}$ are

$$\overline{F}_{1:n}(x) = \prod_{i=1}^{n} \overline{F}_i(x)$$

and

$$f_{1:n}(x) = \left[\prod_{i=1}^{n} \overline{F}_i(x)\right] \left[\sum_{i=1}^{n} \frac{f_i(x)}{\overline{F}_i(x)}\right],$$

respectively. Similarly, the survival and probability density functions of $Y_{1:n}$ are

$$\overline{G}_{1:n}(x) = \prod_{i=1}^{n} \overline{G}_i(x)$$

and

$$g_{1:n}(x) = \left[\prod_{i=1}^{n} \overline{G}_i(x)\right] \left[\sum_{i=1}^{n} \frac{g_i(x)}{\overline{G}_i(x)}\right],$$

respectively.

The orderings that we consider are: $X_{1:n}$ is smaller than $Y_{1:n}$ in stochastic order denoted by $X_{1:n} \leq_{st} Y_{1:n}$ if $\overline{F}_{1:n}(x) \leq \overline{G}_{1:n}(x)$ for all x; $X_{1:n}$ is smaller than $Y_{1:n}$ in hazard rate order denoted by $X_{1:n} \leq_{hr} Y_{1:n}$ if $f_{1:n}(x)/\overline{F}_{1:n}(x) \geq g_{1:n}(x)/\overline{G}_{1:n}(x)$ for all x; $X_{1:n}$ is smaller than $Y_{1:n}$ in likelihood ratio order denoted by $X_{1:n} \leq_{lr} Y_{1:n}$ if $g_{1:n}(x)/f_{1:n}(x)$ is increasing in x for all x for which the ratio is well defined. Further details of these orderings can be found in Shaked and Shanthikumar (2007).

In this chapter, we derive if and only if conditions for stochastic, hazard rate and likelihood ratio orderings of $X_{1:n}$ and $Y_{1:n}$. Section 4.3 gives the results when $X_{1:n}$ is a minimum of independent Weibull random variables and $Y_{1:n}$ is a minimum of an another set of independent Weibull random variables with different shape and scale parameters. Section 4.4 gives the results when $X_{1:n}$ is a minimum of independent lower truncated Weibull random variables and $Y_{1:n}$ is a minimum of an another set of independent lower truncated Weibull random variables with different shape and scale parameters. Some technical lemmas needed for the results in Sections 4.3 and 4.4 are derived in Section 4.2. Section 4.5 presents an application to survival times data.

4.2 Technical lemmas

The proofs of the results in Section 4.3 and 4.4 need the following lemmas.

Lemma 4.2.1. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^r + Sx^s}$$

for $x > 0$, $p > 0$, $q > 0$, $r > 0$, $s > 0$, $P > 0$, $Q > 0$, $R > 0$ and $S > 0$ real numbers. Then $g(x) \geq w$ for all x if one of the following conditions holds

(i) $p \geq r$, $q \geq s$, $p \geq s$, $q \geq r$ and $g(0) \geq w$,

(ii) $p < r$, $q < s$, $p < s$, $q < r$ and $g(\infty) \geq w$.

Proof: The first derivative of $g(x)$ can be expressed as

$$g^{'}(x) = \frac{PR(p-r)x^{p+r-1} + QR(q-r)x^{q+r-1} + PS(p-s)x^{p+s-1} + QS(q-s)x^{q+s-1}}{\left(Rx^r + Sx^s\right)^2}.$$

If $p \geq r$, $q \geq s$, $p \geq s$ and $q \geq r$ then $g^{'}(x) \geq 0$ for all $x > 0$, and so $g(x) \geq w$ for all $x > 0$ if and only if $g(0) \geq w$. If $p < r$, $q < s$, $p < s$ and $q < r$ then $g^{'}(x) < 0$ for all $x > 0$, and so $g(x) \geq w$ for all $x > 0$ if and only if $g(\infty) \geq w$. The proof is complete. □

Lemma 4.2.2. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^p + Sx^q}$$

for $x > 0$, $p > 0$, $q > 0$, $P > 0$, $Q > 0$, $R > 0$ and $S > 0$ real numbers. Then $g(x) \geq w$ for all x if and only if one of the following conditions holds

(i) $P \geq QR/S$ and $Q \geq wS$,

(ii) $P < QR/S$ and $P \geq wR$.

Proof: We can rewrite $g(x)$ as

$$g(x) = \frac{Q}{S} + \frac{P - RQ/S}{R + Sx^{q-p}}.$$

If $q \geq p$ and $P \geq QR/S$ then $g(x)$ is a decreasing function and so $g(x) \geq w$ for all $x > 0$ if and only if $g(\infty) = Q/S \geq w$. If $q < p$ and $P \geq QR/S$ then $g(x)$ is an increasing function and so $g(x) \geq w$ for all $x > 0$ if and only if $g(0) = Q/S \geq w$. If $q \geq p$ and $P < QR/S$ then $g(x)$ is an increasing function and so $g(x) \geq w$ for all $x > 0$ if and only if $g(0) = P/R \geq w$. If $q < p$ and $P < QR/S$ then $g(x)$ is a decreasing function and so $g(x) \geq w$ for all $x > 0$ if and only if $g(\infty) = P/R \geq w$. The proof is complete. □

Lemma 4.2.3. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^p + Sx^q}$$

for $x > 0$, $p > 0$, $q > 0$, $P > 0$, $Q > 0$, $R > 0$ and $S > 0$ real numbers. Then $g(x)$ is a decreasing function of x if and only if one of the following conditions holds

(i) $p = q$,

(ii) $p > q$ *and* $P \leq QR/S$,

(iii) $p < q$ *and* $P \geq QR/S$.

Proof: We can rewrite $g(x)$ as

$$g(x) = \frac{Q}{S} + \frac{P - RQ/S}{R + Sx^{q-p}}.$$

If $p = q$ then $g(x)$ is a constant. If $p > q$ then $g(x)$ is a decreasing function if and only if $P \leq QR/S$. If $p < q$ then $g(x)$ is a decreasing function if and only if $P \geq QR/S$. The proof is complete. □

Lemma 4.2.4. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^r + Sx^s}$$

for $x > c > 0$, $p > 0$, $q > 0$, $r > 0$, $s > 0$, $P > 0$, $Q > 0$, $R > 0$ and $S > 0$ real numbers. Then $g(x) \geq w$ for all x if one of the following conditions holds

(i) $p \geq r$, $q \geq s$, $p \geq s$, $q \geq r$ *and* $g(c) \geq w$,

(ii) $p < r$, $q < s$, $p < s$, $q < r$ *and* $g(\infty) \geq w$.

Proof: The first derivative of $g(x)$ can be expressed as

$$g^{'}(x) = \frac{PR(p-r)x^{p+r-1} + QR(q-r)x^{q+r-1} + PS(p-s)x^{p+s-1} + QS(q-s)x^{q+s-1}}{(Rx^r + Sx^s)^2}.$$

If $p \geq r$, $q \geq s$, $p \geq s$ and $q \geq r$ then $g^{'}(x) \geq 0$ for all $x > c$, and so $g(x) \geq w$ for all $x > c$ if and only if $g(c) \geq w$. If $p < r$, $q < s$, $p < s$, $q < r$ then $g^{'}(x) < 0$ for all $x > c$, and so $g(x) \geq w$ for all $x > c$ if and only if $g(\infty) \geq w$. The proof is complete. □

Lemma 4.2.5. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^p + Sx^q}$$

for $x > c > 0$, $p > 0$, $q > 0$, $P > 0$, $Q > 0$, $R > 0$ and $S > 0$ real numbers. Then $g(x) \geq w$ for all x if and only if one of the following conditions holds *(i)* $q \geq p$, $P \geq QR/S$ *and* $Q/S \geq w$,

(ii) $q < p$, $P \geq QR/S$ *and* $g(c) \geq w$,

(iii) $q \geq p$, $P < QR/S$ *and* $g(c) \geq w$,

(iv) $q < p$, $P < QR/S$ *and* $P/R \geq w$.

Proof: We can rewrite $g(x)$ as

$$g(x) = \frac{Q}{S} + \frac{P - RQ/S}{R + Sx^{q-p}}.$$

If $q \geq p$ and $P \geq QR/S$ then $g(x)$ is a decreasing function and so $g(x) \geq w$ for all $x > c$ if and only if $g(\infty) = Q/S \geq w$. If $q < p$ and $P \geq QR/S$ then $g(x)$ is an increasing function and so $g(x) \geq w$ for all $x > c$ if and only if $g(c) \geq w$. If $q \geq p$ and $P < QR/S$ then $g(x)$ is an increasing function and so $g(x) \geq w$ for all $x > c$ if and only if $g(c) \geq w$. If $q < p$ and $P < QR/S$ then $g(x)$ is a decreasing function and so $g(x) \geq w$ for all $x > c$ if and only if $g(\infty) = P/R \geq w$. The proof is complete. □

Lemma 4.2.6. *Let*

$$g(x) = \frac{Px^p + Qx^q}{Rx^p + Sx^q}$$

for $x > c > 0$, $p > 0$, $q > 0$, $P > 0$, $Q > 0$, $R > 0$ *and* $S > 0$ *real numbers. Then* $g(x)$ *is a decreasing function of* x *if and only if one of the following conditions holds*

(i) $p = q$,

(ii) $p > q$ *and* $P \leq QR/S$,

(iii) $p < q$ *and* $P \geq QR/S$.

Proof: We can rewrite $g(x)$ as

$$g(x) = \frac{Q}{S} + \frac{P - RQ/S}{R + Sx^{q-p}}.$$

If $p = q$ then $g(x)$ is a constant. If $p > q$ then $g(x)$ is a decreasing function if and only if $P \leq QR/S$. If $p < q$ then $g(x)$ is a decreasing function if and only if $P \geq QR/S$. The proof is complete. □

4.3 Results for the smallest of Weibull random variables

Let $X_i \sim W(\alpha_i, \lambda_i)$, $i = 1, 2, \ldots, n$ be independent Weibull random variables. Let $Y_i \sim W(\beta_i, \theta_i)$, $i = 1, 2, \ldots, n$ be independent Weibull random variables. Let $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$ and $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. Note that

$$\overline{F}_{1:n}(x) = \exp\left[-\sum_{i=1}^{n} (\lambda_i x)^{\alpha_i}\right], \tag{4.1}$$

$$f_{1:n}(x) = x^{-1}\left[\sum_{i=1}^{n} \alpha_i (\lambda_i x)^{\alpha_i}\right] \exp\left[-\sum_{i=1}^{n} (\lambda_i x)^{\alpha_i}\right], \tag{4.2}$$

$$\overline{G}_{1:n}(x) = \exp\left[-\sum_{i=1}^{n} (\theta_i x)^{\beta_i}\right] \tag{4.3}$$

and

$$g_{1:n}(x) = x^{-1}\left[\sum_{i=1}^{n} \beta_i (\theta_i x)^{\beta_i}\right] \exp\left[-\sum_{i=1}^{n} (\theta_i x)^{\beta_i}\right] \tag{4.4}$$

for $x > 0$, $\alpha_i > 0$, $\beta_i > 0$, $\lambda_i > 0$ and $\theta_i > 0$.

Theorems 4.3.1 and 4.3.4 give if and only if conditions for the stochastic ordering. Theorems 4.3.2 and 4.3.7 give if and only if conditions for the hazard rate ordering. Theorem 4.3.3 gives an if and only if condition for the likelihood ratio ordering. Theorem 4.3.5 gives a condition for the stochastic ordering. Theorem 4.3.6 gives a condition for the likelihood ratio ordering. Theorem 4.3.8 gives a condition for the hazard rate ordering. Throughout this section, we let $A = \sum_{i=1}^{n} \lambda_i^{\alpha}$, $B = \sum_{i=1}^{n} \theta_i^{\alpha}$, $A_1 = \sum_{i=1}^{p} \lambda_i^{\alpha_1}$, $A_2 = \sum_{i=p+1}^{n} \lambda_i^{\alpha_2}$, $B_1 = \sum_{i=1}^{p} \theta_i^{\beta_1}$ and $B_2 = \sum_{i=p+1}^{n} \theta_i^{\beta_2}$.

Theorem 4.3.1. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $A \geq B$.*

Proof: By (4.1) and (4.3), $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $e^{-Ax^{\alpha}} \leq e^{-Bx^{\alpha}}$ for all x. Hence the result. □

Theorem 4.3.2. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $A \geq B$.*

Proof: By (4.1)-(4.4), $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $\alpha A x^{\alpha-1} \geq \alpha B x^{\alpha-1}$ for all x. Hence the result. □

Theorem 4.3.3. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $A \geq B$.*

Proof: By (4.2) and (4.4), $X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $\alpha B x^{\alpha-1} e^{-Bx^{\alpha}} / \left[\alpha A x^{\alpha-1} e^{-Ax^{\alpha}}\right]$ is an increasing function of x. Hence the result. □

Theorem 4.3.4. *Suppose $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ and $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. Then $X_{1:n} \leq_{st} Y_{1:n}$ if and only if one of the following conditions holds:*

(i) $A_1 \geq A_2 B_1 / B_2$ *and* $A_2 \geq B_2$,

(ii) $A_1 < A_2 B_1 / B_2$ *and* $A_1 \geq B_1$.

Proof: By (4.1) and (4.3), $X_{1:n} \leq_{st} Y_{1:n}$ if and only if

$$\frac{A_1 x^{\alpha_1} + A_2 x^{\alpha_2}}{B_1 x^{\alpha_1} + B_2 x^{\alpha_2}} = g(x) \geq 1$$

for all x. The result follows by Lemma 4.2.2. □

Theorem 4.3.5. *$X_{1:n} \leq_{st} Y_{1:n}$ if one of the following conditions holds:*

(i) $\alpha_1 \geq \beta_1$, $\alpha_2 \geq \beta_2$, $\alpha_1 \geq \beta_2$, $\alpha_2 \geq \beta_1$ *and* $g(0) \geq 1$,

(ii) $\alpha_1 < \beta_1$, $\alpha_2 < \beta_2$, $\alpha_1 < \beta_2$, $\alpha_2 < \beta_1$ *and* $g(\infty) \geq 1$,

where $g(x) = \left(A_1 x^{\alpha_1} + A_2 x^{\alpha_2}\right) / \left(B_1 x^{\beta_1} + B_2 x^{\beta_2}\right)$.

Proof: By (4.1) and (4.3), $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $g(x) \geq 1$ for all x. The result follows by Lemma 4.2.1. □

Theorem 4.3.6. *Suppose $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ and $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. Then $X_{1:n} \leq_{lr} Y_{1:n}$ if one of the following conditions holds:*

(i) $\alpha_1 = \alpha_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$,

(ii) $\alpha_1 > \alpha_2$, $A_1 \leq A_2 B_1 / B_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$,

(iii) $\alpha_1 < \alpha_2$, $A_1 \geq A_2 B_1 / B_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$.

Proof: By (4.2) and (4.4), we can write

$$\begin{aligned}
\frac{f_{1:n}(x)}{g_{1:n}(x)} &= \frac{\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}}{\alpha_1 B_1 x^{\alpha_1 - 1} + \alpha_2 B_2 x^{\alpha_2 - 1}} \exp\left[\left(B_1 - A_1\right) x^{\alpha_1} + \left(B_2 - A_2\right) x^{\alpha_2}\right] \\
&= g(x) \exp\left[\left(B_1 - A_1\right) x^{\alpha_1} + \left(B_2 - A_2\right) x^{\alpha_2}\right]
\end{aligned}$$

say. By Lemma 4.2.3, $g(x)$ is a decreasing function of x if and only if one of the following conditions holds

(i) $\alpha_1 = \alpha_2$,

(ii) $\alpha_1 > \alpha_2$ and $A_1 \leq A_2 B_1 / B_2$,

(iii) $\alpha_1 < \alpha_2$ and $A_1 \geq A_2 B_1 / B_2$.

$\exp\left[(B_1 - A_1) x^{\alpha_1} + (B_2 - A_2) x^{\alpha_2}\right]$ is a decreasing function of x if $B_1 \leq A_1$ and $B_2 \leq A_2$. Hence the result. □

Theorem 4.3.7. *Suppose $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ and $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if one of the following conditions holds:*

(i) $A_1 \geq A_2 B_1 / B_2$ and $A_2 \geq B_2$,

(ii) $A_1 < A_2 B_1 / B_2$ and $A_1 \geq B_1$.

Proof: By (4.1)-(4.4), we can write

$$\frac{f_{1:n}(x)/\overline{F}_{1:n}(x)}{g_{1:n}(x)/\overline{G}_{1:n}(x)} = \frac{\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}}{\alpha_1 B_1 x^{\alpha_1 - 1} + \alpha_2 B_2 x^{\alpha_2 - 1}} = g(x)$$

say. $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $g(x) \geq 1$ for all x. The result follows by Lemma 4.2.2. □

Theorem 4.3.8. *$X_{1:n} \leq_{hr} Y_{1:n}$ if one of the following conditions holds:*

(i) $\alpha_1 \geq \beta_1$, $\alpha_2 \geq \beta_2$, $\alpha_1 \geq \beta_2$, $\alpha_2 \geq \beta_1$ and $g(0) \geq 1$,

(ii) $\alpha_1 < \beta_1$, $\alpha_2 < \beta_2$, $\alpha_1 < \beta_2$, $\alpha_2 < \beta_1$ and $g(\infty) \geq 1$,

where $g(x) = \left(\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}\right) / \left(\beta_1 B_1 x^{\beta_1 - 1} + \beta_2 B_2 x^{\beta_2 - 1}\right)$.

Proof: By (4.1)-(4.4), $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $g(x) \geq 1$ for all x. The result follows by Lemma 4.2.1. □

4.4 Results for the smallest of lower truncated Weibull random variables

Let $X_i \sim LTW\left(\alpha_i, \lambda_i\right)$, $i = 1, 2, \ldots, n$ be independent lower truncated Weibull random variables. Let $Y_i \sim LTW\left(\beta_i, \theta_i\right)$, $i = 1, 2, \ldots, n$ be independent lower truncated Weibull random variables. Let

$X_{1:n} = \min(X_1, X_2, \ldots, X_n)$ and $Y_{1:n} = \min(Y_1, Y_2, \ldots, Y_n)$. Note that

$$\overline{F}_{1:n}(x) = \exp\left[n - \sum_{i=1}^{n} (\lambda_i x)^{\alpha_i}\right] \tag{4.5}$$

and

$$f_{1:n}(x) = x^{-1}\left[\sum_{i=1}^{n} \alpha_i (\lambda_i x)^{\alpha_i}\right] \exp\left[n - \sum_{i=1}^{n} (\lambda_i x)^{\alpha_i}\right] \tag{4.6}$$

for $x > 1/\lambda_{\min}$, $\alpha_i > 0$ and $\lambda_i > 0$, where $\lambda_{\min} = \min(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Similarly,

$$\overline{G}_{1:n}(x) = \exp\left[n - \sum_{i=1}^{n} (\theta_i x)^{\beta_i}\right] \tag{4.7}$$

and

$$g_{1:n}(x) = x^{-1}\left[\sum_{i=1}^{n} \beta_i (\theta_i x)^{\beta_i}\right] \exp\left[n - \sum_{i=1}^{n} (\theta_i x)^{\beta_i}\right] \tag{4.8}$$

for $x > 1/\theta_{\min}$, $\beta_i > 0$ and $\theta_i > 0$, where $\theta_{\min} = \min(\theta_1, \theta_2, \ldots, \theta_n)$.

As noted above, $X_{1:n}$ and $Y_{1:n}$ are defined over different domains. Any comparison of these random variables should be over a common domain. We take the common domain of $X_{1:n}$ and $Y_{1:n}$ as (c, ∞), where $c = \max(1/\lambda_{\min}, 1/\theta_{\min})$.

Theorems 4.4.1 and 4.4.4 give if and only if conditions for the stochastic ordering. Theorems 4.4.2 and 4.4.7 give if and only if conditions for the hazard rate ordering. Theorem 4.4.3 gives an if and only if condition for the likelihood ratio ordering. Theorem 4.4.5 gives a condition for the stochastic ordering. Theorem 4.4.6 gives a condition for the likelihood ratio ordering. Theorem 4.4.8 gives a condition for the hazard rate ordering. Throughout this section, we let $A = \sum_{i=1}^{n} \lambda_i^{\alpha}$, $B = \sum_{i=1}^{n} \theta_i^{\alpha}$, $A_1 = \sum_{i=1}^{p} \lambda_i^{\alpha_1}$, $A_2 = \sum_{i=p+1}^{n} \lambda_i^{\alpha_2}$, $B_1 = \sum_{i=1}^{p} \theta_i^{\beta_1}$ and $B_2 = \sum_{i=p+1}^{n} \theta_i^{\beta_2}$.

Theorem 4.4.1. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $A \geq B$.*

Proof: Similar to the proof of Theorem 4.3.1. □

Theorem 4.4.2. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $A \geq B$.*

Proof: Similar to the proof of Theorem 4.3.2. □

Theorem 4.4.3. *Suppose $\alpha_i = \beta_i = \alpha$ for all i. Then $X_{1:n} \leq_{lr} Y_{1:n}$ if and only if $A > B$.*

Proof: Similar to the proof of Theorem 4.3.3. $\square$

Theorem 4.4.4. *Suppose* $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ *and* $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. *Then* $X_{1:n} \leq_{st} Y_{1:n}$ *if and only if one of the following conditions holds:*

(i) $\alpha_2 \geq \alpha_1$, $A_1 \geq A_2 B_1 / B_2$ *and* $A_2 \geq B_2$,

(ii) $\alpha_2 < \alpha_1$, $A_1 \geq A_2 B_1 / B_2$ *and* $g(c) \geq 1$,

(iii) $\alpha_2 \geq \alpha_1$, $A_1 < A_2 B_1 / B_2$ *and* $g(c) \geq 1$,

(iv) $\alpha_2 < \alpha_1$, $A_1 < A_2 B_1 / B_2$ *and* $A_1 \geq B_1$,

where $g(x) = \left(A_1 x^{\alpha_1} + A_2 x^{\alpha_2}\right) / \left(B_1 x^{\alpha_1} + B_2 x^{\alpha_2}\right)$.

Proof: By (4.5) and (4.7), $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $g(x) \geq 1$ for all $x > c$. The result follows by Lemma 4.2.5. $\square$

Theorem 4.4.5. $X_{1:n} \leq_{st} Y_{1:n}$ *if one of the following conditions holds:*

(i) $\alpha_1 \geq \beta_1$, $\alpha_2 \geq \beta_2$, $\alpha_1 \geq \beta_2$, $\alpha_2 \geq \beta_1$ *and* $g(c) \geq 1$,

(ii) $\alpha_1 < \beta_1$, $\alpha_2 < \beta_2$, $\alpha_1 < \beta_2$, $\alpha_2 < \beta_1$ *and* $g(\infty) \geq 1$,

where $g(x) = \left(A_1 x^{\alpha_1} + A_2 x^{\alpha_2}\right) / \left(B_1 x^{\beta_1} + B_2 x^{\beta_2}\right)$.

Proof: By (4.5) and (4.7), $X_{1:n} \leq_{st} Y_{1:n}$ if and only if $g(x) \geq 1$ for all $x > c$. The result follows by Lemma 4.2.4. $\square$

Theorem 4.4.6. *Suppose* $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ *and* $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. *Then* $X_{1:n} \leq_{lr} Y_{1:n}$ *if one of the following conditions holds:*

(i) $\alpha_1 = \alpha_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$,

(ii) $\alpha_1 > \alpha_2$, $A_1 \leq A_2 B_1 / B_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$,

(iii) $\alpha_1 < \alpha_2$, $A_1 \geq A_2 B_1 / B_2$, $B_1 \leq A_1$ *and* $B_2 \leq A_2$.

Proof: By (4.6) and (4.8), we can write

$$\begin{aligned}
\frac{f_{1:n}(x)}{g_{1:n}(x)} &= \frac{\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}}{\alpha_1 B_1 x^{\alpha_1 - 1} + \alpha_2 B_2 x^{\alpha_2 - 1}} \exp\left[\left(B_1 - A_1\right) x^{\alpha_1} + \left(B_2 - A_2\right) x^{\alpha_2}\right] \\
&= g(x) \exp\left[\left(B_1 - A_1\right) x^{\alpha_1} + \left(B_2 - A_2\right) x^{\alpha_2}\right]
\end{aligned}$$

say. By Lemma 4.2.6, $g(x)$ is a decreasing function of x if and only if one of the following conditions holds

(i) $\alpha_1 = \alpha_2$,

(ii) $\alpha_1 > \alpha_2$ and $A_1 \leq A_2 B_1 / B_2$,

(iii) $\alpha_1 < \alpha_2$ and $A_1 \geq A_2 B_1 / B_2$.

$\exp\left[(B_1 - A_1) x^{\alpha_1} + (B_2 - A_2) x^{\alpha_2}\right]$ is a decreasing function of x if $B_1 \leq A_1$ and $B_2 \leq A_2$. Hence the result. □

Theorem 4.4.7. *Suppose $\alpha_i = \beta_i = \alpha_1$, $i = 1, 2, \ldots, p$ and $\alpha_i = \beta_i = \alpha_2$, $i = p+1, p+2, \ldots, n$. Then $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if one of the following conditions holds:*

(i) $\alpha_2 \geq \alpha_1$, $A_1 \geq A_2 B_1 / B_2$ *and* $A_2 \geq B_2$,

(ii) $\alpha_2 < \alpha_1$, $A_1 \geq A_2 B_1 / B_2$ *and* $g(c) \geq 1$,

(iii) $\alpha_2 \geq \alpha_1$, $A_1 < A_2 B_1 / B_2$ *and* $g(c) \geq 1$,

(iv) $\alpha_2 < \alpha_1$, $A_1 < A_2 B_1 / B_2$ *and* $A_1 \geq B_1$,

where $g(x) = \left(\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}\right) / \left(\alpha_1 B_1 x^{\alpha_1 - 1} + \alpha_2 B_2 x^{\alpha_2 - 1}\right)$.

Proof: By (4.5)-(4.8), $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $g(x) \geq 1$ for all $x > c$. The result follows by Lemma 4.2.5. □

Theorem 4.4.8. $X_{1:n} \leq_{hr} Y_{1:n}$ *if one of the following conditions holds:*

(i) $\alpha_1 \geq \beta_1$, $\alpha_2 \geq \beta_2$, $\alpha_1 \geq \beta_2$, $\alpha_2 \geq \beta_1$ *and* $g(c) \geq 1$,

(ii) $\alpha_1 < \beta_1$, $\alpha_2 < \beta_2$, $\alpha_1 < \beta_2$, $\alpha_2 < \beta_1$ *and* $g(\infty) \geq 1$,

where $g(x) = \left(\alpha_1 A_1 x^{\alpha_1 - 1} + \alpha_2 A_2 x^{\alpha_2 - 1}\right) / \left(\beta_1 B_1 x^{\beta_1 - 1} + \beta_2 B_2 x^{\beta_2 - 1}\right)$.

Proof: By (4.5)-(4.8), $X_{1:n} \leq_{hr} Y_{1:n}$ if and only if $g(x) \geq 1$ for all $x > c$. The result follows by Lemma 4.2.4. □

4.5 Real data application

Here, we illustrate the results of Section 4.3 using a real data set. The data are from one of the first successful trials of adjuvant chemotherapy for colon cancer. The data has several variables, but the ones we use here are: i) days until recurrence or death; ii) number of lymph nodes with detectable cancer (a number from 1 to 10); iii) time from surgery to registration (0=short, 1=long). We shall refer to days until recurrence or death as the time. We shall refer to patients with short time from surgery to registration as short-time patients. We shall refer to patients with long time from surgery to registration as long-time patients. More details of the data can be found in the R contributed package survival (R Development Core Team, 2015).

Note that, the total sample size is 1858, among which there are 36 missing values and 4 patients with no lymph node. There are 1720 patients with the number of lymph nodes ranging from 1 to 10 in either short-time or long-time group. There are also 98 patients having the number of lymph nodes within the region of $[11, 33]$. However, they are too disperse to construct the sufficient paired-sample (at least 5 samples in both short-time and long-time group) for the particular number of lymph nodes. Thus, we are going to focus on the 1720 patients who have less or equal to 10 lymph nodes in our study.

The medical doctor would like to avoid having a small time to recurrence or a small time to death. We ask the questions: is the minimum time smaller for short-term patients than for long-term patients? Is it the other way around? To answer these questions, we take $X_1, X_2, \ldots, X_{10}$ as denoting the times for short-term patients with the number of lymph nodes with detectable cancer ranging from 1 to 10. We take $Y_1, Y_2, \ldots, Y_{10}$ as denoting the times for long-term patients with the number of lymph nodes with detectable cancer ranging from 1 to 10. The aim is to see how $X_{1:10}$ and $Y_{1:10}$ are ordered.

It is reasonable to assume that $X_1, X_2, \ldots, X_{10}$ are independent random variables since they correspond to different numbers of lymph nodes with detectable cancer. Likewise for $Y_1, Y_2, \ldots, Y_{10}$. We suppose X_i has the Weibull distribution with parameters (α_i, λ_i) and Y_i has the Weibull distribution with parameters (β_i, θ_i) for $i = 1, 2, \ldots, 10$.

The log-likelihood function of the X_i and Y_i are

$$L1_i \quad : \quad \log L(x_i) = n_i \log(\alpha_i) + \alpha_i n_i \log \lambda_i + (\alpha_i - 1) \sum_{j=1}^{n_i} \log(x_{ij}) - \sum_{j=1}^{n_i} (\lambda_i x_{ij})^{\alpha_i}$$

and

$$L2_i \quad : \quad \log L(y_i) = m_i \log(\beta_i) + \beta_i m_i \log \theta_i + (\beta_i - 1) \sum_{j=1}^{m_i} \log(y_{ij}) - \sum_{j=1}^{m_i} (\theta_i y_{ij})^{\beta_i}$$

where, n_i and m_i are the sample sizes of X_i and Y_i respectively.

The maximum likelihood estimates of these parameters are shown in Table 4.1. Also given in this table are p-values of the Kolmogorov test for goodness of fit of the Weibull distribution. These

Time	Nodes	Size	Parameter estimates(s.d.)	p-value for Kolmogorov Test
0	1	418	$\widehat{\alpha_1} = 2.180(0.091)$, $\widehat{\lambda_1} = 0.000487(2.00\text{E-}13)$	0.111
0	2	256	$\widehat{\alpha_2} = 2.091(0.114)$, $\widehat{\lambda_2} = 0.000496(2.00\text{E-}13)$	0.385
0	3	186	$\widehat{\alpha_3} = 1.807(0.111)$, $\widehat{\lambda_3} = 0.000571(2.00\text{E-}13)$	0.637
0	4	128	$\widehat{\alpha_4} = 1.377(0.101)$, $\widehat{\lambda_4} = 0.000592(2.00\text{E-}13\ )$	0.783
0	5	68	$\widehat{\alpha_5} = 1.145(0.110)$, $\widehat{\lambda_5} = 0.000827(2.00\text{E-}13)$	0.171
0	6	56	$\widehat{\alpha_6} = 1.215(0.125)$, $\widehat{\lambda_6} = 0.000751(\ 2.00\text{E-}13\ )$	0.051
0	7	60	$\widehat{\alpha_7} = 1.250(0.127)$, $\widehat{\lambda_7} = 0.000792(2.00\text{E-}13)$	0.327
0	8	36	$\widehat{\alpha_8} = 0.976(0.122)$, $\widehat{\lambda_8} = 0.00113(2.00\text{E-}13)$	0.388
0	9	32	$\widehat{\alpha_9} = 1.120(0.029)$, $\widehat{\lambda_9} = 0.000778(2.00\text{E-}13)$	0.569
0	10	18	$\widehat{\alpha_{10}} = 1.092(0.208)$, $\widehat{\lambda_{10}} = 0.000811(2.00\text{E-}13)$	0.063
1	1	130	$\widehat{\beta_1} = 1.940(0.146)$, $\widehat{\theta_1} = 0.000489(2.00\text{E-}13)$	0.549
1	2	132	$\widehat{\beta_2} = 1.437(0.104)$, $\widehat{\theta_2} = 0.000580(2.00\text{E-}13)$	0.283
1	3	64	$\widehat{\beta_3} = 1.753(0.186)$, $\widehat{\theta_3} = 0.000599(2.00\text{E-}13)$	0.057
1	4	40	$\widehat{\beta_4} = 1.083(0.133)$, $\widehat{\theta_4} = 0.000858(2.00\text{E-}13)$	0.218
1	5	24	$\widehat{\beta_5} = 1.110(0.162)$, $\widehat{\theta_5} = 0.00114(2.00\text{E-}13)$	0.980
1	6	30	$\widehat{\beta_6} = 1.026(0.145)$, $\widehat{\theta_6} = 0.000853(2.00\text{E-}13)$	0.078
1	7	16	$\widehat{\beta_7} = 1.248(0.250)$, $\widehat{\theta_7} = 0.000770(2.00\text{E-}13)$	0.258
1	8	10	$\widehat{\beta_8} = 1.071(0.251)$, $\widehat{\theta_8} = 0.000958(2.00\text{E-}13)$	0.482
1	9	8	$\widehat{\beta_9} = 0.973(0.288)$, $\widehat{\theta_9} = 0.000758(2.00\text{E-}13)$	0.880
1	10	8	$\widehat{\beta_{10}} = 1.718(0.447)$, $\widehat{\theta_{10}} = 0.00217(0.00035)$	0.972

Table 4.1: Sample sizes, parameter estimates with standard deviation, and p-values of the Kolmogorov test.

p-values show that the Weibull distribution provides reasonable fit. As metioned in the previous chapter, because the parameters are estimated instead of taking true values, the Kolmogorov test may be generating a conservative p-value. To address this issue, we can use Fisher's method. It assumes that $-2\sum_{i=1}^{k}\ln(p_i) \sim \chi^2_{2k}$, where p_i is the p-value for the i^{th} hypothesis test. Here, in our application, $k = 20$, $-2\sum_{i=1}^{k}\ln(p_i) = 50.09$. The statistic is much smaller than the critical value 70.618 ($= \chi^2_{40}$), which indicates that we fail to reject all the null hypotheses. That is, the Weibull distribution provides resonable fit. An alternative method for goodness-of-fit can be the Stephen's half-sample method (Stephens et. al., 1978). As further evidence, probability plots of the fits are shown in Figures 4.1 and 4.2. These results are expected since Weibull distributions are popular models for lifetime data.

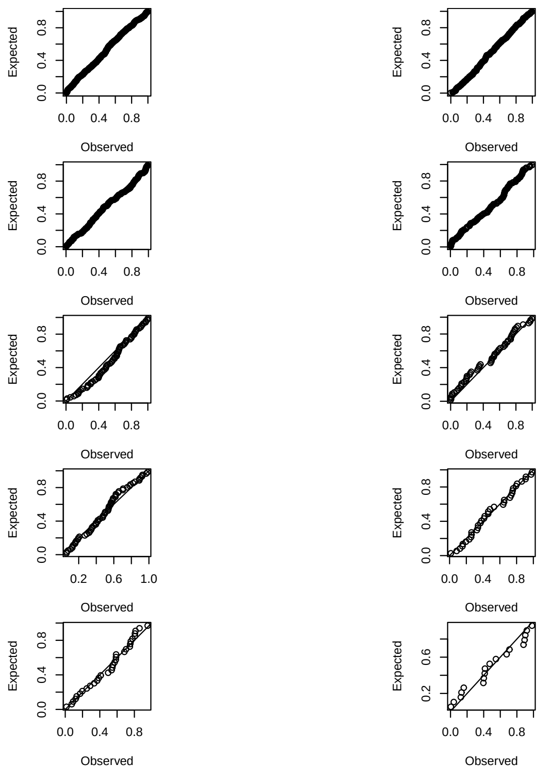

Figure 4.1: Probability plots of the fits of the Weibull distribution for survival times for short-time patients for the 10 different values of the number of lymph nodes with detectable cancer.

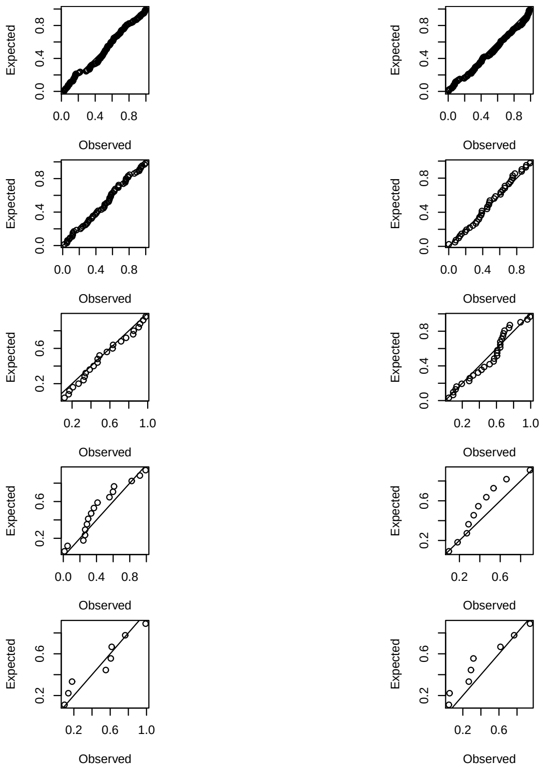

Figure 4.2: Probability plots of the fits of the Weibull distribution for survival times for long-time patients for the 10 different values of the number of lymph nodes with detectable cancer.

Suppose that $\alpha_i = \beta_i = \alpha$ for all i, the log-likelihood function for X and Y can be written as

$$L3 \quad : \quad \log L(x,y) = \sum_{i=1}^{10}\left(n_i \log(\alpha) + \alpha n_i \log\lambda_i + (\alpha-1)\sum_{j=1}^{n_i}\log(x_{ij}) - \sum_{j=1}^{n_i}(\lambda_i x_{ij})^{\alpha}\right)$$
$$+\sum_{i=1}^{10}\left(m_i \log(\alpha) + \alpha m_i \log\theta_i + (\alpha-1)\sum_{j=1}^{m_i}\log(y_{ij}) - \sum_{j=1}^{m_i}(\theta_i y_{ij})^{\alpha}\right)$$

The likelihood ratio test is a statistical tool for testing the goodness of fit of two models which one (the null model) is a special case of the other (the alternative model). Here in our models, the null hyperthesis is $\alpha_i = \beta_i = \alpha$, while the alternative hyperthesis is $\alpha_i \neq \beta_i$. By the definition of likelihood ratio test, the statistic $-2\left(L3 - \sum_{i=1}^{10}(L1_i + L2_i)\right)$ follows the χ^2 distribution with the degree of freedom ν. Here, $\nu = 1$.

The likelihood ratio test of $\alpha_i = \beta_i = \alpha$ for all i gave a p-value of 0.122. So, we can not reject the null hyperthesis at the 5% level of ginificance. That is, $\alpha_i = \beta_i = \alpha$ for all i. Given the parameter estimates in Table 4.1, we have $\widehat{A} = \sum_{i=1}^{10} \widehat{\lambda}_i^{\widehat{\alpha}} = 4.273 \times 10^{-46}$ and $\widehat{B} = \sum_{i=1}^{10} \widehat{\theta}_i^{\widehat{\alpha}} = 6.410 \times 10^{-42}$.

Since $\widehat{A} < \widehat{B}$, Theorem 4.3.1 shows that the minimum lifetime of short-time patients is stochastically greater than the minimum lifetime of long-term patients. Since $\widehat{A} < \widehat{B}$, Theorem 4.3.2 shows that the minimum lifetime of short-time patients is greater than the minimum lifetime of long-term patients with respect to hazard rate order. Since $\widehat{A} < \widehat{B}$, Theorem 4.3.3 shows that the minimum lifetime of short-time patients is greater than the minimum lifetime of long-term patients with respect to likelihood ratio order.

The established orderings between $X_{1:10}$ and $Y_{1:10}$ are confirmed by the plots of the ratio of survival functions, the ratio of probability density functions and the ratio of hazard rate functions shown in Figure 4.3.

There are formal tests for checking orderings of variables in Anderson (1996) and Barrett and Donald (2003).

Since the previous tests involve inverses of matrics, which may not be available under some circumstance, our conditions (e.g. Theorems 4.3.1 to 4.3.3) are easy to be applied. So, we suggest that our method can be used as the first step to target the interest. Those formal tests then can be used in the further study.

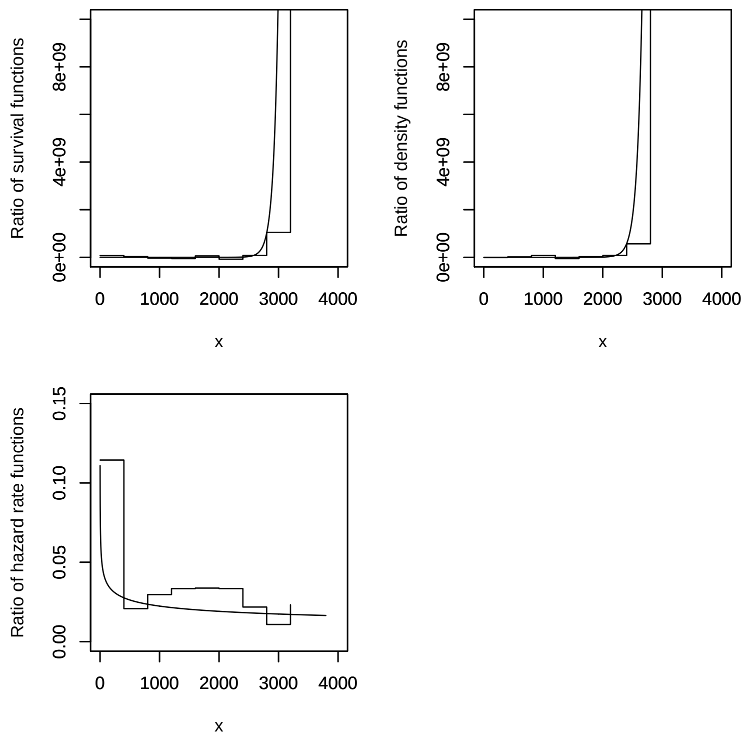

Figure 4.3: Plots of $\widehat{\overline{F}_{1:10}}(x)/\widehat{\overline{G}_{1:10}}(x)$ (top left), $\widehat{f_{1:10}}(x)/\widehat{g_{1:10}}(x)$ (top right) and $\left[\widehat{f_{1:10}}(x)/\widehat{\overline{F}_{1:10}}(x)\right] / \left[\widehat{g_{1:10}}(x)/\widehat{\overline{F}_{1:10}}(x)\right]$ (bottom left). Also shown are the empirical versions of these ratios.

Chapter 5

New discrete bivariate distributions

5.1 Introduction

Dependent random quantities in a wide range of areas have been modelled through the application of bivariate/multivariate distributions. Historically many continuous bivariate/multivariate distributions have been developed and used, in particular the bivariate/multivariate exponential distribution and various extensions of it. There has been relatively little research on the development of discrete bivariate/multivariate distributions at least up until 2010. For instance, a Google Scholar search identified only the following discrete bivariate/multivariate distributions for the period 2001-2009 (A collection of discrete bivariate/multivariate distributions developed before 2000 can be found in Johnson et al. (1997)): discrete multivariate distributions with product-type dependence due to Becker and Utev (2002); discrete multivariate distributions induced by an urn scheme due to Nikulin et al. (2002); bivariate Poisson distributions due to Piperigou and Papageorgiou (2003); discrete multivariate distributions based on generalized linear mixed models due to Tonda (2005); a discrete multivariate distribution resulting from the law of small numbers due to Hoshino (2006); the discrete bivariate generalized hypergeometric factorial moment distribution due to Kumar (2007); conditionally specified discrete multivariate distributions due to Ip and Wang (2009). Only five papers have appeared for the period 2003-2009! The number of papers proposing continuous bivariate/multivariate distributions is several times larger.

In recent years, the need for discrete bivariate/multivariate distributions has become increasingly important in many applied areas: modeling of football matches (McHale and Scarf, 2007); modeling of family size in beef cattle (Garrido et al., 2008); analysis of crime data (Miethe et al., 2008); modeling of the dependence of goals scored by opposing teams in international soccer matches (McHale and Scarf, 2011); risk assessment and fault detection using scarce data (Ahooyi et al., 2014); to mention just a few. This has led to many discrete bivariate/multivariate distributions being developed since 2010: the discrete bivariate normal distribution due to Bairamov and Gultekin (2010); bivariate compound Poisson distributions due to Ozel (2011a, 2011b, 2013); composite discrete bivariate distributions due to Cacoullos and Papageorgiou (2012); the multivariate discrete Poisson-Lindley distribution due to Gomez-Deniz et al. (2012); multivariate discrete distributions based on pair copulas due to Panagiotelis et al. (2012); the discrete bivariate Linnik distribution due to Antony and Jayakumar (2013); multivariate discrete distributions based on general copulas due to Nikoloulopoulos (2013); the bivariate generalization of the noncentral negative binomial distribution due to Ong and Ng (2013); discrete multivariate phase-type distributions due to He and Ren (2015); the composite discrete bivariate distribution with uniform marginals due to Reilly and Sapkota (2015).

The most recent discrete bivariate distribution is due to Lee and Cha (2015). They constructed two general classes of discrete bivariate distributions by applying the minimum and maximum operators to independent discrete random variables. Let U_1, U_2 and U_3 be independent discrete random variables. The distributions constructed in Lee and Cha (2015) are those of $(X, Y) = (\min(U_1, U_3), \min(U_2, U_3))$ and $(X, Y) = (\max(U_1, U_3), \max(U_2, U_3))$. We shall refer to these distributions as Models C and D. Model A shall refer to the joint distribution of the independent variables $(X, Y) = (U_1, U_2)$. Model B shall refer to the joint distribution of $(X, Y) = (U_1 + U_3, U_2 + U_3)$ due to Holgate (1964).

Following the methodology in Lee and Cha (2015), we construct further discrete bivariate distributions by applying possible combinations of common mathematical operators: addition (+), multiplication (×), minimum (min) and maximum (max). It turns out that seven new discrete bivariate distributions can be constructed these operators.

The seven discrete bivariate distributions and expressions for their probability mass function (pmf), cumulative distribution function (cdf), product moments and moment generating functions are given in Section 5.2. The performance of the eighteen distributions versus Models A to D for the football data in Lee and Cha (2015) is investigated in Section 5.3. It turns out some of the newly proposed distributions provide better fits than Models A to D.

5.2 New distributions

Here, we introduce seven new discrete bivariate distributions, in addition to the two proposed by Lee and Cha (2015). The seven distributions are motivated as joint distributions of: **Model 1:** $X = U_1 + U_3$ and $Y = U_2U_3$; **Model 2:** $X = U_1 + U_3$ and $Y = \min(U_2, U_3)$; **Model 3:** $X = U_1 + U_3$ and $Y = \max(U_2, U_3)$; **Model 4:** $X = U_1U_3$ and $Y = U_2U_3$; **Model 5:** $X = U_1U_3$ and $Y = \min(U_2, U_3)$; **Model 6:** $X = U_1U_3$ and $Y = \max(U_2, U_3)$; **Model 7:** $X = \min(U_1, U_3)$ and $Y = \max(U_2, U_3)$. We suppose throughout this section that U_1, U_2 and U_3 are independent discrete random variables each defined on the set of all non-negative integers (zero and all positive integers). We can note that Model 1 and Model 4 only allow for positive correlations. In Model 3 and Model 6, Y and X exhibit the causal relationship when $U_3 > U_2$,. While in Model 2 and Model 5, Y and X exhibit the causal relationship when $U_3 < U_2$,. This character can be applied for many practical uses, e.g. length of hospital stay in days and cost in pounds. We denote the pmf of U_i by p_i for $i = 1, 2, 3$. We denote the cdf of U_i by P_i for $i = 1, 2, 3$. We denote the mgf of U_i by M_i for $i = 1, 2, 3$. For each joint distribution of X and Y, we give derivations and expressions for the joint pmf, joint cdf, product moments and joint moment generating function whenever possible.

Model 1 For $X = U_1 + U_3$ and $Y = U_2U_3$,

$$\Pr(X = x, Y = y) = \begin{cases} p_1(0)p_3(0), & \text{if } x = 0,\ y = 0, \\ \\ p_1(x)p_3(0) + p_2(0)\displaystyle\sum_{k=1}^{x} p_1(x-k)p_3(k), & \text{if } x > 0,\ y = 0, \\ \\ \displaystyle\sum_{k=1}^{x} p_1(x-k)p_2(y/k)p_3(k), & \text{if } x > 0,\ y > 0, \end{cases}$$

$$\Pr(X \le x, Y \le y) = \begin{cases} p_1(0)p_3(0), & \text{if } x = 0,\ y = 0, \\ P_1(x)p_3(0) + p_2(0)\sum_{i=1}^{x}\sum_{k=1}^{i} p_1(i-k)p_3(k), & \text{if } x > 0,\ y = 0, \\ p_1(0)p_3(0) + P_1(x)p_3(0) \\ +p_2(0)\sum_{i=1}^{x}\sum_{k=1}^{i} p_1(i-k)p_3(k) \\ +\sum_{i=1}^{x}\sum_{j=1}^{y}\sum_{k=1}^{i} p_1(i-k)p_2(j/k)p_3(k), & \text{if } x > 0,\ y > 0, \end{cases}$$

$$\begin{aligned} E\left(X^m Y^n\right) &= E\left((U_1+U_3)^m \left(U_2 U_3\right)^n\right) \\ &= E\left(\sum_{i=0}^{m}\binom{m}{i} U_1^i U_3^{m-i} \cdot U_2^n U_3^n\right) \\ &= \sum_{i=0}^{m}\binom{m}{i} E\left(U_1^i\right) E\left(U_2^n\right) E\left(U_3^{m-i+n}\right) \end{aligned}$$

and

$$\begin{aligned} E\left[\exp\left(sX+tY\right)\right] &= E\left[\exp\left(s\left(U_1+U_3\right)+t\left(U_2 U_3\right)\right)\right] \\ &= E\left[\exp\left(s\cdot U_1 + \left(s + t\cdot U_2\right)U_3\right)\right] \\ &= E\left[\exp\left(s\cdot U_1\right)\times\exp\left(\left(s+t\cdot U_2\right)U_3\right)\right] \\ &= E\left[\exp\left(s\cdot U_1\right)\right]\cdot E\left[\exp\left(\left(s+t\cdot U_2\right)U_3\right)\right] \\ &= M_1(s)\sum_{i=0}^{\infty} M_3(s+it)p_2(i). \end{aligned}$$

In particular, Cov $(X,Y) = E\left(U_2\right) Var\left(U_3\right)$, which is always positive.

Model 2 For $X = U_1 + U_3$ and $Y = \min\left(U_2, U_3\right)$,

$$\begin{aligned} \Pr(X=x, Y=y) &= \Pr\left(U_1+U_3=x, U_2=y \mid U_2 \le U_3\right) \\ &\quad + \Pr\left(U_1+U_3=x, U_3=y \mid U_2 > U_3\right) \\ &= \Pr\left(U_1+U_3=x, U_2=y \mid y \le U_3 \le x\right) \\ &\quad + \Pr\left(U_1+U_3=x, U_3=y \mid U_2 > y\right) \\ &= p_2(y)\sum_{k=y+1}^{x} p_1(x-k)p_3(k) + p_1(x-y)p_3(y)\left[1-P_2(y)\right] \end{aligned}$$

and

$$\begin{aligned}\Pr(X \le x, Y \le y) &= \sum_{i=0}^{x}\sum_{j=0}^{y} p_2(j) \sum_{k=j+1}^{i} p_1(i-k)p_3(k) \\ &\quad + \sum_{j=0}^{y} P_1(x-j)p_3(j)\left[1-P_2(j)\right].\end{aligned}$$

Model 3 For $X = U_1 + U_3$ and $Y = \max(U_2, U_3)$,

$$\begin{aligned}\Pr(X = x, Y = y) &= \Pr\left(U_1 + U_3 = x, U_2 = y \mid U_2 > U_3\right) \\ &\quad + \Pr\left(U_1 + U_3 = x, U_3 = y \mid U_2 \le U_3\right) \\ &= p_1(x-y)P_2(y)p_3(y) + p_2(y) \sum_{k=0}^{\min(x,y-1)} p_1(x-k)p_3(k)\end{aligned}$$

and

$$\begin{aligned}\Pr(X \le x, Y \le y) &= \sum_{i=0}^{x}\sum_{j=0}^{y} \left[\Pr(X = i, Y = j)\right] \\ &= \sum_{j=0}^{y} P_1(x-j)P_2(j)p_3(j) + \sum_{i=0}^{x}\sum_{j=0}^{y} p_2(j) \sum_{k=0}^{\min(i,j-1)} p_1(i-k)p_3(k).\end{aligned}$$

Model 4 For $X = U_1U_3$ and $Y = U_2U_3$,

$$\Pr(X = x, Y = y) = \begin{cases} p_1(0)p_2(0)\left[1-p_3(0)\right] + p_3(0), & \text{if } x = 0,\ y = 0, \\ p_2(0)\sum_{k=1}^{\infty} p_1(x/k)p_3(k), & \text{if } x > 0,\ y = 0, \\ p_1(0)\sum_{k=1}^{\infty} p_2(y/k)p_3(k), & \text{if } x = 0,\ y > 0, \\ \sum_{k=1}^{\infty} p_1(x/k)p_2(y/k)p_3(k), & \text{if } x > 0,\ y > 0, \end{cases}$$

$$\Pr(X \le x, Y \le y) = \begin{cases} p_1(0)p_2(0)\left[1-p_3(0)\right]+p_3(0), & \text{if } x=0,\ y=0, \\ \\ p_1(0)p_2(0)\left[1-p_3(0)\right]+p_3(0) \\ +p_2(0)\displaystyle\sum_{k=1}^{\infty}\sum_{i=1}^{x} p_1(i/k)p_3(k), & \text{if } x>0,\ y=0, \\ \\ p_1(0)p_2(0)\left[1-p_3(0)\right]+p_3(0) \\ +p_1(0)\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{y} p_2(j/k)p_3(k), & \text{if } x=0,\ y>0, \\ \\ p_1(0)p_2(0)\left[1-p_3(0)\right]+p_3(0) \\ +p_2(0)\displaystyle\sum_{k=1}^{\infty}\sum_{i=1}^{x} p_1(i/k)p_3(k) \\ +p_1(0)\displaystyle\sum_{k=1}^{\infty}\sum_{j=1}^{y} p_2(j/k)p_3(k) \\ +\displaystyle\sum_{k=1}^{\infty}\sum_{i=1}^{x} p_1(i/k)\sum_{j=1}^{y} p_2(j/k)p_3(k), & \text{if } x>0,\ y>0, \end{cases}$$

$$\begin{aligned} E\left(X^m Y^n\right) &= E\left[\left(U_1U_3\right)^m\left(U_2U_3\right)^n\right] \\ &= E\left[U_1^m U_2^n U_3^{m+n}\right] \\ &= E\left(U_1^m\right)E\left(U_2^n\right)E\left(U_3^{m+n}\right) \end{aligned}$$

and

$$\begin{aligned} E\left[\exp\left(sX+tY\right)\right] &= E\left[\exp\left(sU_1U_3+tU_2U_3\right)\right] \\ &= E\left[\exp\left(sU_1U_3\right)\exp\left(tU_2U_3\right)\right] \\ &= E\left[\exp\left(sU_1U_3\right)\right]\cdot E\left[\exp\left(tU_2U_3\right)\right] \\ &= \sum_{i=0}^{\infty} M_1(si)M_2(ti)p_3(i). \end{aligned}$$

In particular, Cov $(X,Y) = E\left(U_1\right)E\left(U_2\right)Var\left(U_3\right)$, which is always positive.

Model 5 For $X = U_1U_3$ and $Y = \min(U_2, U_3)$,

$$\Pr(X = x, Y = y) = \begin{cases} p_1(0)p_2(0)\left[1 - p_3(0)\right] + p_3(0), & \text{if } x = 0,\ y = 0, \\ p_2(0)\displaystyle\sum_{k=1}^{\infty} p_1(x/k)p_3(k), & \text{if } x > 0,\ y = 0, \\ p_1(0)p_2(y)\left[1 - P_3(y-1)\right] \\ +p_1(0)p_3(y)\left[1 - P_2(y)\right], & \text{if } x = 0,\ y > 0, \\ p_1(x/y)\left[1 - P_2(y-1)\right]p_3(y) \\ +p_2(y)\displaystyle\sum_{k=y+1}^{\infty} p_1(x/k)p_3(k), & \text{if } x > 0,\ y > 0 \end{cases}$$

and

$$\Pr(X \le x, Y \le y) = \begin{cases} p_1(0)p_2(0)\left[1 - p_3(0)\right] + p_3(0), & \text{if } x = 0,\ y = 0, \\ p_1(0)p_2(0)\left[1 - p_3(0)\right] + p_3(0) \\ +p_2(0)\displaystyle\sum_{k=1}^{\infty}\sum_{i=1}^{x} p_1(i/k)p_3(k), & \text{if } x > 0,\ y = 0, \\ p_1(0)p_2(0)\left[1 - p_3(0)\right] + p_3(0) \\ +p_1(0)\displaystyle\sum_{j=1}^{y} p_2(j)\left[1 - P_3(j-1)\right] \\ +p_1(0)\displaystyle\sum_{j=1}^{y} p_3(j)\left[1 - P_2(j)\right], & \text{if } x = 0,\ y > 0, \\ p_1(0)p_2(0)\left[1 - p_3(0)\right] + p_3(0) \\ +p_2(0)\displaystyle\sum_{k=1}^{\infty}\sum_{i=1}^{x} p_1(i/k)p_3(k) \\ +p_1(0)\displaystyle\sum_{j=1}^{y} p_2(j)\left[1 - P_3(j-1)\right] \\ +p_1(0)\displaystyle\sum_{j=1}^{y} p_3(j)\left[1 - P_2(j)\right] \\ +\displaystyle\sum_{i=1}^{x}\sum_{j=1}^{y} p_1(i/j)\left[1 - P_2(j-1)\right]p_3(j) \\ +\displaystyle\sum_{i=1}^{x}\sum_{j=1}^{y} p_2(j)\sum_{k=j+1}^{\infty} p_1(i/k)p_3(k), & \text{if } x > 0,\ y > 0. \end{cases}$$

Model 6 For $X = U_1U_3$ and $Y = \max(U_2, U_3)$,

$$\Pr(X=x,Y=y)=\begin{cases} p_2(0)p_3(0), & \text{if } x=0,\ y=0,\\[2ex] p_1(0)p_2(y)\left[P_3(y)-P_3(0)\right] \\ +p_3(0)p_2(y)+p_1(0)P_2(y-1)p_3(y), & \text{if } x=0,\ y>0,\\[2ex] p_1(x/y)P_2(y)p_3(y)+p_2(y)\displaystyle\sum_{k=1}^{y-1}p_1(x/k)p_3(k), & \text{if } x>0,\ y>0 \end{cases}$$

and

$$\Pr(X\le x,Y\le y)=\begin{cases} p_2(0)p_3(0), & \text{if } x=0,\ y=0,\\[2ex] p_2(0)p_3(0)+p_1(0)\displaystyle\sum_{j=1}^{y}p_2(j)\left[P_3(j)-P_3(0)\right] \\ +p_3(0)\left[P_2(y)-P_2(0)\right]+p_1(0)\displaystyle\sum_{j=1}^{y}P_2(j-1)p_3(j), & \text{if } x=0,\ y>0,\\[2ex] p_2(0)p_3(0)+p_1(0)\displaystyle\sum_{j=1}^{y}p_2(j)\left[P_3(j)-P_3(0)\right] \\ +p_3(0)\left[P_2(y)-P_2(0)\right]+p_1(0)\displaystyle\sum_{j=1}^{y}P_2(j-1)p_3(j) \\ +\displaystyle\sum_{i=1}^{x}\sum_{j=1}^{y}p_1(i/j)P_2(j)p_3(j) \\ +\displaystyle\sum_{i=1}^{x}\sum_{j=1}^{y}p_2(j)\sum_{k=1}^{j-1}p_1(i/k)p_3(k), & \text{if } x>0,\ y>0. \end{cases}$$

Model 7 For $X = \min(U_1, U_3)$ and $Y = \max(U_2, U_3)$,

$$\Pr(X=x,Y=y)=\begin{cases} p_1(x)P_2(x)p_3(x)+\left[1-P_1(x)\right]P_2(x)p_3(x), & \text{if } x=y,\\[2ex] p_1(x)P_2(x+1)p_3(x+1) \\ +\left[1-P_1(x-1)\right]p_2(x+1)p_3(x), & \text{if } y=x+1,\\[2ex] p_1(x)P_2(y)p_3(y)+\left[1-P_1(x-1)\right]p_2(y)p_3(x) \\ +p_1(x)p_2(y)\left[P_3(y)-P_3(x-1)\right], & \text{if } y\ge x+2 \end{cases}$$

and

$$\begin{aligned}
\Pr(X \leq x, Y \leq y) \quad = \quad & \sum_{i=0}^{\min(x,y)} p_1(i)P_2(i)p_3(i) + \sum_{i=0}^{\min(x,y)} \left[1 - P_1(i)\right] P_2(i)p_3(i) \\
& + \sum_{i=0, i+1<y}^{x} p_1(i)P_2(i+1)p_3(i+1) + \sum_{i=0, i+1<y}^{x} \left[1 - P_1(i-1)\right] p_2(i+1)p_3(i) \\
& + \sum_{i=0}^{x} \sum_{j=i+2}^{y} p_1(i)P_2(j)p_3(j) \\
& + \sum_{i=0}^{x} \sum_{j=i+2}^{y} \left[1 - P_1(i-1)\right] p_2(j)p_3(i) \\
& + \sum_{i=0}^{x} \sum_{j=i+2}^{y} p_1(i)p_2(j) \left[P_3(j-1) - P_3(i+1)\right].
\end{aligned}$$

5.3 Data application

Here, we reanalyze the data used in Lee and Cha (2015). The data are the scores of twenty six football matches between ACF Fiorentina and Juventus over the period from 1996 to 2011. The data are reproduced in Table A.1, where X = the number of goals scored by ACF Fiorentina and Y = the number of goals scored by Juventus. Since this dataset spans 16 years, each team may be expected to change substatially. This application is just done as an illustration of the practical use of these newly proposed models. The football data what we use here can be found in Appendix A.

Some summary statistics on X are: minimum = 0, first quartile = 1, median = 1, mean = 1.077, third quartile = 1, maximum = 3, standard deviation = 0.935 and variance = 0.874. Some summary statistics on Y are: minimum = 0, first quartile = 1, median = 1, mean = 1.385, third quartile = 2, maximum = 3, standard deviation = 0.852 and variance = 0.726. The sample correlation coefficient between X and Y is 0.413.

So, on average Juventus has scored more goals than ACF Fiorentina has. But the variability of the number of goals scored by ACF Fiorentina is greater. The highest number of goals scored by both teams are the same. The lowest number of goals scored by both teams are also the same.

Lee and Cha (2015) fitted Models A, B, C and D to the data. They also fitted the same models to a transformed data with X replaced by $3 - X$ and Y replaced by $3 - Y$. We shall refer to the data in Table A.1 as the football data. We shall refer to the transformed data as the transformed football data.

We fitted Models A to D as well as Models 1 to 7 to both data sets. We took U_1, U_2 and U_3 to be independent Poisson random variables with parameters λ_1, λ_2 and λ_3, respectively. The method of maximum likelihood was used for fitting of the models. The nlm function in the R software (R Development Core Team, 2015) was used to maximize the likelihood functions. nlm was executed with a wide range of initial values. nlm always converged and converged to a unique maximum.

Many of the fitted distributions are not nested. Discrimination among them was performed using:

- the Akaike information criterion due to Akaike (1974) defined by

$$\mathrm{AIC} = 2k - 2 \ln L\left(\widehat{\boldsymbol{\Theta}}\right),$$

  where $L(.)$ is the likelihood function and $\boldsymbol{\Theta}$ is the vector of all unknown parameters and k is the number of all unknown parameters;

- the Bayesian information criterion due to Schwarz (1978) defined by

$$\mathrm{BIC} = k \ln n - 2 \ln L\left(\widehat{\boldsymbol{\Theta}}\right),$$

  where $L(.)$ is the likelihood function and n is the sample size.

The smaller the values of these criteria the better the fit. For more discussion on these criteria, see Burnham and Anderson (2004) and Fang (2011).

The parameter estimates and the values of $-\ln L$, AIC, BIC are shown in Table 5.1 for the football data. The parameter estimates and the values of $-\ln L$, AIC, BIC are shown in Table 5.2 for the transformed football data. For the football data, the distinct values of (X, Y) are $(1, 2)$, $(0, 0)$, $(1, 1)$, $(0, 1)$, $(3, 2)$, $(3, 3)$, $(1, 3)$, $(1, 0)$ and $(3, 0)$. The observed number of observations corresponding to these values and the expected number under the fitted models in Table 5.1 are shown in Table 5.3. For the transformed football data, the distinct observations are $(2, 1)$, $(3, 3)$, $(2, 2)$, $(3, 2)$, $(0, 1)$, $(0, 0)$, $(2, 0)$, $(2, 3)$ and $(0, 3)$. The observed number of observations corresponding to these values and the expected number under the fitted models in Table 5.2 are shown in Table 5.4. Tables 5.3 and 5.4 also compute the chisquare goodness of fit statistics.

For the football data, Model 3 gives the smallest values for $-\ln L$, AIC and BIC. Model 1 gives the second smallest values for $-\ln L$, AIC and BIC. Model D gives the third smallest values for $-\ln L$, AIC and BIC. Model 4 gives the largest values for $-\ln L$, AIC and BIC. The smallest, the second smallest and the third smallest chisquare goodness of fit statistic value are given by Models 3, C and D. The largest chisquare goodness of fit statistic value is given by Model 4.

So, we can say that Model 3 is the best fitting model for the football data. It gives the smallest values for $-\ln L$, AIC, BIC and the chisquare goodness of fit statistic. Model D is possibly the second best fitting model in terms of the chisquare goodness of fit statistic.

For the transformed football data, Model C gives the smallest values for $-\ln L$, AIC and BIC. Model D gives the second smallest values for $-\ln L$, AIC and BIC. Model B gives the third smallest values for $-\ln L$, AIC and BIC. Model 1 gives the fourth smallest values for $-\ln L$, AIC and BIC. Model 3 gives the fifth smallest values for $-\ln L$, AIC and BIC. But the numeric values of $-\ln L$, AIC and BIC are very close among above five models. Model 4 gives the largest values for $-\ln L$, AIC and BIC. The smallest, the second smallest and the third smallest chisquare goodness of fit statistic values are given by Models 3, C and D. The largest chisquare goodness of fit statistic value is given by Model 4.

So, we can say that Model C is the best fitting model for the transformed football data in terms of $-\ln L$, AIC and BIC. Model 3 is the best fitting model in terms of chisquare goodness of fit statistic.

Models 2, 5 and 7 could not be fitted to either of the two data sets. Model 2 could not be fitted because it does not allow for observations, where $X < Y$. Model 5 could not be fitted because it does not allow for observations, where $X > 0$, $Y > 0$ and $X/(Y+k)$ is an integer for at least one $k = 1, 2, \ldots$. Model 7 could not be fitted because it does not allow for observations, where $X > Y$.

We should note that Model 3 is more tractable than Models C and D due to Lee and Cha (2015). For instance, the joint pmf X and Y, the marginal pmf of X, the marginal pmf of Y, the mean of X, the mean of Y and the mean of XY for Model 3 are

$$\begin{aligned}\Pr(X = x, Y = y) &= \exp\left(-\lambda_1 - \lambda_2 - \lambda_3\right) \frac{\lambda_1^{x-y}}{(x-y)!} \frac{\lambda_3^y}{y!} \sum_{k=0}^{y} \frac{\lambda_2^k}{k!} \\ &\quad + \exp\left(-\lambda_1 - \lambda_2 - \lambda_3\right) \frac{\lambda_2^y}{y!} \sum_{k=0}^{\min(x, y-1)} \frac{\lambda_1^{x-k}}{(x-k)!} \frac{\lambda_3^k}{k!},\end{aligned}$$

$$\Pr(X = x) = \frac{\exp\left(-\lambda_1 - \lambda_3\right) \left(\lambda_1 + \lambda_3\right)^x}{x!},$$

Model	$\widehat{\lambda_1}$	$\widehat{\lambda_2}$	$\widehat{\lambda_3}$	$-\ln L$	AIC	BIC
A	1.077	1.385		67.604	139.208	141.724
B	0.381	0.688	0.696	64.916	135.832	139.606
C	1.357	2.095	2.271	64.221	134.443	138.217
D	0.671	1.119	0.653	62.662	131.323	135.098
1	0.064	1.312	1.269	55.152	116.304	120.078
3	0.914	0.000	0.650	37.532	81.065	84.839
4	0.929	1.295	1.150	87.279	180.558	184.332
6	1.017	1.159	0.986	67.400	140.800	144.574

Table 5.1: Parameter estimates, log-likelihood values, AIC values and BIC values for the distributions fitted to the football data.

Model	$\widehat{\lambda_1}$	$\widehat{\lambda_2}$	$\widehat{\lambda_3}$	$-\ln L$	AIC	BIC
A	1.923	1.615		75.389	154.777	157.294
B	0.668	0.360	1.255	70.466	146.933	150.707
C	2.918	2.067	2.713	69.020	144.039	147.813
D	1.625	1.051	1.100	69.735	145.469	149.244
1	0.764	1.306	1.320	70.995	147.991	151.765
3	0.555	0.937	1.368	71.003	148.006	151.780
4	1.866	1.520	1.151	88.750	183.501	187.275
6	1.855	1.240	1.202	76.279	158.557	162.332

Table 5.2: Parameter estimates, log-likelihood values, AIC values and BIC values for the distributions fitted to the transformed football data.

(x, y)	Obs	Expected frequencies for the fitted models							
		A	B	C	D	1	3	4	6
(1, 2)	6	2.3	2.5	2.5	2.5	2.0	2.3	0.8	0.7
(0, 0)	1	2.2	4.5	3.4	2.3	6.9	5.4	10.2	3.0
(1, 1)	8	3.3	4.3	5.5	6.9	3.1	5.9	1.2	2.4
(0, 1)	5	3.1	3.1	2.5	2.5	NA	3.1	1.3	5.9
(3, 2)	1	0.4	0.6	0.4	0.2	0.1	0.8	0.1	0.1
(3, 3)	2	0.2	0.6	0.4	0.7	0.8	0.6	0.3	0.6
(1, 3)	1	1.1	0.8	1.0	0.9	0.9	1.7	0.3	0.3
(1, 0)	1	2.4	1.7	1.3	1.5	2.8	5.0	1.0	NA
(3, 0)	1	0.5	0.0	0.2	0.1	0.7	0.7	0.3	NA
χ^2		32.5	38.4	22.3	21.6	29.6	18.4	107.6	63.0

Table 5.3: Observed frequencies, expected frequencies and chisquare goodness of fit statistics for the distributions fitted to the football data.

$$\begin{aligned}\Pr(Y=y) &= \exp\left(-\lambda_1-\lambda_2-\lambda_3\right)\sum_{x=y}^{\infty}\frac{\lambda_1^{x-y}}{(x-y)!}\frac{\lambda_3^{y}}{y!}\sum_{k=0}^{y}\frac{\lambda_2^{k}}{k!}\\ &\quad+\exp\left(-\lambda_1-\lambda_2-\lambda_3\right)\frac{\lambda_2^{y}}{y!}\sum_{x=0}^{\infty}\sum_{k=0}^{\min(x,y-1)}\frac{\lambda_1^{x-k}}{(x-k)!}\frac{\lambda_3^{k}}{k!},\end{aligned}$$

$$\begin{aligned}E(XY) &= \exp\left(-\lambda_2-\lambda_3\right)\sum_{y=0}^{\infty}\left(\lambda_1 y+y^2\right)\frac{\lambda_3^{y}}{y!}\sum_{k=0}^{y}\frac{\lambda_2^{k}}{k!}\\ &\quad+\exp\left(-\lambda_1-\lambda_2-\lambda_3\right)\sum_{y=0}^{\infty}\frac{\lambda_2^{y}}{(y-1)!}\sum_{x=0}^{\infty}\sum_{k=0}^{\min(x,y-1)}\frac{x\lambda_1^{x-k}}{(x-k)!}\frac{\lambda_3^{k}}{k!},\end{aligned}$$

$$E(X)=\lambda_1+\lambda_3,$$

and

$$\begin{aligned}E(Y) &= \exp\left(-\lambda_1-\lambda_2-\lambda_3\right)\sum_{y=0}^{\infty}\sum_{x=y}^{\infty}\frac{\lambda_1^{x-y}}{(x-y)!}\frac{\lambda_3^{y}}{(y-1)!}\sum_{k=0}^{y}\frac{\lambda_2^{k}}{k!}\\ &\quad+\exp\left(-\lambda_1-\lambda_2-\lambda_3\right)\sum_{y=0}^{\infty}\frac{\lambda_2^{y}}{(y-1)!}\sum_{x=0}^{\infty}\sum_{k=0}^{\min(x,y-1)}\frac{\lambda_1^{x-k}}{(x-k)!}\frac{\lambda_3^{k}}{k!},\end{aligned}$$

respectively. These expressions can be used for example to compute the correlation coefficient between X and Y for Model 3. Figures 5.1, 5.2 and 5.3 show possible shapes of the contours of the correlation coefficient. We can observe the following from the figures: the correlation increases as both λ_1 and λ_2 decrease; the correlation decreases as both λ_1 and λ_2 increase; the correlation increases as λ_1 decreases and λ_3 increases; the correlation decreases as λ_1 increases and λ_3 decreases; the correlation increases as λ_2 decreases and λ_3 increases; the correlation decreases as λ_2 increases and λ_3 decreases.

Finally, we note that none of the best fitting models (including Models C and D due to Lee and Cha (2015) and the proposed Model 3) provide adequate fits. The chisquare goodness of fit rejects the null hypothesis at the five percent level. But Model 3 gives the smallest chisquare goodness of fit statistic value for both the football and transformed football datasets.

The estimated mean of X, mean of Y, standard deviation of X, variance of X, standard deviation of Y, variance of Y and the correlation coefficient between X and Y under Model 3 are 1.564, 0.650, 1.251, 1.564, 0.806, 0.650 and 0.645, respectively. These figures do not differ too greatly from the sample values. A future work is to develop discrete bivariate distributions that may provide adequate fits.

(x, y)	Obs	Expected frequencies for the fitted models							
		A	B	C	D	1	3	4	6
(2, 1)	6	2.3	2.4	2.1	2.6	1.2	3.4	0.8	1.6
(3, 3)	1	0.6	1.4	1.4	1.9	0.6	1.8	0.4	0.8
(2, 2)	8	1.8	2.9	3.9	4.6	1.8	3.6	1.2	2.0
(3, 2)	5	1.2	1.7	1.3	1.2	1.0	2.1	0.4	0.3
(0, 1)	1	1.2	1.0	0.5	0.6	NA	1.2	0.5	3.8
(0, 0)	2	0.8	2.7	1.9	0.6	3.2	3.7	8.8	2.3
(2, 0)	1	1.4	0.6	1.0	0.8	2.6	2.0	0.9	NA
(2, 3)	1	1.0	0.9	0.8	0.6	0.3	0.4	0.3	0.4
(0, 3)	1	0.5	0.0	0.2	0.1	NA	0.2	0.3	1.3
χ^2		42.5	67.7	26.2	30.7	60.7	17.3	133.9	95.1

Table 5.4: Observed frequencies, expected frequencies and chisquare goodness of fit statistics for the distributions fitted to the transformed football data.

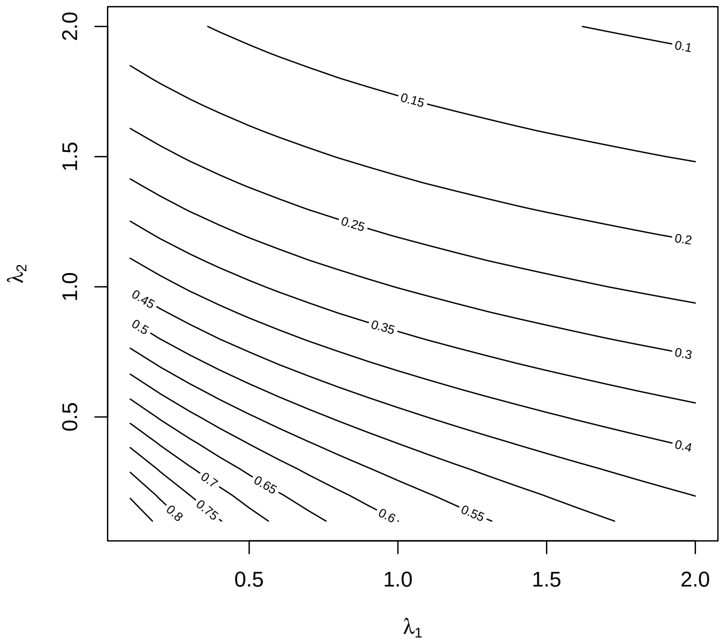

Figure 5.1: Contours of the correlation coefficient between X and Y versus λ_1 and λ_2 for Model 3 and $\lambda_3 = 0.650$.

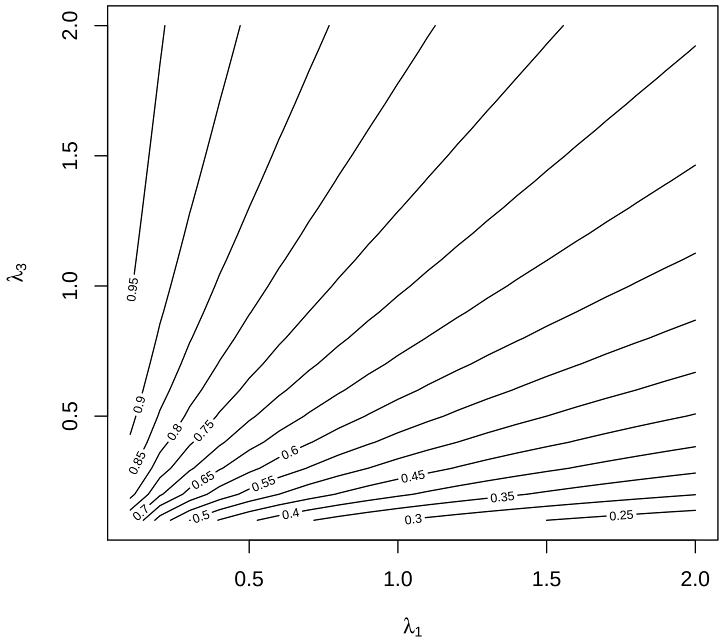

Figure 5.2: Contours of the correlation coefficient between X and Y versus λ_1 and λ_3 for Model 3 and $\lambda_2 = 4.856 \times 10^{-7}$.

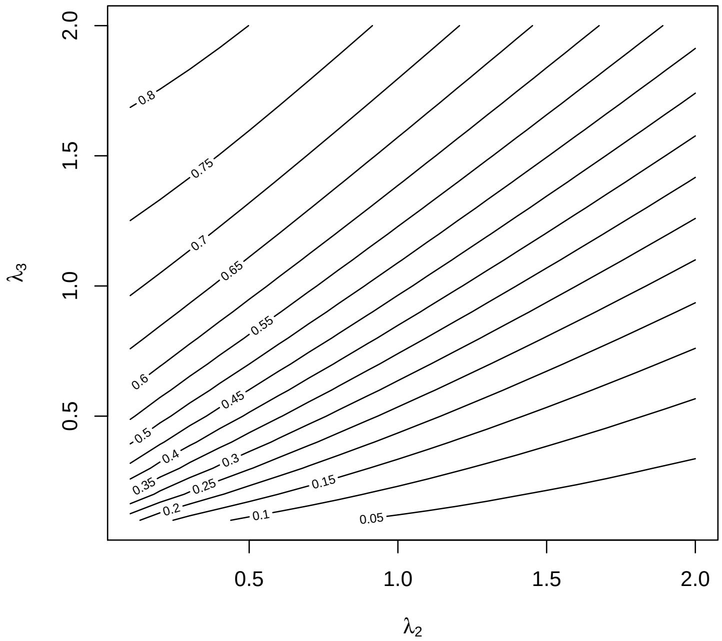

Figure 5.3: Contours of the correlation coefficient between X and Y versus λ_2 and λ_3 for Model 3 and $\lambda_1 = 0.914$.

Chapter 6

Amplitude and phase distributions

6.1 Introduction

Let (X, Y) be a random vector defined on the real space $(-\infty, +\infty) \times (-\infty, +\infty)$. Let $R = \sqrt{X^2 + Y^2}$ and $\Theta = \arctan(Y/X)$. The distributions of R and Θ arise in many areas of the IEEE literature: radar communications, positioning related applications, fading channels, etc. R is usually referred to as the amplitude and Θ the phase. Some published examples in the IEEE literature where the distributions of R and Θ arise directly are: analog data-transmission (Benedetto and Steila, 1977); radiation-pattern of an offset fed paraboloidal reflector antenna (Herben and Maanders, 1979); microwave pulse generation (Brummer et al., 2011); photonic beamforming based on programmable phase shifters (Yi et al., 2011); radiation characteristics of multimode concentric circular microstrip patch antennas (Tran and Sharma, 2012); transmission schemes for Rayleigh block fading channels (Chen and Ueng, 2013); broadband microwave photonic splitters (Li et al., 2014); performance of transmit beamforming codebooks (Dowhuszko and Hamalainen, 2015); AC-AC converters (Zhang and Ruan, 2015); modulation of ultrawideband monocycle pulses on a silicon photonic chip (Xu et al., 2016).

Several papers have derived the distributions of R and Θ: Aalo et al. (2007) and Dharmawansa et al. (2009) derived the distributions of R and Θ when X and Y are correlated and non-identical normal random variables with non-zero means; Yacoub (2010) derived the distributions of R and Θ when X and Y are independent and identical Nakagami type random variables; Coluccia (2013) derived the moments of R when X and Y are uncorrelated normal random variables with zero means and unequal variances; and so on.

Almost all of the known papers have supposed X and Y are normally distributed. However, non-normal distributions are becoming increasingly popular in the IEEE literature: hyperbolic distribution in network modeling (Li and Manikopoulos, 2004); Laplace distribution in signal processing (Eltoft et al., 2006); Cauchy distribution in segmentation of noisy colour images (Wan et al., 2011); Kotz type distribution for multilook polarimetric radar data (Kersten and Anfinsen, 2012); Student's t distribution in medical image segmentation (Nguyen and Wu, 2012); logistic distribution for networked video quality assessment (Zhang et al., 2013); Gumbel distribution for peak sidelobe level for arrays of randomly placed antennas (Krishnamurthy et al., 2015); skew normal distribution for statistical static timing analysis (Vijaykumar and Vasudevan, 2016); and so on. Also, we are not aware of any paper giving expressions for moments of R when X and Y are correlated random variables.

The aim of this chapter is to derive the distribution of R, its moments and the distribution of Θ for a wide range of bivariate distributions, including the correlated bivariate normal distribution. We consider thirty four flexible bivariate distributions in total. These include eleven bivariate normal distributions, eight bivariate t distributions, five bivariate Laplace distributions, two bivariate hyperbolic distributions, two bivariate Gumbel distributions, one bivariate logistic distribution and five other bivariate distributions.

The contents of this chapter are organized as follows. Section 6.2 derives the distribution of R, its moments and the distribution of Θ when X and Y are correlated normal random variables with zero means and unequal variances. These results extend those given in Coluccia (2013). Section 6.3 derives the same when X and Y are correlated random variables following thirty three other bivariate distributions. Details of the derivations are not given. They can be obtained from the corresponding author. To the best of our knowledge, the derived expressions for the distribution of R, its moments and the distribution of Θ are all new and original. Section 6.4 discusses simulation of R and Θ for the bivariate distributions considered.

It is hoped that the details given in Section 6.3 could be a useful reference for the IEEE community. They could also encourage researchers to apply more non-normal distributions to real problems in the EEE. A future work is to extend the results in Section 6.3 for multivariate distributions, that is to derive the distributions of $\sqrt{X_1^2 + X_2^2 + \cdots + X_p^2}$ and $\left(X_1/\sqrt{X_1^2 + X_2^2 + \cdots + X_p^2}, X_2/\sqrt{X_1^2 + X_2^2 + \cdots + X_p^2}, \ldots, X_{p-1}/\sqrt{X_1^2 + X_2^2 + \cdots + X_p^2} \right)$ given a distribution for $(X_1, X_2, \ldots, X_p)$, $p > 2$.

The expressions given in Sections 6.2 and 6.3 involve various special functions, including the gamma function defined by

$$\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t) dt$$

for $a > 0$; the incomplete gamma function defined by

$$\gamma(a, x) = \int_0^x t^{a-1} \exp(-t) dt$$

for $a > 0$ and $x > 0$; the beta function defined by

$$B(a, b) = \int_0^\infty t^{a-1} (1-t)^{b-1} dt$$

for $a > 0$ and $b > 0$; the error function defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt$$

for $x > 0$; the complementary error function defined by

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp\left(-t^2\right) dt$$

for $-\infty < x < +\infty$; the parabolic cylinder function defined by

$$D_\nu(x) = \frac{\exp\left(-x^2/4\right)}{\Gamma(-\nu/2)} \int_0^\infty t^{-\frac{\nu}{2}-1} (1+2t)^{\frac{\nu-1}{2}} \exp\left(-x^2 t\right) dt$$

for $\nu < 0$ and $x^2 > 0$; the Euler number of order n defined by

$$E_{2n} = (-1)^n 2^{2n+1} \int_0^\infty t^{2n} \text{sech}(\pi t) dt;$$

the modified Bessel function of the first kind of order ν defined by

$$I_\nu(x) = \sum_{k=0}^\infty \frac{1}{\Gamma(k+\nu+1) k!} \left(\frac{x}{2}\right)^{2k+\nu};$$

the modified Bessel function of the second kind of order ν defined by

$$K_\nu(x) = \begin{cases} \dfrac{\pi \csc(\pi\nu)}{2}\left[I_{-\nu}(x) - I_\nu(x)\right], & \text{if } \nu \notin \mathbb{Z}, \\ \\ \lim\limits_{\mu\to\nu} K_\mu(x), & \text{if } \nu \in \mathbb{Z}; \end{cases}$$

the confluent hypergeometric function defined by

$${}_1F_1(\alpha;\beta;z) = \sum_{k=0}^{\infty} \frac{(\alpha)_k}{(\beta)_k} \frac{z^k}{k!},$$

where $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$ denotes the ascending factorial; the Gauss hypergeometric function defined by

$${}_2F_1(\alpha,\beta;\gamma;z) = \sum_{k=0}^{\infty} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k} \frac{z^k}{k!};$$

the ${}_2F_2$ hypergeometric function defined by

$${}_2F_2(\alpha,\beta;\gamma,\delta;z) = \sum_{k=0}^{\infty} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k(\delta)_k} \frac{z^k}{k!};$$

the ${}_4F_2$ hypergeometric function defined by

$${}_4F_2(a,b,c,d;e,f;z) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k(c)_k(d)_k}{(e)_k(f)_k} \frac{z^k}{k!};$$

the ${}_4F_3$ hypergeometric function defined by

$${}_4F_3(a,b,c,d;e,f,g;z) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k(c)_k(d)_k}{(e)_k(f)_k(g)_k} \frac{z^k}{k!};$$

and, the Appell function of the first kind defined by

$$F_1(a,b,c,d;z,\xi) = \sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty} \frac{(a)_{k+\ell}(b)_k\,(c)_\ell\, z^k \xi^\ell}{(d)_{k+\ell} k! \ell!}.$$

These special functions are well known and well established in the mathematics literature. Some details of their properties can be found in Prudnikov et al. (1986) and Gradshteyn and Ryzhik (2014). In-built routines for computing them are widely available in packages like Maple, Matlab and Mathematica. For example, the in-built routines in Mathematica for the stated special functions are: GAMMA[a] for the gamma function, GAMMA[a]-GAMMA[a,x] for the incomplete gamma function, Beta[a,b] for the beta function, Erc[x] for the error function, Erfc[x] for the complementary error function, ParabolicCylinderD[nu,x] for the parabolic cylinder function, EulerE[n] for the Euler number of order n, BesselI[nu,x] for the modified Bessel function of the first kind of order ν, BesselK[nu,x] for the modified Bessel function of the second kind of order ν, Hypergeometric1F1[alpha,beta,z] for the confluent hypergeometric function, Hypergeometric2F1[alpha,beta,gamma,z] for the Gauss hypergeometric function, HypergeometricPFQ [{alpha,beta}, {gamma,delta}, z] for the ${}_2F_2$ hypergeometric function, HypergeometricPFQ [{a,b,c,d}, {e,f}, z] for the ${}_4F_2$ hypergeometric function, HypergeometricPFQ [{a,b,c,d}, {e,f,g}, z] for the ${}_4F_3$ hypergeometric function, AppellF1[a,b,c,d,z,xi] for the Appell function of the first kind. Mathematica like other algebraic manipulation packages allows for arbitrary precision, so the accuracy of computations is not an issue.

In-built routines for most of the stated special functions are also available in the freely available R software (R Development Core Team, 2016): gamma(a) in the base package for the gamma function, gamma(a)*pgamma(x,shape=a) in the base package for the incomplete gamma function, beta(a,b) in the base package for the beta function, erf(x) in the contributed package NORMT3 for the error function, erfc(x) in the contributed package NORMT3 for the complementary error function, besselI(x,nu) in the base package for the modified Bessel function of the first kind of order ν, besselK(x,nu) in the base package for the modified Bessel function of the second kind of order ν, kummerM(z,alpha,beta) in the contributed package AsianOptions for the confluent hypergeometric function, hypergeo(alpha,beta,gamma,z) in the contributed package hypergeo for the Gauss hypergeometric function, genhypergeo(U=c(alpha,beta), L=c(gamma,beta), z) in the contributed package hypergeo for the ${}_2F_2$ hypergeometric function, genhypergeo(U=c(a,b,c,d), L=c(e,f), z) in the contributed package hypergeo for the ${}_4F_2$ hypergeometric function, genhypergeo(U=c(a,b,c,d), L=c(e,f,g), z) in the contributed package hypergeo for the ${}_4F_3$ hypergeometric function, F1(a,b,c,d,z,xi) in the contributed package tolerance for the Appell function of the first kind.

6.2 Bivariate normal case

Here, we derive the pdf of R, the pdf of Θ and $E\left(R^p\right)$ when (X,Y) has the bivariate normal pdf

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_x\sigma_y} \exp\left[-\left(\frac{x}{\sigma_x}\right)^2 - \left(\frac{y}{\sigma_y}\right)^2 - 2\rho\frac{x}{\sigma_x}\frac{y}{\sigma_y}\right] \tag{6.1}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\sigma_x > 0$, $\sigma_y > 0$ and $-1 < \rho < 1$. We can write (6.1) in the form

$$f(x, y) = C \exp\left(-ax^2 - by^2 - 2cxy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$ and $-\infty < c < +\infty$, where C denotes the normalizing constant. The corresponding joint pdf of R and Θ can be expressed as

$$\begin{aligned} f(r, \theta) &= Cr \exp\left[-r^2\left(a \sin^2\theta + b\cos^2\theta + 2c\sin\theta\cos\theta\right)\right] \\ &= Cr \exp\left\{-r^2\left[a + (b-a)\cos^2\theta + c\sin(2\theta)\right]\right\} \\ &= Cr \exp\left\{-r^2\left(a + \frac{b-a}{2}\left[1 + \cos(2\theta)\right] + c\sin(2\theta)\right)\right\} \\ &= Cr \exp\left\{-r^2\frac{b+a}{2} - r^2\frac{b-a}{2}\cos(2\theta) - cr^2\sin(2\theta)\right\} \end{aligned} \tag{6.2}$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$. Using the fact

$$\int_0^{2\pi} \exp\left(x\cos\theta + y\sin\theta\right) d\theta = 2\pi I_0\left(\sqrt{x^2 + y^2}\right),$$

we obtain the pdf of R as

$$f_R(r) = 2\pi Cr \exp\left(-\alpha r^2\right) I_0\left(r^2\beta\right)$$

for $r > 0$, where $\alpha = \frac{a+b}{2}$, $\beta = \sqrt{\gamma^2 + c^2}$ and $\gamma = \frac{b-a}{2}$. A straightforward integration of (6.2) shows that the pdf of Θ is

$$f_\Theta(\theta) = C\left[a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)\right]^{-1}$$

for $0 < \theta < 2\pi$. An application of equation (2.15.3.2) in Prudnikov et al. (1986, volume 2) shows that the pth moment of R can be expressed as

$$E\left(R^p\right) = C\pi\alpha^{-\frac{p}{2}-1}\Gamma\left(\frac{p}{2} + 1\right) {}_2F_1\left(\frac{p}{4} + 1, \frac{p}{4} + \frac{1}{2}; 1; \frac{\beta^2}{\alpha^2}\right) \tag{6.3}$$

for $p > 0$. By the transformation formulas of the Gauss hypergeometric function, three equivalent representation of (6.3) are

$$E\left(R^p\right) = C\pi\alpha^{\frac{p}{2}}\left(\alpha^2 - \beta^2\right)^{-\frac{1}{2}-\frac{p}{2}}\Gamma\left(\frac{p}{2} + 1\right) {}_2F_1\left(-\frac{p}{4}, -\frac{p}{4} + \frac{1}{2}; 1; \frac{\beta^2}{\alpha^2}\right), \tag{6.4}$$

$$E\left(R^p\right) = C\pi\alpha\left(\alpha^2-\beta^2\right)^{-1-\frac{p}{4}}\Gamma\left(\frac{p}{2}+1\right){}_2F_1\left(1+\frac{p}{4},-\frac{p}{4}+\frac{1}{2};1;\frac{\beta^2}{\beta^2-\alpha^2}\right) \tag{6.5}$$

and

$$E\left(R^p\right) = C\pi\left(\alpha^2-\beta^2\right)^{-\frac{1}{2}-\frac{p}{4}}\Gamma\left(\frac{p}{2}+1\right){}_2F_1\left(-\frac{p}{4},\frac{p}{4}+\frac{1}{2};1;\frac{\beta^2}{\beta^2-\alpha^2}\right) \tag{6.6}$$

for $p > 0$. If p is a positive integer and a multiple of 4 then (6.4) and (6.6) reduce to the elementary forms

$$E\left(R^p\right) = C\pi\alpha^{\frac{p}{2}}\left(\alpha^2-\beta^2\right)^{-\frac{1}{2}-\frac{p}{2}}\Gamma\left(\frac{p}{2}+1\right)\sum_{k=0}^{p/4}\frac{\left(-\frac{p}{4}\right)_k\left(-\frac{p}{4}+\frac{1}{2}\right)_k}{k!^2}\left(\frac{\beta^2}{\alpha^2}\right)^k$$

and

$$E\left(R^p\right) = C\pi\left(\alpha^2-\beta^2\right)^{-\frac{1}{2}-\frac{p}{4}}\Gamma\left(\frac{p}{2}+1\right)\sum_{k=0}^{p/4}\frac{\left(-\frac{p}{4}\right)_k\left(\frac{p}{4}+\frac{1}{2}\right)_k}{k!^2}\left(\frac{\beta^2}{\beta^2-\alpha^2}\right)^k,$$

respectively.

6.3 The collection

Here, we tabulate expressions for the pdf of R, the pdf of Θ and $E\left(R^p\right)$ when (X,Y) follows thirty four flexible bivariate distributions. We also select the representative models and provide the details of derivation in the Appendix B. Note that, for some distributions, the derivation of $E\left(R^p\right)$ was not possible.

Bivariate normal distribution (Balakrishnan and Lai, 2009, Chapter 11) has the joint pdf specified by

$$f(x,y) = C\exp\left(-ax^2-by^2-2cxy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$ and $-\infty < c < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\exp\left[-\alpha r^2-\frac{b-a}{2}\cos(2\theta)r^2-c\sin(2\theta)r^2\right],$$

$$f_R(r) = 2\pi C r \exp\left(-\alpha r^2\right) I_0\left(r^2 \beta\right),$$

$$f_\Theta(\theta) = C\left[a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)\right]^{-1}$$

and

$$E\left(R^p\right) = C\pi\alpha^{-\frac{p}{2}-1}\Gamma\left(\frac{p}{2}+1\right) {}_2F_1\left(\frac{p}{4}+1, \frac{p}{4}+\frac{1}{2}; 1; \frac{\beta^2}{\alpha^2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$, where $\alpha = \frac{a+b}{2}$, $\beta = \sqrt{\gamma^2 + c^2}$ and $\gamma = \frac{b-a}{2}$.

Bivariate normal distribution with non-zero means (Balakrishnan and Lai, 2009, Chapter 11) has the joint pdf specified by

$$f(x, y) = C\exp\left(-ax^2 - by^2 - 2cxy - dx - ey\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$, $-\infty < c < +\infty$, $d > 0$ and $e > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = Cr\exp\left[-\alpha r^2 - \gamma\cos(2\theta)r^2 - c\sin(2\theta)r^2 - dr\sin\theta - er\cos\theta\right],$$

$$f_R(r) = Cr\exp\left(-\alpha r^2\right)\int_0^{2\pi}\exp\left[-\gamma\cos(2\theta)r^2 - c\sin(2\theta)r^2 - dr\sin\theta - er\cos\theta\right]d\theta$$

and

$$\begin{aligned} f_\Theta(\theta) &= C\left[a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)\right]^{-1} \\ &\cdot \exp\left\{\frac{(d\sin\theta + e\cos\theta)^2}{4\left[a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)\right]}\right\} \\ &\cdot D_{-2}\left(\frac{d\sin\theta + e\cos\theta}{\sqrt{a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)}}\right) \end{aligned}$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$, where $\alpha = \frac{a+b}{2}$ and $\gamma = \frac{b-a}{2}$.

Conditionally specified bivariate normal distribution (Balakrishnan and Lai, 2009, equation (6.8)) has the joint pdf specified by

$$f(x, y) = C\exp\left(-\frac{x^2 + y^2 + cx^2y^2}{2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $-\infty < c < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = Cr \exp\left[-\frac{r^2 + \left(cr^4/4\right)\sin^2(2\theta)}{2}\right],$$

$$f_R(r) = 2Cr \exp\left(-\frac{r^2}{2}\right) \sum_{i=0}^{\infty} \frac{\left(-cr^4\right)^i}{8^i i!} B\left(\frac{1}{2}, i + \frac{1}{2}\right),$$

$$f_\Theta(\theta) = \frac{C}{\sqrt{c}\,|\sin(2\theta)|} \exp\left[\frac{1}{4c\sin^2(2\theta)}\right] D_{-1}\left(\frac{1}{\sqrt{c}\,|\sin(2\theta)|}\right)$$

and

$$E\left(R^p\right) = 2^{1+\frac{p}{2}} C \sum_{i=0}^{\infty} \frac{(-c)^i}{2^i i!} B\left(\frac{1}{2}, i + \frac{1}{2}\right) \Gamma\left(2i + \frac{p}{2} + 1\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Bivariate skew normal distribution (Balakrishnan and Lai, 2009, equation (7.17)) has the joint pdf specified by

$$f(x, y) = 2\phi(x)\phi(y)\Phi\left(\alpha x + \beta y\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < \alpha < +\infty$ and $-\infty < \beta < +\infty$, where $\phi(\cdot)$ and $\Phi(\cdot)$ denote, respectively, the pdf and cdf of a standard normal random variable. For this distribution,

$$f(r, \theta) = \frac{r}{\pi} \exp\left(-\frac{r^2}{2}\right) \Phi\left(\alpha r \sin\theta + \beta r \cos\theta\right),$$

$$f_R(r) = r \exp\left(-\frac{r^2}{2}\right),$$

$$f_\Theta(\theta) = \frac{1}{2\pi} + \frac{\alpha \sin\theta + \beta\cos\theta}{2\pi\sqrt{1 + \left(\alpha\sin\theta + \beta\cos\theta\right)^2}}$$

and

$$E\left(R^p\right) = 2^{\frac{p}{2}}\Gamma\left(\frac{p}{2}+1\right)$$

for $r > 0$, $0 \le \theta \le 2\pi$ and $p > 0$. The given expression for $f_\Theta(\theta)$ follows from $\Phi(x) = \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{x}{\sqrt{2}}\right)$ and by writing

$$\begin{aligned}
f_\Theta(\theta) &= \int_0^\infty \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\Phi\left(\alpha r\sin\theta + \beta r\cos\theta\right)dr \\
&= \int_0^\infty \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\left[\frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{\alpha r\sin\theta + \beta r\cos\theta}{\sqrt{2}}\right)\right]dr \\
&= \frac{1}{2\pi}\int_0^\infty r\exp\left(-\frac{r^2}{2}\right)dr + \frac{1}{2\pi}\int_0^\infty r\exp\left(-\frac{r^2}{2}\right)\text{erf}\left(\frac{\alpha r\sin\theta + \beta r\cos\theta}{\sqrt{2}}\right)dr \\
&= \frac{1}{2\pi} + \frac{1}{2\pi}\int_0^\infty r\exp\left(-\frac{r^2}{2}\right)\text{erf}\left(\frac{\alpha r\sin\theta + \beta r\cos\theta}{\sqrt{2}}\right)dr
\end{aligned}$$

and applying equation (2.8.5.9) in Prudnikov et al. (1986, volume 2) to calculate the integral. The derivation for $f_R(r)$ can be written as

$$f_R(r) = \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\int_0^{2\pi}\Phi\left(\alpha r\sin\theta + \beta r\cos\theta\right)d\theta$$

Note that $\alpha r\sin\theta + \beta r\cos\theta \equiv Rsin(\theta+\phi)$, where $R = r\cdot\sqrt{\alpha^2+\beta^2}$ and $\phi = \tan^{-1}\frac{\beta}{\alpha}$.

$$f_R(r) = \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\int_0^{2\pi}\Phi\left(R\sin(\theta+\phi)\right)d\theta.$$

As this is a integral with repect to θ from 0 to 2π, the element within $\Phi(.)$, $R\sin(\theta+\phi)$, can be simplified as $R\sin(\theta)$.

$$\begin{aligned}
f_R(r) &= \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\int_0^{2\pi}\Phi\left(R\sin\theta\right)d\theta \\
&= \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\left[\int_0^{\pi}\Phi\left(R\sin\theta\right)d\theta + \int_\pi^{2\pi}\Phi\left(R\sin\theta\right)d\theta\right] \\
&= \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\left[\int_0^{\pi}\Phi\left(R\sin\theta\right)d\theta + \int_0^{\pi}\Phi\left(-R\sin\theta\right)d\theta\right]
\end{aligned}$$

The last equation is due to $\sin(\theta - \pi) = -\sin\theta$. Note the fact that $\Phi(-x) = 1 - \Phi(x)$.

$$\begin{aligned} f_R(r) &= \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\left[\int_0^\pi \Phi\left(R\sin\theta\right)d\theta + \int_0^\pi 1 - \Phi\left(R\sin\theta\right)d\theta\right] \\ &= \frac{r}{\pi}\exp\left(-\frac{r^2}{2}\right)\int_0^\pi 1 d\theta \\ &= r\exp\left(-\frac{r^2}{2}\right). \end{aligned}$$

Bivariate skew normal distribution (Arnold and Beaver, 2002, equation (6.5)) has the joint pdf specified by

$$f(x,y) = C\exp\left(-\frac{x^2+y^2}{2}\right)\Phi\left(ax+by+cxy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$, $-\infty < b < +\infty$ and $-\infty < c < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\exp\left(-\frac{r^2}{2}\right)\Phi\left(ar\sin\theta + br\cos\theta + \frac{cr^2}{2}\sin(2\theta)\right),$$

$$f_R(r) = Cr\exp\left(-\frac{r^2}{2}\right)\int_0^{2\pi}\Phi\left(ar\sin\theta + br\cos\theta + \frac{cr^2}{2}\sin(2\theta)\right)d\theta$$

and

$$f_\Theta(\theta) = C\int_0^\infty r\exp\left(-\frac{r^2}{2}\right)\Phi\left(ar\sin\theta + br\cos\theta + \frac{cr^2}{2}\sin(2\theta)\right)dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate skew normal distribution (Arnold and Beaver, 2002, equation (4.11)) has the joint pdf specified by

$$f(x,y) = C\exp\left(-\frac{x^2+y^2}{2}\right)\Phi\left(ax+by+c\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$, $-\infty < b < +\infty$ and $-\infty < c < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\exp\left(-\frac{r^2}{2}\right)\Phi\left(ar\sin\theta + br\cos\theta + c\right),$$

$$f_R(r) = Cr \exp\left(-\frac{r^2}{2}\right) \int_0^{2\pi} \Phi\left(ar\sin\theta + br\cos\theta + c\right) d\theta$$

and

$$\begin{aligned} f_\Theta(\theta) = & \frac{C}{2} + \frac{C}{2}\operatorname{erf}\left(\frac{c}{\sqrt{2}}\right) + \frac{C}{2}\frac{a\sin\theta + b\cos\theta}{\sqrt{(a\sin\theta + b\cos\theta)^2 + 1}} \\ & \cdot \exp\left[-\frac{c^2}{2(a\sin\theta + b\cos\theta)^2 + 2}\right] \operatorname{erfc}\left(\frac{c(a\sin\theta + b\cos\theta)}{\sqrt{2(a\sin\theta + b\cos\theta)^2 + 2}}\right) \end{aligned}$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate skew normal distribution (Arnold et al., 2002, equation (5.1)) has the joint pdf specified by

$$f(x, y) = C \exp\left(-\frac{x^2 + y^2}{2}\right) \Phi\left(a + bx + cy + dx^2 + ey^2 + fxy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$, $-\infty < b < +\infty$, $-\infty < c < +\infty$, $-\infty < d < +\infty$, $-\infty < e < +\infty$ and $-\infty < f < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = Cr \exp\left(-\frac{r^2}{2}\right) \Phi\left(a + br\sin\theta + cr\cos\theta + dr^2\sin^2\theta + er^2\cos^2\theta + \frac{fr^2}{2}\sin(2\theta)\right),$$

$$f_R(r) = Cr \exp\left(-\frac{r^2}{2}\right) \int_0^{2\pi} \Phi\left(a + br\sin\theta + cr\cos\theta + dr^2\sin^2\theta + er^2\cos^2\theta + \frac{fr^2}{2}\sin(2\theta)\right) d\theta$$

and

$$f_\Theta(\theta) = C \int_0^{\infty} r \exp\left(-\frac{r^2}{2}\right) \Phi\left(a + br\sin\theta + cr\cos\theta + dr^2\sin^2\theta + er^2\cos^2\theta + \frac{fr^2}{2}\sin(2\theta)\right) dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate skew normal distribution (Balakrishnan and Lai, 2009, page 525) has the joint pdf specified by

$$f(x, y) = C \exp\left(-\frac{x^2 + y^2}{2a^2}\right) \Phi(\alpha x) \Phi(\beta y)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $-\infty < \alpha < +\infty$ and $-\infty < \beta < +\infty$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr \exp\left(-\frac{r^2}{2a^2}\right) \Phi\left(\alpha r \sin\theta\right) \Phi\left(\beta r \cos\theta\right),$$

$$f_R(r) = \frac{C\pi r}{2} \exp\left(-\frac{r^2}{2a^2}\right),$$

$$\begin{aligned} f_\Theta(\theta) = & \frac{Ca^2}{4} + \frac{Ca^3\alpha \sin\theta}{4\sqrt{\alpha^2 a^2 \sin^2\theta + 1}} + \frac{Ca^3\beta\cos\theta}{4\sqrt{\beta^2 a^2 \cos^2\theta + 1}} \\ & + \frac{Ca^3\alpha\sin\theta}{2\pi\sqrt{\alpha^2 a^2 \sin^2\theta + 1}} \arctan \frac{\beta a \cos\theta}{\sqrt{\alpha^2 a^2 \sin^2\theta + 1}} \\ & + \frac{Ca^3\beta\cos\theta}{2\pi\sqrt{\beta^2 a^2 \cos^2\theta + 1}} \arctan \frac{\alpha a \sin\theta}{\sqrt{\beta^2 a^2 \cos^2\theta + 1}} \end{aligned}$$

and

$$E\left(R^p\right) = C\pi a^{p+2} 2^{\frac{p}{2}-1} \Gamma\left(\frac{p}{2} + 1\right)$$

for $r > 0$, $0 \le \theta \le 2\pi$ and $p > 0$.

Conditionally specified bivariate skew normal distribution (Balakrishnan and Lai, 2009, equation (6.78)) has the joint pdf specified by

$$f(x,y) = 2\phi(x)\phi(y)\Phi\left(\lambda xy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $-\infty < \lambda < +\infty$. For this distribution,

$$f(r,\theta) = \frac{r}{\pi} \exp\left(-\frac{r^2}{2}\right) \Phi\left(\frac{\lambda r^2}{2} \sin 2\theta\right),$$

$$f_R(r) = r \exp\left(-\frac{r^2}{2}\right),$$

$$f_\Theta(\theta) = \frac{1}{2\pi}\left\{1 + \exp\left[\frac{1}{4\lambda^2 \sin^2(2\theta)}\right] \operatorname{erfc}\left(\frac{1}{2\lambda\sin(2\theta)}\right)\right\}$$

and

$$E\left(R^p\right)=2^{\frac{p}{2}}\Gamma\left(\frac{p}{2}+1\right)$$

for $r>0$, $0\leq\theta\leq 2\pi$ and $p>0$.

Bivariate alpha-skew-normal distribution (Louzada et al., 2016) has the joint pdf specified by

$$f(x,y)=C\left[1+(1-ax-by)^2\right]\exp\left(-x^2-y^2-2\rho xy\right)$$

for $-\infty<x<+\infty$, $-\infty<y<+\infty$, $-\infty<a<+\infty$, $-\infty<b<+\infty$ and $-1\leq b\leq 1$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta)=Cr\left[2-2r\left(a\sin\theta+b\cos\theta\right)+r^2\left(a\sin\theta+b\cos\theta\right)^2\right]\exp\left[-r^2-\rho r^2\sin(2\theta)\right],$$

$$f_\Theta(\theta)=\frac{C}{1+\rho\sin(2\theta)}-\frac{\sqrt{\pi}C\left(a\sin\theta+b\cos\theta\right)}{2\left[1+\rho\sin(2\theta)\right]^{3/2}}+\frac{C\left(a\sin\theta+b\cos\theta\right)^2}{2\left[1+\rho\sin(2\theta)\right]^2},$$

$$\begin{aligned}f_R(r)&=Cr\exp\left(-r^2\right)\sum_{k=0}^{\infty}\frac{\left(-2\rho r^2\right)^k}{k!^2}\left[1+(-1)^k\right]\Gamma^2\left(\frac{k+1}{2}\right)\\&+abCr^3\exp\left(-r^2\right)\sum_{k=0}^{\infty}\frac{\left(-2\rho r^2\right)^k}{k!(k+1)!}\left[1+(-1)^{k+1}\right]\Gamma^2\left(\frac{k+2}{2}\right)\\&+\left(a^2+b^2\right)Cr^3\exp\left(-r^2\right)\sum_{k=0}^{\infty}\frac{\left(-2\rho r^2\right)^k}{2k!(k+1)!}\left[1+(-1)^k\right]\left[1+(-1)^{k+2}\right]\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{k+3}{2}\right)\end{aligned}$$

and

$$\begin{aligned}E\left(R^p\right)&=C\sum_{k=0}^{\infty}\frac{(-2\rho)^k}{k!^2}\left[1+(-1)^k\right]\Gamma\left(\frac{p}{2}+k+1\right)\Gamma^2\left(\frac{k+1}{2}\right)\\&+abC\sum_{k=0}^{\infty}\frac{(-2\rho)^k}{k!(k+1)!}\Gamma\left(\frac{p}{2}+k+2\right)\left[1+(-1)^{k+1}\right]\Gamma^2\left(\frac{k+2}{2}\right)\\&+\left(a^2+b^2\right)C\sum_{k=0}^{\infty}\frac{(-2\rho)^k}{2k!(k+1)!}\Gamma\left(\frac{p}{2}+k+2\right)\left[1+(-1)^k\right]\left[1+(-1)^{k+2}\right]\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{k+3}{2}\right)\end{aligned}$$

for $r>0$ and $0\leq\theta\leq 2\pi$.

Bivariate generalized skew-symmetric normal distribution (Fathi-Vajargah and Hasanalipour, 2013, equation (1)) has the joint pdf specified by

$$f(x,y) = C \exp\left(-\frac{x^2+y^2}{2}\right) \Phi\left(\frac{axy}{1+bx^2y^2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$ and $b \geq 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr \exp\left(-\frac{r^2}{2}\right) \Phi\left(\frac{2ar^2\sin(2\theta)}{4+br^4\sin^2(2\theta)}\right),$$

$$f_R(r) = Cr \exp\left(-\frac{r^2}{2}\right) \int_0^{2\pi} \Phi\left(\frac{2ar^2\sin(2\theta)}{4+br^4\sin^2(2\theta)}\right) d\theta$$

and

$$f_\Theta(\theta) = C \int_0^{\infty} r \exp\left(-\frac{r^2}{2}\right) \Phi\left(\frac{2ar^2\sin(2\theta)}{4+br^4\sin^2(2\theta)}\right) dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate Kotz type distribution (Balakrishnan and Lai, 2009, Section 13.6.1) has the joint pdf specified by

$$f(x,y) = C\left(x^2+y^2\right)^{N-1} \exp\left[-\phi\left(x^2+y^2\right)^s\right]$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $N > 0$, $s > 0$ and $\phi > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr^{2(N-1)} \exp\left(-\phi r^{2s}\right),$$

$$f_R(r) = 2\pi C r^{2(N-1)} \exp\left(-\phi r^{2s}\right),$$

$$f_\Theta(\theta) = \frac{C}{2s} \phi^{-\frac{N-\frac{1}{2}}{s}} \Gamma\left(\frac{N-\frac{1}{2}}{s}\right)$$

and

$$E\left(R^p\right) = \frac{C\pi}{s}\phi^{-\frac{2N+p-1}{2s}}\Gamma\left(\frac{2N+p-1}{2s}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Bivariate t distribution (Kotz and Nadarajah, 2004, Chapter 1) has the joint pdf specified by

$$f(x, y) = C\left(1 + ax^2 + by^2 + 2cxy\right)^{-\frac{\nu+2}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$, $-\infty < c < +\infty$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = Cr\left\{1 + \frac{a+b}{2}r^2 + \left[\frac{b-a}{2}\cos(2\theta) + c\sin(2\theta)\right]r^2\right\}^{-\frac{\nu+2}{2}},$$

$$\begin{aligned}f_R(r) = & \frac{C}{2}\sum_{k=0}^{\infty}\sum_{\ell=0}^{k}\binom{-\frac{\nu+2}{2}}{k}\binom{k}{\ell}c^\ell\left(\frac{b-a}{2}\right)^{k-\ell}\left[1 + (-1)^k + (-1)^\ell + (-1)^{k-\ell}\right] \\ & \cdot B\left(\frac{\ell+1}{2}, \frac{k-\ell+1}{2}\right)r^{2k+1}\left[1 + \frac{a+b}{2}r^2\right]^{-\frac{\nu+2}{2}-k},\end{aligned}$$

$$f_\Theta(\theta) = \frac{2C}{\nu\left[a + b + (b-a)\cos(2\theta) + 2c\sin(2\theta)\right]}$$

and

$$\begin{aligned}E\left(R^p\right) = & \frac{2^{\frac{p}{2}-1}C}{(b+a)^{\frac{p}{2}+1}}\sum_{k=0}^{\infty}\sum_{\ell=0}^{k}\binom{-\frac{\nu+2}{2}}{k}\binom{k}{\ell}(2c)^\ell\frac{(b-a)^{k-\ell}}{(b+a)^k}\left[1 + (-1)^k + (-1)^\ell + (-1)^{k-\ell}\right] \\ & \cdot B\left(\frac{\ell+1}{2}, \frac{k-\ell+1}{2}\right)B\left(k + \frac{p}{2} + 1, \frac{\nu-p}{2}\right)\end{aligned}$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < \nu$.

Bivariate Cauchy distribution (Kotz and Nadarajah, 2004, Chapter 1) has the joint pdf specified by

$$f(x, y) = C\left(1 + ax^2 + by^2\right)^{-\frac{3}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$ and $b > 0$, where C denotes the normalizing constant.

For this distribution,

$$f(r,\theta) = Cr\left\{1 + ar^2 + \frac{b-a}{2}\left[1+\cos(2\theta)\right]r^2\right\}^{-\frac{3}{2}},$$

$$f_R(r) = C\sum_{k=0}^{\infty}\binom{-\frac{3}{2}}{k}\left(\frac{b-a}{2}\right)^k\left[1+(-1)^k\right]B\left(\frac{1}{2},\frac{k+1}{2}\right)r^{2k+1}\left[1+\frac{a+b}{2}r^2\right]^{-\frac{3}{2}-k},$$

$$f_\Theta(\theta) = \frac{2C}{a+b+(b-a)\cos(2\theta)}$$

and

$$E\left(R^p\right) = \frac{2^{\frac{p}{2}}C}{(b+a)^{\frac{p}{2}+1}}\sum_{k=0}^{\infty}\binom{-\frac{3}{2}}{k}\left(\frac{b-a}{b+a}\right)^k\left[1+(-1)^k\right]B\left(\frac{1}{2},\frac{k+1}{2}\right)B\left(k+\frac{p}{2}+1,\frac{1-p}{2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < 1$.

Bivariate skew t distribution (Azzalini and Capitanio, 2003) has the joint pdf specified by

$$f(x,y) = \frac{1}{\pi}\left(1+\frac{x^2+y^2}{\nu}\right)^{-\frac{\nu+2}{2}}T_{\nu+2}\left((ax+by)\sqrt{\frac{\nu+2}{x^2+y^2+\nu}}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$, $-\infty < b < +\infty$ and $\nu > 0$, where $T_a(\cdot)$ denotes the cdf of a standard Student's t random variable with degree of freedom a. For this distribution,

$$f(r,\theta) = \frac{r}{\pi}\left(1+\frac{r^2}{\nu}\right)^{-\frac{\nu+2}{2}}T_{\nu+2}\left(r\left(a\sin\theta+b\cos\theta\right)\sqrt{\frac{\nu+2}{r^2+\nu}}\right),$$

$$f_R(r) = r\left(1+\frac{r^2}{\nu}\right)^{-\frac{\nu+2}{2}},$$

$$f_\Theta(\theta) = \frac{1}{2\pi} + \frac{\nu\Gamma\left(\frac{\nu+3}{2}\right)}{2\pi^{3/2}\Gamma\left(\frac{\nu+2}{2}\right)}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{1}{2}\right)_k\left(\frac{\nu+3}{2}\right)_k}{k!\left(\frac{3}{2}\right)_k}\left(a\sin\theta+b\cos\theta\right)^{2k+1}B\left(k+\frac{3}{2},\frac{\nu}{2}\right)$$

and

$$E\left(R^p\right) = 2^{-1}\nu^{\frac{p}{2}+1}B\left(\frac{p}{2}+1,\frac{\nu-p}{2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < \nu$.

Bivariate skew Cauchy distribution (Azzalini and Capitanio, 2003) has the joint pdf specified by

$$f(x,y) = \frac{1}{\pi} \left(1 + x^2 + y^2\right)^{-\frac{3}{2}} T_3 \left((ax + by) \sqrt{\frac{3}{x^2 + y^2 + 1}} \right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$ and $-\infty < b < +\infty$. For this distribution,

$$f(r,\theta) = \frac{r}{\pi} \left(1 + r^2\right)^{-\frac{3}{2}} T_3 \left(r \left(a \sin\theta + b \cos\theta\right) \sqrt{\frac{3}{r^2 + 1}} \right),$$

$$f_R(r) = r \left(1 + r^2\right)^{-\frac{3}{2}},$$

$$f_\Theta(\theta) = \frac{1}{2\pi} + \frac{1}{\pi^2} \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{1}{2}\right)_k (2)_k}{\left(\frac{3}{2}\right)_k} \left(a \sin\theta + b \cos\theta\right)^{2k+1} B \left(k + \frac{3}{2}, \frac{1}{2}\right)$$

and

$$E\left(R^p\right) = 2^{-1} B \left(\frac{p}{2} + 1, \frac{1-p}{2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < 1$.

Standard bivariate t distribution (Kotz and Nadarajah, 2004, Chapter 1) has the joint pdf specified by

$$f(x,y) = C \left(a^2 + x^2 + y^2\right)^{-\frac{\nu+2}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr \left(a^2 + r^2\right)^{-\frac{\nu+2}{2}},$$

$$f_R(r) = 2\pi Cr \left(a^2 + r^2\right)^{-\frac{\nu+2}{2}},$$

$$f_\Theta(\theta) = \frac{C}{\nu a^\nu}$$

and

$$E\left(R^p\right) = \pi C a^{p-\nu} B\left(\frac{p}{2}+1, \frac{\nu-p}{2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < \nu$.

Conditionally specified bivariate t distribution (Kotz and Nadarajah, 2004, equation (4.26)) has the joint pdf specified by

$$f(x,y) = C\left(a + bx^2 + by^2 + cx^2y^2\right)^{-\frac{\nu+1}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$, $-\infty < c < +\infty$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\left[a + br^2 + \frac{c}{4}r^4\sin^2(2\theta)\right]^{-\frac{\nu+1}{2}},$$

$$f_R(r) = 2Cr\left(a+br^2\right)^{-\frac{\nu+1}{2}}\sum_{k=0}^{\infty}\binom{-\frac{\nu+1}{2}}{k}\left[\frac{cr^4}{4\left(a+br^2\right)}\right]^k B\left(\frac{1}{2}, k+\frac{1}{2}\right)$$

and

$$f_\Theta(\theta) = \frac{C}{\sqrt{c}\nu a^{\nu/2}\left|\sin(2\theta)\right|} {}_2F_1\left(\frac{1}{2}, \frac{\nu}{2}; \frac{\nu}{2}+1; 1-\frac{b^2}{ac\sin^2(2\theta)}\right)$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate poly t distribution (Dickey, 1968) has the joint pdf specified by

$$f(x,y) = C\left(1 + \alpha x^2 + \alpha y^2\right)^{-\frac{\mu+1}{2}}\left(1 + \beta x^2 + \beta y^2\right)^{-\frac{\nu+1}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$, $\beta > 0$, $\mu > 0$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\left(1+\alpha r^2\right)^{-\frac{\mu+1}{2}}\left(1+\beta r^2\right)^{-\frac{\nu+1}{2}},$$

$$f_R(r) = 2\pi C r \left(1 + \alpha r^2\right)^{-\frac{\mu+1}{2}} \left(1 + \beta r^2\right)^{-\frac{\nu+1}{2}},$$

$$f_\Theta(\theta) = \frac{C}{\alpha(\mu+\nu)} {}_2F_1\left(1, \frac{\nu+1}{2}; \frac{\mu+\nu}{2} + 1; 1 - \frac{\beta}{\alpha}\right)$$

and

$$E\left(R^p\right) = \pi \alpha^{-\frac{p}{2}-1} C B\left(\frac{p}{2} + 1, \frac{\mu+\nu-p}{2}\right) {}_2F_1\left(\frac{p}{2} + 1, \frac{\nu+1}{2}; \frac{\mu+\nu}{2} + 1; 1 - \frac{\beta}{\alpha}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < \mu + \nu$.

Bivariate poly Cauchy distribution (Dickey, 1968) has the joint pdf specified by

$$f(x, y) = C\left(1 + \alpha x^2 + \alpha y^2\right)^{-1} \left(1 + \beta x^2 + \beta y^2\right)^{-1}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$ and $\beta > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = C r \left(1 + \alpha r^2\right)^{-1} \left(1 + \beta r^2\right)^{-1},$$

$$f_R(r) = 2\pi C r \left(1 + \alpha r^2\right)^{-1} \left(1 + \beta r^2\right)^{-1},$$

$$f_\Theta(\theta) = \frac{C}{2\alpha} {}_2F_1\left(1, 1; 2; 1 - \frac{\beta}{\alpha}\right)$$

and

$$E\left(R^p\right) = \pi \alpha^{-\frac{p}{2}-1} C B\left(\frac{p}{2} + 1, \frac{2-p}{2}\right) {}_2F_1\left(\frac{p}{2} + 1, 1; 2; 1 - \frac{\beta}{\alpha}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p < 2$.

Bivariate heavy tailed distribution (Balakrishnan and Lai, 2009, equation (9.22)) has the joint pdf specified by

$$f(x, y) = C\left(1 + x^2\right)^{-\frac{\alpha}{2}} \left(1 + y^2\right)^{-\frac{\beta}{2}} \left(1 + x^2 + y^2\right)^{-\frac{\gamma}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\left(1 + r^2 \sin^2\theta\right)^{-\frac{\alpha}{2}} \left(1 + r^2 \cos^2\theta\right)^{-\frac{\beta}{2}} \left(1 + r^2\right)^{-\frac{\gamma}{2}},$$

$$f_R(r) = 2\pi Cr\left(1 + r^2\right)^{-\frac{\gamma+\beta}{2}} F_1\left(\frac{1}{2}, \frac{\alpha}{2}, \frac{\beta}{2}, 1; -r^2, \frac{r^2}{1+r^2}\right)$$

and

$$f_\Theta(\theta) = C\int_0^\infty \left(1 + 2y\sin^2\theta\right)^{-\frac{\alpha}{2}} \left(1 + 2y\cos^2\theta\right)^{-\frac{\beta}{2}} (1+2y)^{-\frac{\gamma}{2}}\, dy$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Standard symmetric bivariate Laplace distribution (Kotz et al., 2001, equation (5.1.2)) has the joint pdf specified by

$$f(x,y) = \frac{1}{\pi} K_0\left(\sqrt{2\left(x^2+y^2\right)}\right)$$

for $-\infty < x < +\infty$ and $-\infty < y < +\infty$. For this distribution,

$$f(r,\theta) = \frac{1}{\pi} r K_0\left(\sqrt{2}r\right),$$

$$f_R(r) = 2rK_0\left(\sqrt{2}r\right),$$

$$f_\Theta(\theta) = \frac{1}{2\pi}$$

and

$$E\left(R^p\right) = 2^{\frac{p}{2}}\Gamma^2\left(\frac{p}{2}+1\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

General symmetric bivariate Laplace distribution (Kotz et al., 2001, equation (5.2.2)) has the joint pdf specified by

$$f(x,y) = C\left(x^2+y^2\right)^{\frac{\nu}{2}} K_\nu\left(\sqrt{2\left(x^2+y^2\right)}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = C r^{1+\nu} K_\nu \left(\sqrt{2} r \right),$$

$$f_R(r) = 2\pi C r^{1+\nu} K_\nu \left(\sqrt{2} r \right),$$

$$f_\Theta(\theta) = 2^{\frac{\nu}{2}-1} C \Gamma (\nu + 1)$$

and

$$E\left(R^p\right) = \pi C 2^{\frac{p+\nu}{2}} \Gamma \left(1 + \nu + \frac{p}{2} \right) \Gamma \left(1 + \frac{p}{2} \right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Asymmetric bivariate Laplace distribution (Kotz et al., 2001, equation (6.5.3)) has the joint pdf specified by

$$f(x, y) = C \exp\left(\alpha x + \beta y\right) \left(x^2 + y^2\right)^{\frac{\nu}{2}} K_\nu \left(\gamma \sqrt{x^2 + y^2} \right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha < 0$, $\beta < 0$, $\gamma > 0$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = C r^{1+\nu} \exp\left(\alpha r \sin\theta + \beta r \cos\theta\right) K_\nu \left(\gamma r\right),$$

$$f_R(r) = 2\pi C r^{1+\nu} I_0 \left(r \sqrt{\alpha^2 + \beta^2} \right) K_\nu \left(\gamma r\right),$$

$$f_\Theta(\theta) = \frac{\sqrt{\pi} C (2\gamma)^\nu \Gamma(2 + 2\nu)}{\Gamma\left(\frac{5}{2} + \nu\right) \left(\gamma - \alpha \sin\theta - \beta \cos\theta\right)^{2+2\nu}} {}_2F_1 \left(2 + 2\nu, \frac{1}{2} + \nu; \frac{5}{2} + \nu; -\frac{\gamma + \alpha \sin\theta + \beta \cos\theta}{\gamma - \alpha \sin\theta - \beta \cos\theta} \right)$$

and

$$E\left(R^p\right) = 2^{\nu+p+1} \pi C \gamma^{-2-\nu-p} \Gamma \left(1 + \nu + \frac{p}{2} \right) \Gamma \left(1 + \frac{p}{2} \right) {}_2F_1 \left(1 + \nu + \frac{p}{2}, 1 + \frac{p}{2}; 1; \frac{\alpha^2 + \beta^2}{\gamma^2} \right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Standard asymmetric bivariate Laplace distribution (Kotz et al., 2001, page 302) has the joint pdf specified by

$$f(x, y) = C \exp\left(\alpha x + \beta y\right) K_0 \left(\gamma \sqrt{x^2 + y^2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha < 0$, $\beta < 0$ and $\gamma > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = C r \exp\left(\alpha r \sin\theta + \beta r \cos\theta\right) K_0\left(\gamma r\right),$$

$$f_R(r) = 2\pi C r I_0 \left(r \sqrt{\alpha^2 + \beta^2}\right) K_0\left(\gamma r\right),$$

$$f_\Theta(\theta) = \frac{4C}{3\left(\gamma - \alpha\sin\theta - \beta\cos\theta\right)^2} {}_2F_1\left(2, \frac{1}{2}; \frac{5}{2}; -\frac{\gamma + \alpha\sin\theta + \beta\cos\theta}{\gamma - \alpha\sin\theta - \beta\cos\theta}\right)$$

and

$$E\left(R^p\right) = 2^{p+1}\pi C \gamma^{-2-p}\Gamma^2\left(1 + \frac{p}{2}\right) {}_2F_1\left(1 + \frac{p}{2}, 1 + \frac{p}{2}; 1; \frac{\alpha^2 + \beta^2}{\gamma^2}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Bivariate poly Laplace distribution (Aravkin et al., 2011) has the joint pdf specified by

$$f(x, y) = C \exp\left(\alpha x + \beta y\right) \left(x^2 + y^2\right)^\lambda K_\mu\left(\gamma\sqrt{x^2 + y^2}\right) K_\nu\left(\gamma\sqrt{x^2 + y^2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $\lambda > 0$, $\mu > 0$ and $\nu > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r, \theta) = C r^{1+2\lambda} \exp\left(\alpha r \sin\theta + \beta r \cos\theta\right) K_\mu\left(\gamma r\right) K_\nu\left(\gamma r\right),$$

$$f_R(r) = 2\pi C r^{1+2\lambda} I_0\left(r\sqrt{\alpha^2 + \beta^2}\right) K_\mu\left(\gamma r\right) K_\nu\left(\gamma r\right),$$

$$f_\Theta(\theta) = C \int_0^\infty r^{1+2\lambda} \exp\left(\alpha r \sin\theta + \beta r \cos\theta\right) K_\mu\left(\gamma r\right) K_\nu\left(\gamma r\right) dr$$

and

$$\begin{aligned} E\left(R^p\right) = & \frac{2^{2\lambda+p}\pi C}{\gamma^{2+2\lambda+p}\Gamma\left(2+2\lambda+p\right)} \Gamma\left(1+\lambda+\frac{p+\mu+\nu}{2}\right)\Gamma\left(1+\lambda+\frac{p+\mu-\nu}{2}\right) \\ & \cdot\Gamma\left(1+\lambda+\frac{p-\mu+\nu}{2}\right)\Gamma\left(1+\lambda+\frac{p-\mu-\nu}{2}\right)\Gamma^{-1}\left(2\lambda+p+2\right) \\ & \cdot {}_4F_3\left(1+\lambda+\frac{p+\mu+\nu}{2}, 1+\lambda+\frac{p+\mu-\nu}{2}, 1+\lambda+\frac{p-\mu+\nu}{2}, 1+\lambda+\frac{p-\mu-\nu}{2};\right. \\ & \left. 1+\lambda+\frac{p}{2}, 1+\lambda+\frac{p+1}{2}, 1; \frac{\alpha^2+\beta^2}{4\gamma^2}\right) \end{aligned}$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Bivariate hyperbolic distribution (Balakrishnan and Lai, 2009, Section 13.14) has the joint pdf specified by

$$f(x,y) = C \exp\left[-\alpha\left(x^2+y^2\right) - \beta x - \gamma y\right]$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr \exp\left(-\alpha r^2 - \beta r \sin\theta - \gamma r \cos\theta\right),$$

$$f_R(r) = 2\pi Cr \exp\left(-\alpha r^2\right) I_0\left(r\sqrt{\beta^2+\gamma^2}\right),$$

$$f_\Theta(\theta) = \frac{C}{2\alpha} \exp\left[\frac{\left(\beta\sin\theta+\gamma\cos\theta\right)^2}{8\alpha}\right] D_{-2}\left(\frac{\beta\sin\theta+\gamma\cos\theta}{\sqrt{2\alpha}}\right)$$

and

$$E\left(R^p\right) = \pi C \alpha^{-\frac{p}{2}-1}\Gamma\left(\frac{p}{2}+1\right) {}_1F_1\left(\frac{p}{2}+1; 1; \frac{\beta^2+\gamma^2}{4\alpha}\right)$$

for $r > 0$, $0 \leq \theta \leq 2\pi$ and $p > 0$.

Bivariate hyperbolic secant distribution (Balakrishnan and Lai, 2009, Chapter 13) has the joint pdf specified by

$$f(x,y)=\frac{1}{4}(1+cxy)\operatorname{sech}\left(\frac{\pi x}{2}\right)\operatorname{sech}\left(\frac{\pi y}{2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $-1 < c < 1$. For this distribution,

$$f(r,\theta)=\frac{r}{4}\left(1+\frac{c}{2}r^2\sin 2\theta\right)\operatorname{sech}\left(\frac{\pi r\sin\theta}{2}\right)\operatorname{sech}\left(\frac{\pi r\cos\theta}{2}\right),$$

$$f_R(r)=\frac{r}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{E_{2i}E_{2j}\pi^{2i+2j}}{(2i)!(2j)!4^{i+j}}B\left(i+\frac{1}{2},j+\frac{1}{2}\right)r^{2i+2j},$$

$$\begin{aligned}f_\Theta(\theta)=\ &\pi^{-2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\left[\left(i+\frac{1}{2}\right)\sin\theta+\left(j+\frac{1}{2}\right)\cos\theta\right]^{-2}\\&\cdot\left\{1+3c\pi^{-2}\sin(2\theta)\left[\left(i+\frac{1}{2}\right)\sin\theta+\left(j+\frac{1}{2}\right)\cos\theta\right]^{-2}\right\}\end{aligned}$$

if $\sin\theta\cos\theta\geq 0$, and

$$\begin{aligned}f_\Theta(\theta)=\ &\pi^{-2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}(-1)^{i+j}\left[\left(i+\frac{1}{2}\right)\sin\theta-\left(j+\frac{1}{2}\right)\cos\theta\right]^{-2}\\&\cdot\left\{1+3c\pi^{-2}\sin(2\theta)\left[\left(i+\frac{1}{2}\right)\sin\theta-\left(j+\frac{1}{2}\right)\cos\theta\right]^{-2}\right\}\end{aligned}$$

if $\sin\theta\cos\theta < 0$ for $r > 0$, $0\leq\theta\leq 2\pi$ and $p > 0$.

Conditionally specified bivariate Gumbel distribution (Balakrishnan and Lai, 2009, Section 12.13.1) has the joint pdf specified by

$$f(x,y)=C\exp\left[-x-y-\exp(-x)-\exp(-y)-\alpha\exp(-x-y)\right]$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $0 < \alpha < 1$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta)=Cr\exp\left[-r\sin\theta-r\cos\theta-\exp(-r\sin\theta)-\exp(-r\cos\theta)-\alpha\exp(-r\sin\theta-r\cos\theta)\right],$$

$$f_R(r) = 2\pi C r \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(-1)^{i+j+k}\alpha^k}{i!j!k!} I_0\left(r\sqrt{(1+i+k)^2+(1+j+k)^2}\right)$$

and

$$f_\Theta(\theta) = C \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \frac{(-1)^{i+j+k}\alpha^k}{i!j!k!} \left[(1+i+k)\sin\theta + (1+j+k)\cos\theta\right]^{-2}$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate logistic distribution (Balakrishnan and Lai, 2009, Section 2.3.1) has the joint pdf specified by

$$f(x,y) = 2\exp\left(-\alpha x - \beta y\right)\left[1+\exp(-\alpha x)+\exp(-\beta y)\right]^{-3}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$ and $\beta > 0$. For this distribution,

$$f(r,\theta) = 2r\exp\left(-\alpha r\sin\theta - \beta r\cos\theta\right)\left[1+\exp(-\alpha r\sin\theta)+\exp(-\beta r\cos\theta)\right]^{-3},$$

$$f_R(r) = 4\pi r \sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \binom{-3}{k}\binom{k}{\ell} I_0\left(r\sqrt{\alpha^2(\ell+1)^2+\beta^2(k-\ell+1)^2}\right)$$

and

$$f_\Theta(\theta) = 2\sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \binom{-3}{k}\binom{k}{\ell} \left[(1+\ell)\alpha\sin\theta + (1+k-\ell)\beta\cos\theta\right]^{-2}$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate Gumbel type distribution (Balakrishnan and Lai, 2009, equation (13.27)) has the joint pdf specified by

$$f(x,y) = C\exp\left\{-a\left(x^2+y^2\right) - b\exp\left[-a\left(x^2+y^2\right)\right]\right\}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$ and $b > 0$, where C denotes the normalizing constant. For this distribution,

$$f(r,\theta) = Cr\exp\left\{-ar^2 - b\exp\left(-ar^2\right)\right\},$$

$$f_R(r) = 2\pi Cr \exp\left\{-ar^2 - b\exp\left(-ar^2\right)\right\},$$

$$f_\Theta(\theta) = \frac{C}{2ab}\left(1 - \exp(-b)\right),$$

$$E\left(R^p\right) = \frac{\pi C}{a(-a)^{p/2}}\left(\frac{\partial}{\partial\alpha}\right)^{p/2}\left[b^{-\alpha-1}\gamma(\alpha+1,b)\right]\Bigg|_{\alpha=0}$$

if p is even, and

$$E\left(R^p\right) = \frac{\pi C}{a(-a)^{p/2}}\left(\frac{\partial}{\partial\alpha}\right)^{(p-1)/2}\left[\int_0^1 \sqrt{\log t}\, t^\alpha \exp(-bt)dt\right]\Bigg|_{\alpha=0}$$

if p is odd, where $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate skew elliptical distribution (Arnold and Beaver, 2002, Section 10) has the joint pdf specified by

$$f(x,y) = Cg(x,y)H\left(a + bx + cy\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $-\infty < a < +\infty$, $-\infty < b < +\infty$ and $-\infty < c < +\infty$, where C denotes the normalizing constant, g denotes a valid joint pdf and H denotes a valid univariate cdf. For this distribution,

$$f(r,\theta) = Crg(r\sin\theta, r\cos\theta)H\left(a + br\sin\theta + cr\cos\theta\right),$$

$$f_R(r) = Cr\int_0^{2\pi} g\left(r\sin\theta, r\cos\theta\right)H\left(a + br\sin\theta + cr\cos\theta\right)d\theta$$

and

$$f_\Theta(\theta) = C\int_0^{\infty} rg\left(r\sin\theta, r\cos\theta\right)H\left(a + br\sin\theta + cr\cos\theta\right)dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate Sarmanov distribution (Sarmanov, 1966) has the joint pdf specified by

$$f(x,y) = g_1(x)g_2(y)\left\{1 + \alpha\theta_1(x)\theta_2(y)\right\}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $-1 \leq \alpha \leq 1$, where g_1, g_2 are valid probability density functions and θ_1, θ_2 are bounded nonconstant functions such that

$$\int_{-\infty}^{+\infty} \theta_1(x) g_1(x) dx = 0, \qquad \int_{-\infty}^{+\infty} \theta_2(y) g_2(y) dy = 0.$$

For this distribution,

$$f(r, \theta) = r g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha \theta_1 (r \sin \theta) \theta_2 (r \cos \theta)\},$$

$$f_R(r) = r \int_0^{2\pi} g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha \theta_1 (r \sin \theta) \theta_2 (r \cos \theta)\} d\theta$$

and

$$f_\Theta(\theta) = \int_0^{\infty} r g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha \theta_1 (r \sin \theta) \theta_2 (r \cos \theta)\} dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

Bivariate Farlie-Gumbel-Morgenstern distribution (Farlie, 1960; Gumbel, 1960; Morgenstern, 1956) has the joint pdf specified by

$$f(x, y) = g_1(x) g_2(y) \{1 + \alpha [1 - 2G_1(x)] [1 - 2G_2(y)]\}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$ and $-1 \leq \alpha \leq 1$, where g_1, g_2 are valid probability density functions and G_1, G_2 are the corresponding cumulative distribution functions. For this distribution,

$$f(r, \theta) = r g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha [1 - 2G_1 (r \sin \theta)] [1 - 2G_2 (r \cos \theta)]\},$$

$$f_R(r) = r \int_0^{2\pi} g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha [1 - 2G_1 (r \sin \theta)] [1 - 2G_2 (r \cos \theta)]\} d\theta$$

and

$$f_\Theta(\theta) = \int_0^{\infty} r g_1 (r \sin \theta) g_2 (r \cos \theta) \{1 + \alpha [1 - 2G_1 (r \sin \theta)] [1 - 2G_2 (r \cos \theta)]\} dr$$

for $r > 0$ and $0 \leq \theta \leq 2\pi$.

6.4 Simulation

Simulation of R and Θ from the stated bivariate distributions is simple given there are algorithms for simulating (X, Y):

- simulate (X, Y) from the stated pdf $f(x, y)$;
- set $R = \sqrt{X^2 + Y^2}$;
- set $\Theta = \arctan(Y/X)$.

Algorithms for simulating from each of the stated bivariate distributions are available in the literature. For example, a random vector (X, Y) having the bivariate normal distribution with means (μ_x, μ_y), variances $\left(\sigma_x^2, \sigma_y^2\right)$ and correlation coefficient ρ can be simulated as

- simulate Z_1 and Z_2 independently from a standard normal distribution;
- set $X = \left(1 - \rho^2\right) \sigma_x Z_1 + \rho \sigma_x Z_2 + \mu_x$;
- set $Y = \sigma_y Z_2 + \mu_y$.

Similarly, a random vector (X, Y) having the bivariate t distribution with means (μ_x, μ_y), scale parameters $\left(\sigma_x^2, \sigma_y^2\right)$, correlation coefficient ρ and degree of freedom ν can be simulated as

- simulate Z_1 and Z_2 independently from a standard normal distribution;
- set $P = \left(1 - \rho^2\right) \sigma_x Z_1 + \rho \sigma_x Z_2$;
- set $Q = \sigma_y Z_2$;
- simulate U independently from a chisquare distribution with degree of freedom ν;
- set $X = \mu_x + P \sqrt{\frac{\nu}{U}}$;
- set $Y = \mu_y + Q \sqrt{\frac{\nu}{U}}$.

Chapter 7

Characteristic functions of product

7.1 Introduction

Many variables in the real world (including the signal processing area) can be assumed to follow the normal distribution. That is, we can write $U = \mu + \sigma X$, where X is a standard normal variable, μ is the mean and σ is the standard deviation. But often the mean and standard deviation are themselves random variables, so U involves a product of two random variables. For the rest of this chapter, we use Y to represent σ.

Schoenecker and Luginbuhl (2016) derived the characteristic function of the product, i.e. XY, when X is a standard normal random variable and Y is an independent random variable following either the normal or gamma distribution. They expressed the distribution of $XY = W$ say in terms of its characteristic function $\phi_W(t) = E\left[\exp\left(\mathrm{i}tW\right)\right]$, where $\mathrm{i} = \sqrt{-1}$ is the complex unit.

The charateristic function defines the probability distribution of any real-valued random variable. It is the Fourier transform of the probability density function. The tail behavior of the charateristic function defines the smoothness character of the density function. Literatures where the characteristic function arises include: estimation of the joint reperesentations in signal analysis (Cohen, 1996); estimation of affine asset pricing models (Singleton, 2001); estimation of continuous-time stochastic processes (Jiang and Knight, 2002).

The aim of this chapter is to derive closed form expressions for the characteristic function $\phi_W(t)$ when X is a standard normal random variable and Y is an independent random variable following a wide range of other distributions. For a variety of applications, it is needed that the

σ follows different types of distributions. For example, as introduced in the previous chapter, in signal transfer, amplitude is the height from the central line to peak. There is another measure which refers to the distance between two peaks, so called, period. The period of different waves may be different, which results in the difference in shape of parabolic curves. Therefore, we need to adjust the shape parameter (σ) to provide the goodness of fit to different signal data. Here we construct a comprehensive study on the light-tailed distributions (e.g. nomal distribution, exponential distribution), and the heavy tailed distributions (e.g. the Pareto distribution, the Cauchy distribution, the student's t distribution) for Y. The full list of distributions for Y considered in this chapter is: Pareto distribution (Pareto, 1964), triangular distribution, Argus distribution (Albrecht, 1990), Cauchy distribution, Student's t distribution (Gosset, 1908), skewed Student's t distribution (Zhu and Galbraith, 2010), asymmetric skewed Student's t distribution (Zhu and Galbraith, 2010), Rice distribution (Rice, 1945), symmetric Laplace distribution (Laplace, 1774), Laplace distribution (Laplace, 1774), asymmetric Laplace distribution (Kozubowski and Podgorski, 2000), Poiraud-Casanova-Thomas-Agnan Laplace distribution (Poiraud-Casanova and Thomas-Agnan, 2000), Holla-Bhattacharya Laplace distribution (Holla and Bhattacharya, 1968), McGill Laplace distribution (McGill, 1962), log Laplace distribution, exponential distribution, gamma distribution, chi distribution, variance gamma distribution (Madan and Seneta, 1990), normal inverse gamma distribution, Nakagami distribution (Nakagami, 1960), reciprocal distribution, Maxwell distribution (Maxwell, 1860), quadratic distribution, uniform distribution, power function distribution, Rayleigh distribution (Weibull, 1951), exponentiated Rayleigh distribution, beta Rayleigh distribution, normal distribution (de Moivre, 1738; Gauss, 1809), truncated normal distribution, split normal distribution, q-Gaussian distribution (Tsallis, 2009), normal exponential gamma distribution, Wigner semicircle distribution, Kumaraswamy distribution (Kumaraswamy, 1980), linear failure rate distribution (Bain, 1974) and Irwin Hall distribution (Irwin, 1927; Hall, 1927).

The given expressions are as explicit as possible. They involve various special functions, including the gamma function defined by

$$\Gamma(a) = \int_0^{+\infty} t^{a-1} \exp(-t)\, dt$$

for $a > 0$; the incomplete gamma function defined by

$$\gamma(a, x) = \int_0^{x} t^{a-1} \exp(-t)\, dt$$

for $a > 0$ and $x > 0$; the complementary incomplete gamma function defined by

$$\Gamma(a, x) = \int_x^{+\infty} t^{a-1} \exp(-t)\, dt$$

for $x > 0$; the beta function defined by

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

for $a > 0$ and $b > 0$; the incomplete beta function defined by

$$B_x(a, b) = \int_0^x t^{a-1}(1-t)^{b-1} dt$$

for $0 < x < 1$, $a > 0$ and $b > 0$; the error function defined by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt$$

for $x > 0$; the complementary error function defined by

$$\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} \exp\left(-t^2\right) dt$$

for $-\infty < x < +\infty$; the parabolic cylinder function of order ν defined by

$$D_\nu(x) = \frac{\exp\left(-x^2/4\right)}{\Gamma(-\nu/2)} \int_0^{+\infty} t^{-\frac{\nu}{2}-1}(1+2t)^{\frac{\nu-1}{2}} \exp\left(-x^2 t\right) dt$$

for $\nu < 0$ and $x^2 > 0$; the Whittaker W function of orders ν, μ defined by

$$W_{\nu,\mu}(x) = \frac{x^{\mu+\frac{1}{2}}}{\Gamma\left(\mu-\nu+\frac{1}{2}\right)} \int_{\frac{1}{2}}^{+\infty} \left(t-\frac{1}{2}\right)^{\mu-\nu-\frac{1}{2}} \left(t+\frac{1}{2}\right)^{\mu+\nu-\frac{1}{2}} \exp(-xt)\, dt$$

for $\mu - \nu > -\frac{1}{2}$ and $x > 0$; the modified Bessel function of the first kind of order ν defined by

$$I_\nu(x) = \sum_{k=0}^{+\infty} \frac{1}{\Gamma(k+\nu+1)k!} \left(\frac{x}{2}\right)^{2k+\nu};$$

the modified Bessel function of the second kind of order ν defined by

$$K_\nu(x) = \begin{cases} \dfrac{\pi \csc(\pi\nu)}{2}\left[I_{-\nu}(x) - I_\nu(x)\right], & \text{if } \nu \notin \mathbb{Z}, \\ \\ \lim\limits_{\mu\to\nu} K_\mu(x), & \text{if } \nu \in \mathbb{Z}; \end{cases}$$

the confluent hypergeometric function defined by

$${}_1F_1(\alpha;\beta;x) = \sum_{k=0}^{+\infty} \frac{(\alpha)_k}{(\beta)_k}\frac{x^k}{k!},$$

where $(\alpha)_k = \alpha(\alpha+1)\cdots(\alpha+k-1)$ denotes the ascending factorial; the Gauss hypergeometric function defined by

$${}_2F_1(\alpha,\beta;\gamma;x) = \sum_{k=0}^{+\infty} \frac{(\alpha)_k(\beta)_k}{(\gamma)_k}\frac{x^k}{k!};$$

the standard normal density function defined by

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right);$$

the standard normal distribution function defined by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt.$$

and the function $\Psi(a,c;z)$ defined by

$$\Psi(a,c;z) = \frac{\Gamma(1-c)}{\Gamma(1+a-c)}\,{}_1F_1(a;c;z) + \frac{\Gamma(c-1)}{\Gamma(a)} z^{1-c}\,{}_1F_1(1+a-c;2-c;z)$$

These special functions are well known and well established in the mathematics literature. Some details of their properties can be found in Prudnikov et al. (1986) and Gradshteyn and Ryzhik (2000). In-built routines for computing them are available in packages like Maple, Matlab and Mathematica. For example, the in-built routines in Mathematica for the stated special functions are: GAMMA[a] for the gamma function; GAMMA[a]-GAMMA[a,x] for the incomplete gamma function; GAMMA[a,x] for the complementary incomplete gamma function; Beta[a,b] for the beta function; Beta[x,a,b] for the incomplete beta function; Erf[x] for the error function; Erfc[x] for the complementary error function; ParabolicCylinderD[nu,x] for the parabolic cylinder function; Whittaker[nu,mu,x] for the

Whittaker W function of orders ν, μ; BesselI[nu,x] for the modified Bessel function of the first kind of order ν; BesselK[nu,x] for the modified Bessel function of the second kind of order ν; Hypergeometric1F1[alpha,beta,x] for the confluent hypergeometric function; Hypergeometric2F1[alpha,beta,gamma,x] for the Gauss hypergeometric function; PDF[NormalDistribution[0,1],x] for the standard normal density function; CDF[NormalDistribution[0,1],x] for the standard normal distribution function. Mathematica like other algebraic manipulation packages allows for arbitrary precision, so the accuracy of computations is not an issue.

The contents of this chapter are organized as follows. Section 7.2 provides simple derivations of the characteristic functions due to Schoenecker and Luginbuhl (2016). Section 7.3 lists explicit expressions for $\phi_W(t)$ for nearly fifty distributions for Y. Section 7.4 presents simulation results that verify correctness of the expressions in Section 7.3.

7.2 Simpler derivations for normal and gamma cases

Here, we present simpler derivations of the characteristic function of $W = XY$ when: i) X is a standard normal random variable and Y is an independent normal random variable with mean μ and standard deviation σ; ii) X is a standard normal random variable and Y is an independent gamma random variable with shape parameter α and scale parameter β. For any distribution of Y, we can write

$$\phi_W(t) = E\left[\exp(\mathrm{i}tW)\right] = E\left[\exp(\mathrm{i}tXY)\right] = E\left\{E\left[\exp(\mathrm{i}tXY) \mid Y\right]\right\} = E\left\{\exp\left(-\frac{t^2Y^2}{2}\right)\right\}.$$

If Y is a normal random variable with mean μ and standard deviation σ then

$$\begin{aligned}
\phi_W(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left[-\frac{t^2y^2}{2}-\frac{(y-\mu)^2}{2\sigma^2}\right]dy \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left[-\frac{\left(\sigma^2t^2+1\right)y^2-2\mu y+\mu^2}{2\sigma^2}\right]dy \\
&= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{\mu^2t^2}{2\left(\sigma^2t^2+1\right)}\right]\int_{-\infty}^{\infty}\exp\left[-\frac{\sigma^2t^2+1}{2\sigma^2}\left(y-\frac{\mu}{\sigma^2t^2+1}\right)^2\right]dy \\
&= \frac{1}{\sqrt{\sigma^2t^2+1}}\exp\left[-\frac{\mu^2t^2}{2\left(\sigma^2t^2+1\right)}\right],
\end{aligned}$$

where the last step follows from the fact that any probability density function must integrate to one. If Y is a gamma random variable with shape parameter α and scale parameter β then

$$\phi_W(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty y^{\alpha-1} \exp\left(-\beta y - \frac{t^2 y^2}{2}\right) dy = \left(\frac{\beta}{t}\right)^\alpha \exp\left(\frac{\beta^2}{4t^2}\right) D_{-\alpha}\left(\frac{\beta}{t}\right),$$

where the last step follows by direct application of equation (2.3.15.3) in Prudnikov et al. (1986, volume 1).

Note that the n-th moment of W can be generated from the characteristic function $\phi_W(t)$ by the following equation.

$$E\left[W^n\right] = i^{-n} \phi_W^{(n)}(0) = i^{-n} \left[\frac{\partial^n}{\partial t^n} \phi_W(t)\right]_{t=0},$$

where $\mathrm{i} = \sqrt{-1}$ is the complex unit. Moments can be used for obtaining the statistical characters of W. For example, the first moment indicates the mean of W. The second moment gives the information on the scale of the distribution of W; The third moment defines the skewness of the distribution of W. If the third moment is zero, the distribution is symmetric. The fourth moment is used for measuring the flatness or peakedness of the distribution.

7.3 Expressions for characteristic functions

Here, we list explicit expressions for $\phi_W(t)$ when X is a standard normal random variable and Y is an independent random variable following nearly fifty other distributions. We also select the representative models and provide the details of derivation in the Appendix C.

Pareto distribution (Pareto, 1964): for this distribution,

$$f_Y(y) = \alpha K^\alpha y^{-\alpha-1},$$
$$\phi_W(t) = \alpha K^\alpha 2^{-\frac{\alpha}{2}-1} t^\alpha \Gamma\left(-\frac{\alpha}{2}, \frac{K^2 t^2}{2}\right)$$

for $y \geq K > 0$ and $\alpha > 0$.

Triangular distribution: for this distribution,

$$
f_Y(y) = \begin{cases} \dfrac{2(y-a)}{(b-a)(c-a)}, & \text{if } a < y < c, \\[2em] \dfrac{2(b-y)}{(b-a)(b-c)}, & \text{if } c \le y < b, \\[2em] 0, & \text{elsewhere,} \end{cases}
$$

$$
\begin{aligned}
\phi_W(t) = {} & \frac{2}{t^2(b-a)(c-a)} \left[\exp\left(-\frac{t^2a^2}{2}\right) - \exp\left(-\frac{t^2c^2}{2}\right) \right] \\
& - \frac{2}{t^2(b-a)(b-c)} \left[\exp\left(-\frac{t^2c^2}{2}\right) - \exp\left(-\frac{t^2b^2}{2}\right) \right] \\
& - \frac{\sqrt{2}a}{t(b-a)(c-a)} \left[\Gamma\left(\frac{1}{2}, \frac{t^2a^2}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{t^2c^2}{2}\right) \right] \\
& + \frac{\sqrt{2}b}{t(b-a)(b-c)} \left[\Gamma\left(\frac{1}{2}, \frac{t^2c^2}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{t^2b^2}{2}\right) \right]
\end{aligned}
$$

for $a < y < b$.

Argus distribution (Albrecht, 1990): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{a^3}{\sqrt{2\pi}\Psi(a)c^2} \exp\left[-\frac{a^2}{2}\left(1 - \frac{y^2}{c^2}\right) \right] y \sqrt{1 - \frac{y^2}{c^2}}, \\
\phi_W(t) &= \frac{a^3}{3\sqrt{2\pi}\Psi(a)} \exp\left(-\frac{a^2}{2}\right) {}_1F_1\left(1; \frac{5}{2}; \frac{a^2 - c^2t^2}{2}\right)
\end{aligned}
$$

for $0 < y < c$ and $a > 0$, where $\Psi(x) = \Phi(x) - x\phi(x) - \frac{1}{2}$.

Cauchy distribution: for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{\gamma}{\pi\left(\gamma^2 + y^2\right)}, \\
\phi_W(t) &= \frac{\sqrt{2}}{\sqrt{\pi}} \exp\left(\frac{\gamma^2t^2}{4}\right) D_{-1}(\gamma t)
\end{aligned}
$$

for $-\infty < y < +\infty$ and $\gamma > 0$.

Student's t distribution (Gosset, 1908): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{y^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}}, \\
\phi_W(t) &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \Psi\left(\frac{1}{2}, \frac{3}{2} - \frac{\nu+1}{2}; \frac{\nu\sigma^2t^2}{2}\right)
\end{aligned}
$$

for $-\infty < y < +\infty$, $\nu > 0$ and $\sigma > 0$.

Skewed Student's t distribution (Zhu and Galbraith, 2010): for this distribution,

$$f_Y(y) = \begin{cases} K(\nu)\left[1+\frac{1}{\nu}\left(\frac{y}{2\alpha}\right)^2\right]^{-\frac{\nu+1}{2}}, & \text{if } y \leq 0, \\ \\ K(\nu)\left[1+\frac{1}{\nu}\left(\frac{y}{2(1-\alpha)}\right)^2\right]^{-\frac{\nu+1}{2}}, & \text{if } y > 0, \end{cases}$$

$$\phi_W(t) = K(\nu)\alpha\sqrt{\pi\nu}\Psi\left(\frac{1}{2}, \frac{3}{2} - \frac{\nu+1}{2}; 2\nu\alpha^2 t^2\right)$$
$$+K(\nu)(1-\alpha)\sqrt{\pi\nu}\Psi\left(\frac{1}{2}, \frac{3}{2} - \frac{\nu+1}{2}; 2\nu(1-\alpha)^2 t^2\right)$$

for $-\infty < y < +\infty$, $\nu > 0$ and $0 < \alpha < 1$, where $K(\nu) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)}$.

Asymmetric skewed Student's t distribution (Zhu and Galbraith, 2010): for this distribution,

$$f_Y(y) = \begin{cases} \frac{\alpha}{\alpha^*}K(\nu_1)\left[1+\frac{1}{\nu_1}\left(\frac{y}{2\alpha^*}\right)^2\right]^{-\frac{\nu_1+1}{2}}, & \text{if } y \leq 0, \\ \\ \frac{1-\alpha}{1-\alpha^*}K(\nu_2)\left[1+\frac{1}{\nu_2}\left(\frac{y}{2(1-\alpha^*)}\right)^2\right]^{-\frac{\nu_2+1}{2}}, & \text{if } y > 0, \end{cases}$$

$$\phi_W(t) = \alpha\sqrt{\pi\nu_1}K(\nu_1)\Psi\left(\frac{1}{2}, \frac{3}{2} - \frac{\nu_1+1}{2}; 2\nu_1(\alpha^*)^2\right)$$
$$+(1-\alpha)\sqrt{\pi\nu_2}K(\nu_2)\Psi\left(\frac{1}{2}, \frac{3}{2} - \frac{\nu_2+1}{2}; 2\nu_2(1-\alpha^*)^2 t^2\right)$$

for $-\infty < y < +\infty$, $\nu_1 > 0$, $\nu_2 > 0$ and $0 < \alpha < 1$, where $K(\nu) = \frac{\Gamma((\nu+1)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)}$ and $\alpha^* = \frac{\alpha K(\nu_1)}{\alpha K(\nu_1)+(1-\alpha)K(\nu_2)}$.

Rice distribution (Rice, 1945): for this distribution,

$$f_Y(y) = \frac{y}{\sigma^2}\exp\left(-\frac{y^2+\nu^2}{2\sigma^2}\right)I_0\left(\frac{\nu y}{\sigma^2}\right),$$
$$\phi_W(t) = \frac{1}{1+\sigma^2 t^2}\exp\left[-\frac{\nu^2}{2\sigma^2}\right]{}_1F_1\left(1; 1; \frac{\nu^2}{2\sigma^2+2\sigma^4 t^2}\right)$$

for $y > 0$, $\sigma > 0$ and $\nu \geq 0$.

Symmetric Laplace distribution (Laplace, 1774): for this distribution,

$$f_Y(y) = \frac{1}{2\lambda}\exp\left(-\frac{\mid y \mid}{\lambda}\right),$$
$$\phi_W(t) = \frac{\sqrt{\pi}}{\sqrt{2}\lambda t}\exp\left(\frac{1}{2\lambda^2 t^2}\right)\mathrm{erfc}\left(\frac{1}{\sqrt{2}\lambda t}\right)$$

for $-\infty < y < +\infty$ and $\lambda > 0$.

Laplace distribution (Laplace, 1774): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{1}{2b} \exp\left(-\frac{\mid y-\mu \mid}{b}\right),\\
\phi_W(t) &= \frac{1}{2bt} \exp\left[\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(\frac{1}{b} + \frac{\mu t^2}{2}\right)\right] D_{-1}\left(\frac{1-\mu b t^2}{bt}\right)\\
&\quad + \frac{1}{2bt} \exp\left[\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(-\frac{1}{b} + \frac{\mu t^2}{2}\right)\right] D_{-1}\left(\frac{\mu b t^2 + 1}{bt}\right)
\end{aligned}
$$

for $-\infty < y < +\infty$, $-\infty < \mu < +\infty$ and $b > 0$.

Asymmetric Laplace distribution (Kozubowski and Podgorski, 2000): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{1}{\kappa + \frac{1}{\kappa}} \begin{cases} \exp\left[\frac{\lambda}{\kappa}(y-m)\right], & \text{if } y < m,\\ \\ \exp\left[\lambda\kappa(y-m)\right], & \text{if } y \geq m, \end{cases}\\
\phi_W(t) &= \frac{1}{\left(\kappa + \frac{1}{\kappa}\right)t} \exp\left[\frac{\lambda^2}{4\kappa^2t^2} - \frac{m}{2}\left(\frac{\lambda}{\kappa} + \frac{mt^2}{2}\right)\right] D_{-1}\left(\frac{\lambda}{\kappa t} - mt\right)\\
&\quad + \frac{1}{\left(\kappa + \frac{1}{\kappa}\right)t} \exp\left[\frac{\lambda^2\kappa^2}{4t^2} - \frac{m}{2}\left(\lambda\kappa + \frac{mt^2}{2}\right)\right] D_{-1}\left(mt - \frac{\lambda\kappa}{t}\right)
\end{aligned}
$$

for $-\infty < y < +\infty$, $-\infty < m < +\infty$, $\lambda > 0$ and $\kappa > 0$.

Poiraud-Casanova-Thomas-Agnan Laplace distribution (Poiraud-Casanova and Thomas-Agnan, 2000): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \begin{cases} \alpha(1-\alpha)\exp\left\{(1-\alpha)(y-\theta)\right\}, & \text{if } y \leq \theta,\\ \\ \alpha(1-\alpha)\exp\left\{\alpha(\theta-y)\right\}, & \text{if } y > \theta, \end{cases}\\
\phi_W(t) &= \frac{\alpha(1-\alpha)}{t} \exp\left[\frac{(1-\alpha)^2}{4t^2} - \frac{(1-\alpha)\theta}{2} - \frac{\theta^2t^2}{4}\right] D_{-1}\left(\frac{1-\alpha}{t} - \theta t\right)\\
&\quad + \frac{\alpha(1-\alpha)}{t} \exp\left[\frac{\alpha^2}{4t^2} + \frac{\alpha\theta}{2} - \frac{\theta^2t^2}{4}\right] D_{-1}\left(\theta t + \frac{\alpha}{t}\right)
\end{aligned}
$$

for $-\infty < y < +\infty$, $-\infty < \theta < +\infty$ and $0 < \alpha < 1$.

Holla-Bhattacharya Laplace distribution (Holla and Bhattacharya, 1968): for this distribution,

$$f_Y(y) = \begin{cases} a\phi \exp\left\{\phi\left(y-\theta\right)\right\}, & \text{if } y \le \theta, \\ \\ (1-a)\,\phi \exp\left\{\phi\left(\theta-y\right)\right\}, & \text{if } y > \theta, \end{cases}$$

$$\begin{aligned} \phi_W(t) = & \frac{a\phi}{t} \exp\left(\frac{\phi^2}{4t^2} - \frac{\theta\phi}{2} - \frac{\theta^2 t^2}{4}\right) D_{-1}\left(\frac{\phi}{t} - \theta t\right) \\ & + \frac{(1-a)\phi}{t} \exp\left(\frac{\phi^2}{4t^2} + \frac{\theta\phi}{2} - \frac{\theta^2 t^2}{4}\right) D_{-1}\left(\theta t + \frac{\phi}{t}\right) \end{aligned}$$

for $-\infty < y < +\infty$, $-\infty < \theta < +\infty$, $\phi > 0$ and $0 < a < 1$.

McGill Laplace distribution (McGill, 1962): for this distribution,

$$f_Y(y) = \begin{cases} \dfrac{1}{2\psi} \exp\left(\dfrac{y-\theta}{\psi}\right), & \text{if } y \le \theta, \\ \\ \dfrac{1}{2\phi} \exp\left(\dfrac{\theta-y}{\phi}\right), & \text{if } y > \theta, \end{cases}$$

$$\begin{aligned} \phi_W(t) = & \frac{1}{2\psi t} \exp\left(\frac{1}{4\psi^2 t^2} - \frac{\theta}{2\psi} - \frac{\theta^2 t^2}{4}\right) D_{-1}\left(\frac{1}{\psi t} - \theta t\right) \\ & + \frac{1}{2\phi t} \exp\left(\frac{1}{4\phi^2 t^2} + \frac{\theta}{2\phi} - \frac{\theta^2 t^2}{4}\right) D_{-1}\left(\theta t + \frac{1}{\phi t}\right) \end{aligned}$$

for $-\infty < y < +\infty$, $-\infty < \theta < +\infty$, $\phi > 0$ and $\psi > 0$.

Log Laplace distribution: for this distribution,

$$f_Y(y) = \frac{1}{2b} \begin{cases} y^{\frac{1}{b}-1} \exp\left(-\dfrac{\mu}{b}\right), & \text{if } y < \mu, \\ \\ y^{-\frac{1}{b}-1} \exp\left(\dfrac{\mu}{b}\right), & \text{if } y \ge \mu, \end{cases}$$

$$\phi_W(t) = b^{-1} t^{-\frac{1}{b}} 2^{\frac{1}{2b}-2} \exp\left(-\frac{\mu}{b}\right) \gamma\left(\frac{1}{2b}, \frac{\mu^2 t^2}{2}\right) + b^{-1} t^{\frac{1}{b}} 2^{-\frac{1}{2b}-2} \exp\left(\frac{\mu}{b}\right) \Gamma\left(-\frac{1}{2b}, \frac{\mu^2 t^2}{2}\right)$$

for $y > 0$, $b > 0$ and $\mu > 0$.

Exponential distribution: for this distribution,

$$\begin{aligned} f_Y(y) &= \lambda \exp(-\lambda y), \\ \phi_W(t) &= \frac{\sqrt{\pi}\lambda}{\sqrt{2}t} \exp\left(\frac{\lambda^2}{2t^2}\right) \operatorname{erfc}\left(\frac{\lambda}{\sqrt{2}t}\right) \end{aligned}$$

for $y > 0$ and $\lambda > 0$.

Chi distribution: for this distribution,

$$f_Y(y) = \frac{2^{1-\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} y^{k-1} \exp\left(-\frac{y^2}{2}\right),$$
$$\phi_W(t) = \left(1+t^2\right)^{-\frac{k}{2}}$$

for $y > 0$ and $k > 0$.

Variance gamma distribution (Madan and Seneta, 1990): for this distribution,

$$f_Y(y) = \frac{\mid \alpha \mid^{2\lambda}}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-\frac{1}{2}}} \mid y \mid^{\lambda-\frac{1}{2}} K_{\lambda-\frac{1}{2}}(\alpha \mid y \mid),$$
$$\phi_W(t) = \frac{\mid \alpha \mid^{2\lambda} t^{-\lambda-\frac{1}{2}} 2^{3/4-\lambda/2}}{\alpha^{\lambda+\frac{1}{2}}} exp\left[\frac{\alpha^2}{4t^2}\right] W_{-\frac{1}{4}-\frac{\lambda}{2},\frac{\lambda}{2}-\frac{1}{4}}\left(\frac{\alpha^2}{2t^2}\right)$$

for $-\infty < y < +\infty$, $-\infty < \alpha < +\infty$ and $\lambda > 0$.

Normal inverse gamma distribution: for this distribution,

$$f_Y(y) = \frac{\sqrt{\lambda}\beta^\alpha}{\sqrt{2\pi}\sigma^{2\alpha+3}\Gamma(\alpha)} \exp\left[-\frac{2\beta+\lambda(y-\mu)^2}{2\sigma^2}\right],$$
$$\phi_W(t) = \frac{\sqrt{\lambda}\beta^\alpha}{\sigma^{2\alpha+2}\Gamma(\alpha)} \frac{1}{\sqrt{\sigma^2 t^2+\lambda}} \exp\left[-\frac{2\beta+\lambda\mu^2}{2\sigma^2} + \frac{\lambda^2\mu^2}{2\sigma^2\left(\lambda+\sigma^2 t^2\right)}\right]$$

for $-\infty < y < +\infty$, $\lambda > 0$, $\alpha > 0$ and $\beta > 0$.

Nakagami distribution (Nakagami, 1960): for this distribution,

$$f_Y(y) = \frac{2m^m}{\Omega^m\Gamma(m)} y^{2m-1} \exp\left[-\frac{my^2}{\Omega}\right],$$
$$\phi_W(t) = \frac{m^m}{\Omega^m}\left(\frac{t^2}{2}+\frac{m}{\Omega}\right)^{-m}$$

for $y > 0$, $m > 0$ and $\Omega > 0$.

Reciprocal distribution (Hamming, 1970): for this distribution,

$$f_Y(y) = \frac{C}{y},$$
$$\phi_W(t) = \frac{C}{2}\left[\Gamma\left(0, \frac{t^2a^2}{2}\right) - \Gamma\left(0, \frac{t^2b^2}{2}\right)\right]$$

for $0 < a < y < b$, where C denotes the normalizing constant.

Maxwell distribution (Maxwell, 1860): for this distribution,

$$f_Y(y) = \sqrt{\frac{2}{\pi}}\frac{1}{a^3}y^2 \exp\left(-\frac{y^2}{2a^2}\right),$$
$$\phi_W(t) = \frac{1}{a^3}\left(t^2 + \frac{1}{a^2}\right)^{-\frac{3}{2}}$$

for $y > 0$ and $a > 0$.

Quadratic distribution: for this distribution,

$$f_Y(y) = \alpha(y-\beta)^2,$$
$$\phi_W(t) = \frac{\sqrt{2}\alpha}{t^3}\left[\Gamma\left(\frac{3}{2}, \frac{t^2a^2}{2}\right) - \Gamma\left(\frac{3}{2}, \frac{t^2b^2}{2}\right)\right] - \frac{2\alpha\beta}{t^2}\left[\exp\left(-\frac{t^2a^2}{2}\right) - \exp\left(-\frac{t^2b^2}{2}\right)\right] + \frac{\alpha\beta^2}{\sqrt{2}t}\left[\Gamma\left(\frac{1}{2}, \frac{t^2a^2}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{t^2b^2}{2}\right)\right]$$

for $-\infty < a < y < b < +\infty$, where $\beta = \frac{a+b}{2}$ and $\alpha = \frac{12}{(b-a)^3}$.

Uniform distribution: for this distribution,

$$f_Y(y) = \frac{1}{b-a},$$
$$\phi_W(t) = \frac{1}{(b-a)\sqrt{2}t}\left[\Gamma\left(\frac{1}{2}, \frac{a^2t^2}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{b^2t^2}{2}\right)\right]$$

for $-\infty < a < y < b < +\infty$.

Power function distribution: for this distribution,

$$f_Y(y) = ay^{a-1},$$
$$\phi_W(t) = a2^{\frac{a}{2}-1}t^{-a}\gamma\left(\frac{a}{2}, \frac{t^2}{2}\right)$$

for $0 < y < 1$ and $a > 0$.

Rayleigh distribution (Weibull, 1951): for this distribution,

$$f_Y(y) = 2\lambda^2 y \exp\left(-\lambda^2 y^2\right),$$
$$\phi_W(t) = \frac{2\lambda^2}{t^2 + 2\lambda^2}$$

for $y > 0$ and $\lambda > 0$.

Exponentiated Rayleigh distribution (Kundu and Raqab, 2005): for this distribution,

$$
\begin{aligned}
f_Y(y) &= 2\alpha\lambda^2 y \exp\left(-\lambda^2 y^2\right)\left[1-\exp\left(-\lambda^2 y^2\right)\right]^{\alpha-1}, \\
\phi_W(t) &= \alpha B\left(1+\frac{t^2}{2\lambda^2}, \alpha\right)
\end{aligned}
$$

for $y > 0$, $\alpha > 0$ and $\lambda > 0$.

Beta Rayleigh distribution (Kundu and Raqab, 2005): for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{2\alpha\lambda^2}{B(\alpha,\beta)} y \exp\left(-\beta\lambda^2 y^2\right)\left[1-\exp\left(-\lambda^2 y^2\right)\right]^{\alpha-1}, \\
\phi_W(t) &= \frac{\alpha}{B(\alpha,\beta)} B\left(\beta+\frac{t^2}{2\lambda^2}, \alpha\right)
\end{aligned}
$$

for $y > 0$, $\alpha > 0$, $\beta > 0$ and $\lambda > 0$.

Truncated normal distribution: for this distribution,

$$
\begin{aligned}
f_Y(y) &= \frac{1}{\sqrt{2\pi}\sigma\left[\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)\right]} \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right], \\
\phi_W(t) &= \frac{\Phi\left(\dfrac{b\left(\sigma^2t^2+1\right)-\mu}{\sigma\sqrt{\sigma^2t^2+1}}\right)-\Phi\left(\dfrac{a\left(\sigma^2t^2+1\right)-\mu}{\sigma\sqrt{\sigma^2t^2+1}}\right)}{\left[\Phi\left(\dfrac{b-\mu}{\sigma}\right)-\Phi\left(\dfrac{a-\mu}{\sigma}\right)\right]} \frac{1}{\sqrt{1+\sigma^2t^2}} \exp\left[-\frac{\mu^2t^2}{2\left(\sigma^2t^2+1\right)}\right]
\end{aligned}
$$

for $a < y < b$, $-\infty < \mu < +\infty$ and $\sigma > 0$.

Split normal distribution: for this distribution,

$$
\begin{aligned}
f_Y(y) &= C\begin{cases} \exp\left[-\dfrac{(y-\mu)^2}{2\sigma_1^2}\right], & \text{if } y < \mu, \\[2ex] \exp\left[-\dfrac{(y-\mu)^2}{2\sigma_2^2}\right], & \text{if } y \geq \mu, \end{cases} \\
\phi_W(t) &= \frac{C\sigma_1}{\sqrt{\sigma_1^2t^2+1}} \exp\left(\frac{-2\mu^2t^2-\mu^2\sigma_1^2t^4}{4\left(1+\sigma_1^2t^2\right)}\right) D_{-1}\left(\frac{-\mu\sigma_1t^2}{\sqrt{\sigma_1^2t^2+1}}\right) \\
&\quad + \frac{C\sigma_2}{\sqrt{\sigma_2^2t^2+1}} \exp\left(\frac{-2\mu^2t^2-\mu^2\sigma_2^2t^4}{4\left(1+\sigma_2^2t^2\right)}\right) D_{-1}\left(\frac{\mu\sigma_2t^2}{\sqrt{\sigma_2^2t^2+1}}\right)
\end{aligned}
$$

for $-\infty < y < +\infty$, $-\infty < \mu < +\infty$, $\sigma_1 > 0$ and $\sigma_2 > 0$, where C denotes the normalizing constant.

q-Gaussian distribution (Tsallis, 2009): for this distribution,

$$f_Y(y) = \frac{\sqrt{\beta}}{C}\left[1-(1-q)\beta y^2\right]^{\frac{1}{1-q}},$$
$$\phi_W(t) = \begin{cases} \dfrac{\sqrt{2}}{C\sqrt{q-1}}\Psi\left(\dfrac{1}{2}, \dfrac{3}{2}-\dfrac{1}{q-1}; \dfrac{t^2}{2\beta(q-1)}\right), & \text{if } 1 \leq q < 3, \\ \\ \dfrac{1}{C\sqrt{1-q}} B\left(\dfrac{1}{2}, \dfrac{2-q}{1-q}\right) {}_1F_1\left(\dfrac{1}{2}, \dfrac{1}{2}+\dfrac{2-q}{1-q}; \dfrac{-t^2}{2(1-q)\beta}\right), & \text{if } q < 1 \end{cases}$$

for $-\infty < y < +\infty$ if $1 \leq q < 3$, $-\frac{1}{\sqrt{\beta(1-q)}} < y < +\frac{1}{\sqrt{\beta(1-q)}}$ if $q < 1$ and $\beta > 0$.

Normal exponential gamma distribution: for this distribution,

$$f_Y(y) = C\exp\left(-\frac{y^2}{4\theta^2}\right) D_{-2k-1}\left(\frac{\mid y \mid}{\theta}\right),$$
$$\phi_W(t) = C\sqrt{\pi}\theta 2^{-k}\Gamma^{-1}\left(k+\frac{3}{2}\right) {}_2F_1\left(\frac{1}{2}, 1; k+\frac{3}{2}; -t^2\theta^2\right)$$

for $-\infty < y < +\infty$, $k > 0$ and $\theta > 0$, where C denotes the normalizing constant.

Wigner semicircle distribution: for this distribution,

$$f_Y(y) = \frac{2\sqrt{R^2-y^2}}{\pi R^2},$$
$$\phi_W(t) = {}_1F_1\left(\frac{1}{2}; 2; -\frac{t^2R^2}{2}\right)$$

for $-R < y < R$.

Kumaraswamy distribution (Kumaraswamy, 1980): for this distribution,

$$f_Y(y) = 2ay\left(1-y^2\right)^{a-1},$$
$$\phi_W(t) = a\left(-\frac{t^2}{2}\right)^{-a}\exp\left(-\frac{t^2}{2}\right)\gamma\left(a, -\frac{t^2}{2}\right)$$

for $0 < y < 1$ and $a > 0$.

Linear failure rate distribution (Bain, 1974): for this distribution,

$$f_Y(y) = (a+by)\exp\left(-ay-\frac{by^2}{2}\right),$$
$$\phi_W(t) = \frac{a}{\sqrt{b+t^2}}\exp\left[\frac{a^2}{4\left(b+t^2\right)}\right] D_{-1}\left(\frac{a}{\sqrt{b+t^2}}\right) + \frac{b}{b+t^2}\exp\left[\frac{a^2}{4\left(b+t^2\right)}\right] D_{-2}\left(\frac{a}{\sqrt{b+t^2}}\right)$$

for $y > 0$, $a > 0$ and $b \geq 0$.

Irwin Hall distribution (Irwin, 1927; Hall, 1927): for this distribution,

$$f_Y(y) = \frac{1}{2(n-1)!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}(y-k)^{n-1}\mathrm{sign}(y-k),$$

$$\phi_W(t) = \frac{1}{2(n-1)!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\sum_{m=0}^{n-1}\binom{n-1}{m}(-k)^{n-1-m}2^{\frac{m-1}{2}}t^{-m-1}\gamma\left(\frac{m+1}{2},\frac{n^2t^2}{2}\right)$$
$$-\frac{1}{2(n-1)!}\sum_{k=0}^{n}(-1)^k\binom{n}{k}\sum_{m=0}^{n-1}\binom{n-1}{m}(-k)^{n-1-m}2^{\frac{m+1}{2}}t^{-m-1}\gamma\left(\frac{m+1}{2},\frac{k^2t^2}{2}\right)$$

for $0 < y < n$ and $n \geq 1$.

7.4 Simulation results

Here, we perform simulations to check the mathematical derivations for Section 7.3. We simulated the distribution of W for given distributions of X and Y as follows:

1. simulate 1000 random numbers from the distribution of X;
2. simulate 1000 random numbers from the distribution of Y;
3. set $W = XY$;
4. construct a histogram of the 1000 values of W.

The simulated histograms can be compared to the theoretical probability density functions of W computed by the method of inverse Fourier transform using the characteristic functions in Section 7.3 in software Mathematica.

The comparisons are illustrated in Figures 7.1 to 7.5 for five of the distributions of Y considered in Section 7.3: Figure 7.1 for the exponential distribution; Figure 7.2 for the uniform distribution; Figure 7.3 for the power function distribution; Figure 7.4 for the Rayleigh distribution; Figure 7.5 for the exponentiated Rayleigh distribution.

We see that the simulated histogram and the theoretical probability density function agree well in each of the five figures. We have considered the five distributions for illustration. But the conclusions were the same for other distributions in Section 7.3.

7.5 Future work: density function of the product

Both the characteristic function and the density function define a random variable's probability distribution. Benefit from the formula of Rohaty, we can also obtain the density function of the product $W = X \times Y$.

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x) f_Y\left(\frac{w}{x}\right)\frac{1}{|x|}dx$$

Here, we present two derivations of the density function of $W = XY$ as examples when: i) X is a standard normal random variable and Y is an independent Pareto random variable with density function $f_Y(y) = \alpha K^\alpha y^{-\alpha-1}$; ii) X is a standard normal random variable and Y is an independent Chi distributed random variable with density function $f_Y(y) = \frac{2^{1-\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} y^{k-1}\exp\left(-\frac{y^2}{2}\right)$.

Example 1: Y is an independent Pareto random variable.

$$\begin{aligned}
f_W(w) &= \int_{-\infty}^{\infty} f_X(x) f_Y\left(\frac{w}{x}\right)\frac{1}{|x|}dx \\
&= \int_{-\infty}^{0} f_X(x) f_Y\left(\frac{w}{x}\right)\frac{-1}{x}dx + \int_{0}^{\infty} f_X(x) f_Y\left(\frac{w}{x}\right)\frac{1}{x}dx \\
&= \int_{-\infty}^{0} \frac{-1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\cdot\alpha K^\alpha\left(\frac{w}{x}\right)^{-\alpha-1}dx \\
&\quad + \int_{0}^{\infty} \frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\cdot\alpha K^\alpha\left(\frac{w}{x}\right)^{-\alpha-1}dx \\
&= \frac{1}{\sqrt{2\pi}}\alpha K^\alpha w^{-\alpha-1}\left(\int_0^\infty x^\alpha\exp\left(-\frac{x^2}{2}\right)dx - \int_{-\infty}^0 x^\alpha\exp\left(-\frac{x^2}{2}\right)dx\right) \\
&= \frac{1}{\sqrt{2\pi}}\alpha K^\alpha w^{-\alpha-1}\left(1-(-1)^\alpha\right)\int_0^\infty x^\alpha\exp\left(-\frac{x^2}{2}\right)dx.
\end{aligned}$$

Because $\int_0^\infty x^{\alpha-1}\exp\left(-px^\mu\right)dx = \frac{1}{\mu}p^{-\alpha/\mu}\Gamma\left(\frac{\alpha}{\mu}\right)$, here $\alpha = \alpha+1$, $p = \frac{1}{2}$, $\mu = 2$.

$$\begin{aligned}
f_W(w) &= \frac{1}{\sqrt{2\pi}}\alpha K^\alpha w^{-\alpha-1}\frac{1}{2}\left(\frac{1}{2}\right)^{-\frac{\alpha+1}{2}}\Gamma\left(\frac{\alpha+1}{2}\right) \\
&= \frac{\alpha 2^{-\alpha-1}K^\alpha}{\sqrt{\pi}}w^{-\alpha-1}\Gamma\left(\frac{\alpha+1}{2}\right).
\end{aligned}$$

Example 2: Y is an independent Chi distributed random variable.

$$\begin{aligned}
f_W(w) &= \int_{-\infty}^{\infty} f_X(x) f_Y\left(\frac{w}{x}\right) \frac{1}{|x|} dx \\
&= \int_{-\infty}^{\infty} \frac{1}{|x|} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{2^{1-\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{w}{x}\right)^{k-1} \exp\left(-\frac{w^2}{2}x^{-2}\right) dx \\
&= \frac{2^{1-\frac{k}{2}} w^{k-1}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} \int_{-\infty}^{0} \frac{1}{-x} \left(\frac{1}{x}\right)^{k-1} \exp\left(-\frac{x^2}{2} - \frac{w^2}{2}x^{-2}\right) dx \\
&\quad + \frac{2^{1-\frac{k}{2}} w^{k-1}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} \int_{0}^{\infty} \frac{1}{x} \left(\frac{1}{x}\right)^{k-1} \exp\left(-\frac{x^2}{2} - \frac{w^2}{2}x^{-2}\right) dx \\
&= \frac{2^{1-\frac{k}{2}} w^{k-1}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} [(-1)^k \int_{0}^{\infty} \frac{1}{z} \left(\frac{1}{z}\right)^{k-1} \exp\left(-\frac{z^2}{2} - \frac{w^2}{2}z^{-2}\right) dz \\
&\quad + \int_{0}^{\infty} \frac{1}{x} \left(\frac{1}{x}\right)^{k-1} \exp\left(-\frac{x^2}{2} - \frac{w^2}{2}x^{-2}\right) dx]
\end{aligned}$$

If k is a odd value, $f_W(w) = 0$. If k is a even value,

$$f_W(w) = \frac{2^{2-\frac{k}{2}} w^{k-1}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} \int_{0}^{\infty} x^{-k} \exp\left(-\frac{1}{2}x^2 - \frac{w^2}{2}x^{-2}\right) dx$$

Because $\int_0^{\infty} x^{\nu-1} \exp\left(-\beta x^p - \gamma x^{-p}\right) dx = \frac{2}{p}\left(\frac{\gamma}{\beta}\right)^{\frac{\nu}{2p}} K_{\frac{\nu}{p}}\left(2\sqrt{\beta\gamma}\right)$, here, $\nu = 1-k$, $\beta = \frac{1}{2}$,p=2,$\gamma = \frac{w^2}{2}$.

$$\begin{aligned}
f_W(w) &= \frac{2^{2-\frac{k}{2}} w^{k-1}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} \left(w^2\right)^{\frac{1-k}{4}} K_{\frac{1-k}{2}}(w) \\
&= \frac{2^{2-\frac{k}{2}} w^{\frac{k-1}{2}}}{\sqrt{2\pi}\Gamma\left(\frac{k}{2}\right)} K_{\frac{1-k}{2}}(w).
\end{aligned}$$

As we can see, the explicit form of the density functions of some products can be derived by the formula of Rohaty. However, there are two points we need to consider before applying this method.

1. Because there is an absolute item $\frac{1}{|x|}$ in Rohaty's formula, we need to split the formula into two integrals due to the domain of x, i.e. $f_W(w) = \int_{-\infty}^{0} \frac{1}{-x} f_X(x) f_Y\left(\frac{w}{x}\right) dx + \int_0^{\infty} \frac{1}{x} f_X(x) f_Y\left(\frac{w}{x}\right) dx$. This may require accurate substitution in derivations.

2. If there is an exponential function in the density function of Y, we need to solve the integral $\int A(x) \exp\left(ax^p + bx^{-q}\right) dx$, where $A(x)$ is a function of x, a and b are real values, $p > 0$, $q > 0$. If $p \neq q$, we can not derive the closed expression of the integral.

Hence, it remains further investigation in finding the sufficient method for the derivation of the density function of the product distribution.

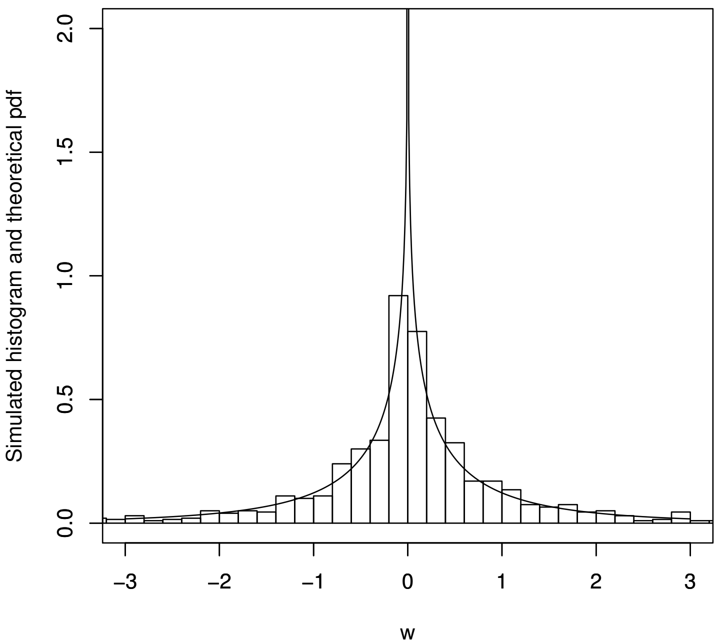

Figure 7.1: Simulated histogram and theoretical probability density function of $W = XY$ when X is a standard normal random variable and Y is an independent unit exponential random variable.

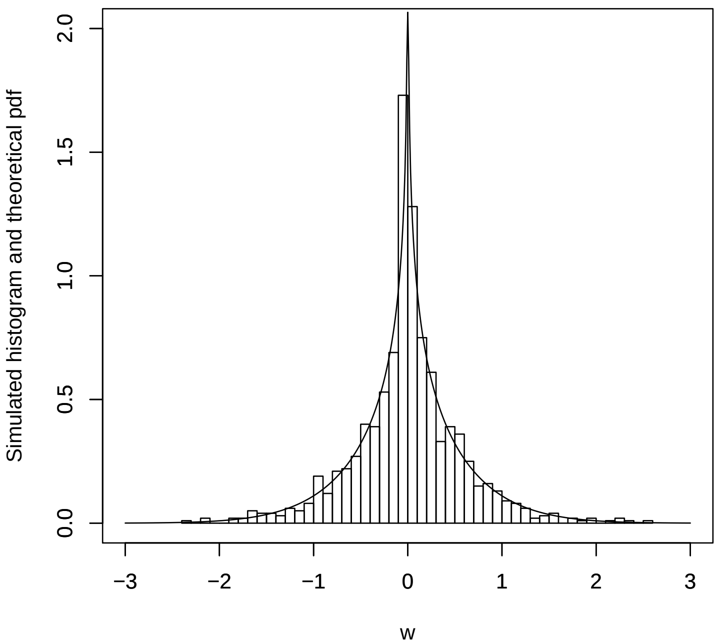

Figure 7.2: Simulated histogram and theoretical probability density function of $W = XY$ when X is a standard normal random variable and Y is an independent uniform $[0, 1]$ random variable.

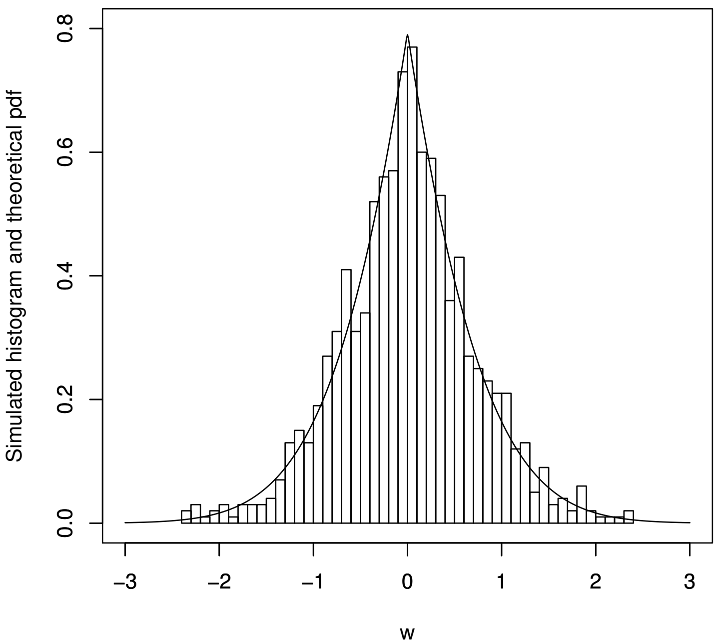

Figure 7.3: Simulated histogram and theoretical probability density function of $W = XY$ when X is a standard normal random variable and Y is an independent power function random variable with $a = 2$.

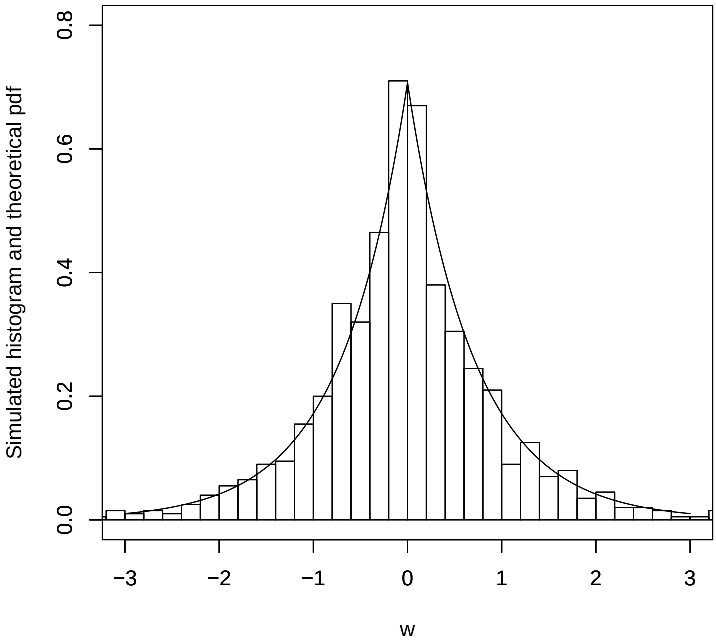

Figure 7.4: Simulated histogram and theoretical probability density function of $W = XY$ when X is a standard normal random variable and Y is an independent Rayleigh random variable with $\lambda = 1$.

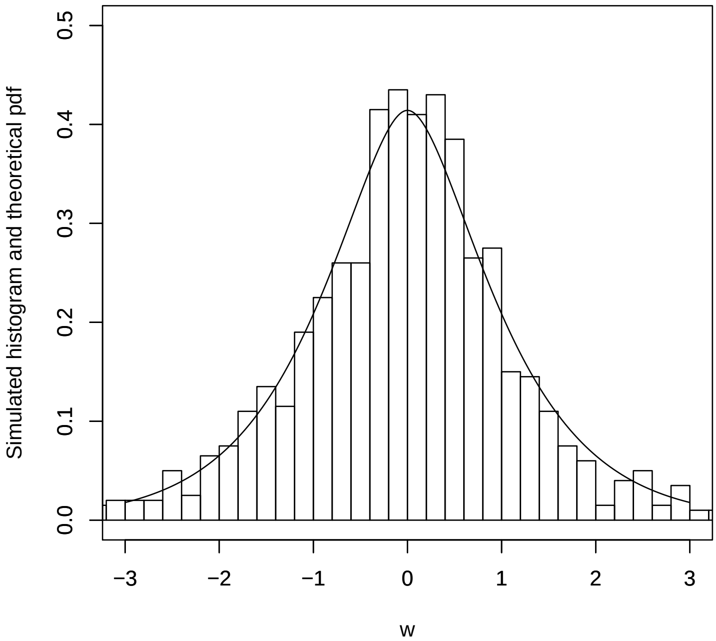

Figure 7.5: Simulated histogram and theoretical probability density function of $W = XY$ when X is a standard normal random variable and Y is an independent exponentiated Rayleigh random variable with $\alpha = 2$, $\lambda = 1$.

Chapter 8

Distribution of Aggregated risk and its TVaR

8.1 Introduction

Results related to the sum of dependent risks are of interest in the calculation of the accumulated risks for portfolio investment and risk measures for decision making. Besides, strategic planning also requires the knowledge of the cumulative distribution function of the sum of dependent random variables (Cossette et al., 2015). Therefore, risk measures like Value at Risk and Tail-Value at Risk are worthwhile being explored.

In recent years, several closed-form expressions for the distribution of aggregate risks, its TVaR and TVaR based allocations have been developed, based on an allocation method due to Tasche (1999). These expressions are based on a given joint distribution between the components of a portfolio. The joint distributions considered so far are the: multivariate normal distribution (Panjer, 2002); multivariate elliptical distributions (Landsman and Valdez, 2003; Dhaene et al., 2008); multivariate gamma distribution (Furman and Landsman, 2005); multivariate Tweedie distribution (Furman and Landsman, 2008); multivariate Pareto distribution (Chiragiev and Landsman, 2007); Farlie-Gumbel-Morgenstern copula (Barges et al., 2009); Farlie-Gumbel-Morgenstern copula with mixed Erlang marginals (Cossette et al., 2013); multivariate compound distributions (Cossette et al., 2012); bivariate exponential and bivariate mixed Erlang distributions (Cossette et al., 2015).

The distribution of bivariate aggregate risks can be applied for many practical uses. Here is an example in automobile insurance. Let $S = X + Y$ and suppose X and Y are the total claims from safe drivers and high-risk drivers in a fixed time period. Assume that $X = r_0 + r_1$ and $Y = r_0 + r_2$. r_0 is the random variable which stands for the amount of claims due to some objective reasons, e.g. car condition, weather condition, road condition, etc. r_1 and r_2 are the random variables that stand for the amount of claims due to subjective reasons in X and Y repectively, e.g. driving behavior, number of overspeed, etc. We can see there is a correlation of X and Y due to some common affected events. Distribution of S is important to a insurer in estimating the total amount of claims from X and Y. An insurer can also optimize the insurance contract by understanding the contribution of each group to the aggregated loss. More details on the bivariate aggregate risks distributions can be found in Hesselager (1994).

The aim of this chapter is to derive expressions for the distribution of aggregate risks, its TVaR and TVaR based allocations for a comprehensive collection of bivariate distributions. We consider thirty-three families of bivariate distributions each defined on $(0, \infty) \times (0, \infty)$ or $(\beta, \infty) \times (\beta, \infty)$ for some $\beta > 0$. They include mixtures of independent exponential distributions, Mirhosseini et al. (2015)'s bivariate exponential distribution, Crovelli (1973)'s bivariate exponential distribution, Gumbel's bivariate exponential distribution, Lawrance and Lewis' (1980) bivariate exponential distribution, Block and Basu (1976)'s bivariate exponential distribution, Arnold and Strauss (1991)'s bivariate exponential distribution, mixtures of independent gamma distributions with real shape parameters, Nadarajah and Gupta (2006)'s bivariate gamma distribution with equal scale parameters, Nadarajah and Gupta (2006)'s bivariate gamma distribution with unequal scale parameters, Nagar and Sepulveda-Murillo (2011)'s bivariate confluent hypergeometric distribution, Becker and Roux (1981)'s bivariate gamma distribution, Mohsin et al. (2013)'s bivariate gamma distribution, Cheriyan (1941)'s bivariate gamma distribution, Dussauchoy and Berland (1975)'s bivariate gamma distribution, mixtures of independent two piece gamma distributions, beta Stacey distribution in equation (5.38) of Balakrishnan and Lai (2009), Mardia (1970)'s bivariate distributions, bivariate Liouville distributions, bivariate equilibrium distributions due to Unnikrishnan Nair and Sankaran (2014), Chacko and Thomas (2007)'s bivariate Pareto distribution, bivariate Pareto distribution in equation (10.68) of Balakrishnan and Lai (2009), bivariate Pareto distribution with equal scale parameters, bivariate Pareto distribution with unequal scale parameters, mixtures of independent Pareto distributions, mixtures of bivariate Pareto distributions with equal scale parameters, mixtures of bivariate Pareto distributions with unequal scale parameters, generalized bivariate Pareto distribution, Lee and Cha (2014)'s bivariate distribution and truncated bivariate normal distribution. We have not considered the bivariate normal or other distributions defined over $(-\infty, \infty) \times (-\infty, \infty)$, as they have already been considered by others.

8.2 Mathematical notation

Let (X, Y) be non-negative continuous risks with joint probability density function $f(x, y)$. Let $S = X + Y$ denote the aggregated risk. We are interested in: the probability density function of S given by

$$f_S(s) = \int_0^s f(x, s - x)dx; \tag{8.1}$$

the cumulative distribution function of S given by

$$F_S(s) = \int_0^s f_S(t)dt; \tag{8.2}$$

the truncated expectation of S given by

$$E\left[S 1_{\{S>b\}}\right] = \int_b^\infty s f_S(s)ds; \tag{8.3}$$

the contribution of each risk to the aggregated risk given by

$$E\left[X 1_{\{S>b\}}\right] = \int_b^\infty g_{X,S}(s)ds, \tag{8.4}$$

where

$$g_{X,S}(s) = \int_0^s x f(x, s - x)dx. \tag{8.5}$$

We derive expressions for (8.1)-(8.5) for thirty-three bivariate distributions.

The derived expressions given in Section 8.3 involve several special functions, including the gamma function defined by

$$\Gamma(a) = \int_0^\infty t^{a-1} \exp(-t)dt;$$

the incomplete gamma function defined by

$$\gamma(a, x) = \int_0^x t^{a-1} \exp(-t)dt;$$

the complementary incomplete gamma function defined by

$$\Gamma(a,x) = \int_x^{\infty} t^{a-1} \exp(-t) dt;$$

the error function defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-t^2\right) dt;$$

the beta function defined by

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt;$$

the incomplete beta function defined by

$$B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1} dt;$$

the standard normal distribution function defined by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt;$$

the confluent hypergeometric function defined by

$${}_1F_1\left(a;b;x\right) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{x^k}{k!};$$

the Whittaker W function of orders ν, μ defined by

$$W_{\nu,\mu}(x) = \frac{x^{\mu+\frac{1}{2}}}{\Gamma\left(\mu-\nu+\frac{1}{2}\right)} \int_{\frac{1}{2}}^{+\infty} \left(t-\frac{1}{2}\right)^{\mu-\nu-\frac{1}{2}} \left(t+\frac{1}{2}\right)^{\mu+\nu-\frac{1}{2}} \exp\left(-xt\right) dt$$

for $\mu-\nu > -\frac{1}{2}$ and $x > 0$; the Gauss hypergeometric function defined by

$${}_2F_1\left(a;b;x\right) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!};$$

the Appell hypergeometric function of the first kind of two variables defined by

$$F_1\left(a,b,c,d;x,y\right) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n}(b)_m(c)_n x^m y^n}{(d)_{m+n} m! n!};$$

and the degenerate hypergeometric series of two variables defined by

$$\Phi_1(a, b, c, x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n}(b)_n x^m y^n}{(c)_{m+n} m! n!},$$

where $(e)_k = e(e+1) \cdots (e+k-1)$ denotes the ascending factorial.

In-built routines for computing these special functions are available in packages like Matlab, Maple and Mathematica. For example, Erf[x], Gamma[a, x], Beta[a, b], Beta[x, a, b], Hypergeometric1F1[a, b, x], Hypergeometric2F1[a, b, c, x] and AppellF1[a, b, c, d, x, y] in Mathematica compute the error, complementary incomplete gamma, beta, incomplete beta, confluent hypergeometric, Gauss hypergeometric and Appell hypergeometric functions. Mathematica allows for arbitrary precision, so the accuracy of computation is not an issue.

8.3 The collection

Here, we give expressions for (8.1)-(8.5) for thirty three bivariate distributions. We also select the representative models and provide the details of derivation in the Appendix D.

Mixtures of independent exponential distributions:

$$\begin{aligned}
f(x, y) &= \sum_{k=1}^{m} C_k \exp(-\alpha_k x - \beta_k y), \\
f_S(s) &= \sum_{k=1}^{m} \frac{C_k}{\beta_k - \alpha_k} \left[\exp(-\alpha_k s) - \exp(-\beta_k s)\right], \\
F_S(s) &= \sum_{k=1}^{m} \frac{C_k}{\beta_k - \alpha_k} \left[\frac{1 - \exp(-\alpha_k s)}{\alpha_k} - \frac{1 - \exp(-\beta_k s)}{\beta_k}\right], \\
E\left[S 1_{\{S>b\}}\right] &= \sum_{k=1}^{m} \frac{C_k}{\beta_k - \alpha_k} \left[\frac{(b\alpha_k + 1)\exp(-b\alpha_k)}{\alpha_k^2} - \frac{(b\beta_k + 1)\exp(-b\beta_k)}{\beta_k^2}\right], \\
g_{X,S}(s) &= \sum_{k=1}^{m} C_k \exp(-\beta_k s) \left[\frac{1 - \exp(\beta_k s - \alpha_k s)}{(\alpha_k - \beta_k)^2} - \frac{s \exp(\beta_k s - \alpha_k s)}{\alpha_k - \beta_k}\right], \\
E\left[X 1_{\{S>b\}}\right] &= \sum_{k=1}^{m} C_k \left[-\frac{(\alpha_k b + 1)\exp(-\alpha_k b)}{\alpha_k^2 (\alpha_k - \beta_k)} + \frac{\exp(-\beta_k b)}{\beta_k (\alpha_k - \beta_k)^2} - \frac{\exp(-\alpha_k b)}{\alpha_k (\alpha_k - \beta_k)^2}\right]
\end{aligned}$$

for $\alpha_k > 0$, $\beta_k > 0$, $x > 0$ and $y > 0$.

Mirhosseini et al. (2015)'s bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= \lambda^2 \alpha \exp\left(-\lambda x - \lambda y\right)\left[1-\alpha \exp\left(-\lambda x - \lambda y\right)\right]\left[1-\exp\left(-\lambda x - \lambda y\right)\right]^{\alpha-2}, \\
f_S(s) &= \lambda^2 \alpha s \exp\left(-\lambda s\right)\left[1-\alpha \exp\left(-\lambda s\right)\right]\left[1-\exp\left(-\lambda s\right)\right]^{\alpha-2}, \\
F_S(s) &= \alpha \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+1)^2} \left[1-(1+(k+1)\lambda s)\exp\left(-(k+1)\lambda s\right)\right] \\
&\quad -\alpha^2 \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+2)^2} \left[1-(1+(k+2)\lambda s)\exp\left(-(k+2)\lambda s\right)\right], \\
E\left[S 1_{\{S>b\}}\right] &= \frac{2\alpha}{\lambda} \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+1)^3} \left(1+(k+1)\lambda b + \frac{(k+1)^2\lambda^2 b^2}{2}\right) \exp\left(-(k+1)\lambda b\right) \\
&\quad - \frac{2\alpha^2}{\lambda} \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+2)^3} \left(1+(k+2)\lambda b + \frac{(k+2)^2\lambda^2 b^2}{2}\right) \exp\left(-(k+2)\lambda b\right), \\
g_{X,S}(s) &= 2^{-1}\lambda^2 \alpha s^2 \exp\left(-\lambda s\right)\left[1-\alpha \exp\left(-\lambda s\right)\right]\left[1-\exp\left(-\lambda s\right)\right]^{\alpha-2}, \\
E\left[X 1_{\{S>b\}}\right] &= \frac{\alpha}{\lambda} \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+1)^3} \left(1+(k+1)\lambda b + \frac{(k+1)^2\lambda^2 b^2}{2}\right) \exp\left(-(k+1)\lambda b\right) \\
&\quad - \frac{\alpha^2}{\lambda} \sum_{k=0}^{\infty} \binom{\alpha-2}{k} \frac{(-1)^k}{(k+2)^3} \left(1+(k+2)\lambda b + \frac{(k+2)^2\lambda^2 b^2}{2}\right) \exp\left(-(k+2)\lambda b\right)
\end{aligned}
$$

for $\alpha > 0$, $\lambda > 0$, $x > 0$ and $y > 0$.

Crovelli (1973)'s bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= \begin{cases} \alpha\beta\exp(-\beta y)\left[1-\exp(-\alpha x)\right], & \text{if } 0 \le \alpha x \le \beta y, \\ \alpha\beta\exp(-\beta x)\left[1-\exp(-\beta y)\right], & \text{if } 0 \le \beta y \le \alpha x, \end{cases} \\
f_S(s) &= (\alpha+\beta)\exp\left(-\frac{\alpha\beta s}{\alpha+\beta}\right) + \frac{\alpha^2}{\alpha-\beta}\exp(-\beta s) - \frac{\beta^2}{\alpha-\beta}\exp(-\alpha s), \\
F_S(s) &= \frac{(\alpha+\beta)^2}{\alpha\beta}\left[1-\exp\left(-\frac{\alpha\beta s}{\alpha+\beta}\right)\right] + \frac{\alpha^2}{\beta(\beta-\alpha)}\left[1-\exp\left(-\beta s\right)\right] \\
&\quad - \frac{\beta^2}{\alpha(\beta-\alpha)}\left[1-\exp\left(-\alpha s\right)\right], \\
E\left[S1_{\{S>b\}}\right] &= \frac{(\alpha+\beta)^3}{(\alpha\beta)^2}\exp\left(-\frac{\alpha\beta b}{\alpha+\beta}\right)\left[1+\frac{\alpha\beta b}{\alpha+\beta}\right] + \frac{\alpha^2}{\beta(\alpha-\beta)}\exp(-\beta b) \\
&\quad - \frac{\beta^2}{\alpha(\alpha-\beta)}\exp(-\alpha b), \\
g_{X,S}(s) &= \frac{\beta^2}{\alpha-\beta}s\exp(-\alpha s) + \frac{\beta^2(2\alpha-\beta)}{\alpha(\alpha-\beta)^2}\exp(-\alpha s) + \frac{\alpha^2(\alpha-2\beta)}{\beta(\alpha-\beta)^2}\exp(-\beta s) \\
&\quad + \beta s\exp\left(-\frac{\alpha\beta s}{\alpha+\beta}\right) + \frac{-\alpha^4+2\alpha^3+\beta^4-2\alpha\beta^3}{\alpha\beta(\alpha-\beta)^2}\exp\left(-\frac{\alpha\beta s}{\alpha+\beta}\right), \\
E\left[X1_{\{S>b\}}\right] &= \frac{\beta^2}{\alpha^2(\alpha-\beta)}(1+\alpha b)\exp\left(-\alpha b\right) + \frac{\beta^2(2\alpha-\beta)}{\alpha^2(\alpha-\beta)^2}\exp\left(-\alpha b\right) \\
&\quad + \frac{\alpha^2(\alpha-2\beta)}{\beta^2(\alpha-\beta)^2}\exp\left(-\beta b\right) + \frac{(\alpha+\beta)^2}{\alpha^2\beta}\left(1+\frac{\alpha\beta b}{\alpha+\beta}\right)\exp\left(-\frac{\alpha\beta b}{\alpha+\beta}\right) \\
&\quad + \frac{(\alpha+\beta)\left(-\alpha^4+2\alpha^3+\beta^4-2\alpha\beta^3\right)}{(\alpha\beta)^2(\alpha-\beta)^2}\exp\left(-\frac{\alpha\beta b}{\alpha+\beta}\right)
\end{aligned}
$$

for $\alpha > 0$, $\beta > 0$, $x > 0$ and $y > 0$.

Gumbel's bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= [(1+\theta x)(1+\theta y)-\theta]\exp\left(-x-y-\theta xy\right), \\
f_S(s) &= -\mathrm{i}\theta^{3/2}\sqrt{\pi}\frac{s^2}{4}\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad -\mathrm{i}\sqrt{\theta\pi}s\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad +\frac{\mathrm{i}\sqrt{\theta\pi}}{2}\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad -\mathrm{i}\sqrt{\pi/\theta}\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right)-\frac{\theta}{2}s\exp\left(-s\right), \\
g_{X,S}(s) &= -\frac{\mathrm{i}\sqrt{\pi}\theta^{3/2}}{8}s^3\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad -\frac{\mathrm{i}3\sqrt{\theta\pi}}{4}s^2\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad +\frac{\mathrm{i}\sqrt{\theta\pi}}{2}s\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right) \\
&\quad -\frac{\mathrm{i}\sqrt{\pi/\theta}}{2}s\exp\left(-s-\frac{\theta s^2}{4}\right)\mathrm{erf}\left(\frac{\mathrm{i}\sqrt{\theta}}{2}s\right)+\left(\frac{1}{\theta}+s-\frac{\theta s^3}{4}\right)\exp\left(-s\right) \\
&\quad -\left(\frac{1}{\theta}+s+\frac{\theta s^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)
\end{aligned}
$$

for $0<\theta<1$, $x>0$ and $y>0$, where $\mathrm{i}=\sqrt{-1}$.

Lawrance and Lewis' (1980) bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= I\left\{\beta y<x<y/\beta\right\}\frac{1}{1+\beta}\exp\left(-\frac{x+y}{1+\beta}\right)+\frac{1}{\beta}\exp\left(-\frac{x+y}{\beta}\right), \\
f_S(s) &= \frac{(1-\beta)s}{(1+\beta)^2}\exp\left(-\frac{s}{1+\beta}\right)+\frac{s}{\beta}\exp\left(-\frac{s}{\beta}\right), \\
F_S(s) &= \frac{1-\beta}{(1+\beta)^2}\left[(1+\beta)^2-(1+\beta)^2\exp\left(-\frac{s}{1+\beta}\right)-(1+\beta)s\exp\left(-\frac{s}{1+\beta}\right)\right] \\
&\quad +\frac{1}{\beta}\left[\beta^2-\beta^2\exp\left(-\frac{s}{\beta}\right)-\beta s\exp\left(-\frac{s}{\beta}\right)\right], \\
E\left[S1_{\{S>b\}}\right] &= \frac{(1-\beta)(1+\beta+b)}{1+\beta}\exp\left(-\frac{b}{1+\beta}\right)+(\beta+b)\exp\left(-\frac{b}{\beta}\right), \\
g_{X,S}(s) &= \frac{\left(1-\beta^2\right)s^2}{2(1+\beta)^3}\exp\left(-\frac{s}{1+\beta}\right)+\frac{s^2}{2\beta}\exp\left(-\frac{s}{\beta}\right), \\
E\left[X1_{\{S>b\}}\right] &= \left(1-\beta^2\right)\left[1+\frac{b}{1+\beta}+\frac{b^2}{2(1+\beta)^2}\right]\exp\left(-\frac{b}{1+\beta}\right) \\
&\quad +\left(\beta^2+b\beta+\frac{b^2}{2}\right)\exp\left(-\frac{b}{\beta}\right)
\end{aligned}
$$

for $0 < \beta \leq 1$, $x > 0$ and $y > 0$.

Block and Basu (1976)'s bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= \begin{cases} C \exp\left(-\alpha x - \beta y\right), & \text{if } x > y, \\ D \exp\left(-\gamma x - \delta y\right), & \text{if } x \leq y, \end{cases} \\
f_S(s) &= \frac{C}{\beta-\alpha}\left[\exp(-\alpha s) - \exp\left(-\frac{(\alpha+\beta)s}{2}\right)\right] - \frac{D}{\delta-\gamma}\left[\exp(-\delta s) - \exp\left(-\frac{(\delta+\gamma)s}{2}\right)\right], \\
F_S(s) &= \frac{C}{\beta-\alpha}\left[\frac{1-\exp(-\alpha s)}{\alpha} - \frac{2}{\alpha+\beta}\left(1 - \exp\left(-\frac{(\alpha+\beta)s}{2}\right)\right)\right] \\
&\quad - \frac{D}{\delta-\gamma}\left[\frac{2}{\delta+\gamma}\left(1-\exp\left(-\frac{(\delta+\gamma s}{2}\right)\right) - \frac{1-\exp(-\delta s)}{\delta}\right], \\
E\left[S 1_{\{S>b\}}\right] &= \frac{C}{\beta-\alpha}\left[\frac{(b\alpha+1)\exp(-b\alpha)}{\alpha^2} - \frac{4}{(\alpha+\beta)^2}\left(b\frac{\alpha+\beta}{2}+1\right)\exp\left(-b\frac{\alpha+\beta}{2}\right)\right] \\
&\quad - \frac{D}{\delta-\gamma}\left[\frac{4}{(\delta+\gamma)^2}\left(b\frac{\delta+\gamma}{2}+1\right)\exp\left(-b\frac{\delta+\gamma}{2}\right) - \frac{(b\delta+1)\exp(-b\delta)}{\delta^2}\right], \\
g_{X,S}(s) &= C\left[\frac{\exp(-\beta s)-\exp(-\alpha s)}{(\alpha-\beta)^2} - \frac{s\exp(-\alpha s)}{\alpha-\beta}\right] \\
&\quad + D\left[\frac{\exp(-\delta s) - \exp\left(-(\delta+\gamma)s/2\right)}{(\delta-\gamma)^2} - \frac{s\exp\left(-(\delta+\gamma)s/2\right)}{2(\delta-\gamma)}\right] \\
&\quad - C\left[\frac{\exp(-\beta s) - \exp\left(-\beta s - (\gamma-\delta)s/2\right)}{(\delta-\gamma)^2} - \frac{s\exp\left(-\beta s-(\gamma-\delta)s/2\right)}{2(\delta-\gamma)}\right], \\
E\left[X 1_{\{S>b\}}\right] &= C\left[-\frac{(b\alpha+1)\exp(-b\alpha)}{\alpha^2(\alpha-\beta)} + \frac{\exp(-b\beta)}{\beta(\alpha-\beta)^2} - \frac{\exp(-b\alpha)}{\alpha(\alpha-\beta)^2}\right] \\
&\quad + D\left[\frac{\exp(-b\delta)}{\delta(\delta-\gamma)^2} - \frac{2\exp\left(-(\delta+\gamma)b/2\right)}{(\delta-\gamma)^2(\delta+\gamma)} - \frac{(2+(\delta+\gamma)b)\exp\left(-(\delta+\gamma)b/2\right)}{(\delta+\gamma)^2(\delta-\gamma)}\right] \\
&\quad - C\left[\frac{\exp(-b\beta)}{\beta(\delta-\gamma)^2} - \frac{\exp\left(-b\beta-(\gamma-\delta)b/2\right)}{(\delta-\gamma)^2\left(\beta+(\gamma-\delta)/2\right)}\right] \\
&\quad + C\frac{\left(2+2b\beta+b(\gamma-\delta)\right)\exp\left(-b\beta-(\gamma-\delta)b/2\right)}{(\delta-\gamma)\left(2\beta+\gamma-\delta\right)^2}
\end{aligned}
$$

for $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $\delta > 0$, $x > 0$ and $y > 0$.

Arnold and Strauss (1991)'s bivariate exponential distribution:

$$
\begin{aligned}
f(x,y) &= C\exp\left(-\gamma x - \delta y - \theta xy\right), \\
f_S(s) &= C\sqrt{\pi}p^{-1/2}\exp\left(-\delta s + \frac{q^2}{4p}\right)\left[\Phi\left(\sqrt{2p}s - (2p)^{-1/2}q\right) - \Phi\left(-(2p)^{-1/2}q\right)\right], \\
g_{X,S}(s) &= 2^{-1}C\sqrt{\pi}p^{-3/2}q\exp\left(-\delta s + \frac{q^2}{4p}\right)\left[\Phi\left(\sqrt{2p}s - (2p)^{-1/2}q\right) - \Phi\left(-(2p)^{-1/2}q\right)\right] \\
&\quad - 2^{-1}Cp^{-1}\exp\left(-\delta s + \frac{q^2}{4p}\right)\left[\exp\left(-ps^2+qs\right)-1\right]
\end{aligned}
$$

for $-\infty < \gamma < \infty$, $-\infty < \delta < \infty$, $-\infty < \theta < \infty$, $x > 0$ and $y > 0$, where $p = -\theta$ and $q = \delta - \theta s$.

Mixtures of independent gamma distributions with real shape parameters:

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty} C_k x^{\alpha_k-1} y^{\beta_k-1} \exp\left(-\gamma_k x - \delta_k y\right), \\
f_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) s^{\alpha_k+\beta_k-1} \exp\left(-\delta_k s\right) {}_1F_1\left(\alpha_k;\alpha_k+\beta_k;\left(\delta_k-\gamma_k\right)s\right), \\
F_S(s) &= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) \frac{\left(\alpha_k\right)_i \left(\delta_k-\gamma_k\right)^i \gamma\left(i+\alpha_k+\beta_k,\delta_k s\right)}{i!\left(\alpha_k+\beta_k\right)_i \delta_k^{\alpha_k+\beta_k+i}}, \\
E\left[S1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) \frac{\left(\alpha_k\right)_i \left(\delta_k-\gamma_k\right)^i \Gamma\left(i+\alpha_k+\beta_k+1,\delta_k b\right)}{i!\left(\alpha_k+\beta_k\right)_i \delta_k^{\alpha_k+\beta_k+i+1}}, \\
g_{X,S}(s) &= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} C_k B\left(\alpha_k+1,\beta_k\right) \frac{\left(\alpha_k+1\right)_i \left(\delta_k-\gamma_k\right)^i s^{i+\alpha_k+\beta_k} \exp\left(-\delta_k s\right)}{\left(1+\alpha_k+\beta_k\right)_i i!}, \\
E\left[X1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} C_k B\left(\alpha_k+1,\beta_k\right) \frac{\left(\alpha_k+1\right)_i \left(\delta_k-\gamma_k\right)^i \Gamma\left(i+\alpha_k+\beta_k+1,\delta_k b\right)}{\left(1+\alpha_k+\beta_k\right)_i i! \delta_k^{\alpha_k+\beta_k+i+1}}
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $\delta_k > 0$, $x > 0$ and $y > 0$.

Nadarajah and Gupta (2006)'s bivariate gamma distribution with equal scale parameters:

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty} C_k x^{\alpha_k-1} y^{\beta_k-1} (x+y)^{\gamma_k} \exp\left(-p_k x - p_k y\right), \\
f_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) s^{\alpha_k+\beta_k+\gamma_k-1} \exp\left(-p_k s\right), \\
F_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) p_k^{-\alpha_k-\beta_k-\gamma_k} \gamma\left(\alpha_k+\beta_k+\gamma_k, p_k s\right), \\
g_{X,S}(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k+1,\beta_k\right) s^{\alpha_k+\beta_k+\gamma_k} \exp\left(-p_k s\right), \\
E\left[S1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k,\beta_k\right) p_k^{-\alpha_k-\beta_k-\gamma_k-1} \Gamma\left(\alpha_k+\beta_k+\gamma_k+1, p_k b\right), \\
E\left[X1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k+1,\beta_k\right) p_k^{-\alpha_k-\beta_k-\gamma_k-1} \Gamma\left(\alpha_k+\beta_k+\gamma_k+1, p_k b\right)
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $p_k > 0$, $x > 0$ and $y > 0$.

Nadarajah and Gupta (2006)'s bivariate gamma distribution with unequal scale parameters:

$$\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty} C_k x^{\alpha_k-1} y^{\beta_k-1} \left(p_k x + q_k y\right)^{\gamma_k} \exp\left(-p_k x - q_k y\right), \\
f_S(s) &= \sum_{k=0}^{\infty} C_k q_k^{\gamma_k} \exp\left(-q_k s\right) \sum_{i=0}^{\infty} \binom{\gamma_k}{i} \left(\frac{p_k}{q_k} - 1\right)^i B\left(i+\alpha_k, \beta_k\right) s^{\alpha_k+\beta_k+\gamma_k-1} \\
&\quad \cdot {}_1F_1\left(i+\alpha_k; i+\alpha_k+\beta_k; -\left(p_k-q_k\right)s\right), \\
F_S(s) &= \sum_{k=0}^{\infty} C_k q_k^{\gamma_k} \sum_{i=0}^{\infty} \binom{\gamma_k}{i} \left(\frac{p_k}{q_k} - 1\right)^i B\left(i+\alpha_k, \beta_k\right) \\
&\quad \cdot \sum_{j=0}^{\infty} \frac{\left(i+\alpha_k\right)_j (-1)^j \left(p_k-q_k\right)^i}{\left(i+\alpha_k+\beta_k\right)_j j! q_k^{\alpha_k+\beta_k+\gamma_k+j}} \gamma\left(\alpha_k+\beta_k+\gamma_k+j, q_k s\right), \\
g_{X,S}(s) &= \sum_{k=0}^{\infty} C_k q_k^{\gamma_k} \exp\left(-q_k s\right) \sum_{i=0}^{\infty} \binom{\gamma_k}{i} \left(\frac{p_k}{q_k} - 1\right)^i B\left(i+\alpha_k+1, \beta_k\right) s^{\alpha_k+\beta_k+\gamma_k} \\
&\quad \cdot {}_1F_1\left(i+\alpha_k+1; i+\alpha_k+\beta_k+1; -\left(p_k-q_k\right)s\right)
\end{aligned}$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $p_k > 0$, $q_k > 0$, $x > 0$ and $y > 0$.

Nagar and Sepulveda-Murillo (2011)'s bivariate confluent hypergeometric distribution:

$$\begin{aligned}
f(x,y) &= C x^{p-1} y^{q-1} {}_1F_1\left(\alpha; \beta; -x-y\right), \\
f_S(s) &= C B(p,q) s^{p+q-1} {}_1F_1\left(\alpha; \beta; -s\right), \\
F_S(s) &= C B(p,q) \sum_{k=0}^{\infty} \frac{(\alpha)_k (-1)^k}{(\beta)_k k!} \frac{s^{p+q+k}}{p+q+k}, \\
g_{X,S}(s) &= C B(p+1,q) s^{p+q} {}_1F_1\left(\alpha; \beta; -s\right)
\end{aligned}$$

for $p > 0$, $q > 0$, $\alpha > 0$, $\beta > 0$, $x > 0$ and $y > 0$.

Becker and Roux (1981)'s bivariate gamma distribution:

$$
\begin{aligned}
f(x,y) &= \begin{cases} Cx^{\alpha-1}(x+py)^{\beta-1}\exp(-\gamma x-\delta y), & \text{if } x<y, \\ Dx^{c-1}(x+qy)^{d-1}\exp(-ex-fy), & \text{if } x\geq y, \end{cases} \\
f_S(s) &= 2^{-\alpha}\alpha^{-1}Cs^{\alpha+\beta-1}\exp(-\delta s)\Phi_1\left(\alpha,1-\beta,1+\alpha,(\delta-\gamma)\frac{s}{2},-\frac{1-p}{2p}\right) \\
&\quad +q^{d-1}c^{-1}Ds^{c+d-1}\exp(-fs)\Phi_1\left(c,1-d,1+c,(f-e)s,\frac{q-1}{q}\right) \\
&\quad -2^{-c}q^{d-1}Ds^{c+d-1}\exp(-fs)\Phi_1\left(c,1-d,1+c,(f-e)\frac{s}{2},\frac{q-1}{2q}\right), \\
g_{X,S}(s) &= 2^{-\alpha+1}(\alpha+1)^{-1}Cs^{\alpha+\beta}\exp(-\delta s)\Phi_1\left(\alpha+1,1-\beta,2+\alpha,(\delta-\gamma)\frac{s}{2},-\frac{1-p}{2p}\right) \\
&\quad +q^{d-1}(1+c)^{-1}Ds^{c+d}\exp(-fs)\Phi_1\left(1+c,1-d,2+c,(f-e)s,\frac{q-1}{q}\right) \\
&\quad -2^{-1-c}q^{d-1}Ds^{c+d}\exp(-fs)\Phi_1\left(1+c,1-d,2+c,(f-e)\frac{s}{2},\frac{q-1}{2q}\right)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $\delta>0$, $c>0$, $d>0$, $e>0$, $f>0$, $p>0$, $q>0$, $x>0$ and $y>0$.

Mohsin et al. (2013)'s bivariate gamma distribution:

$$
\begin{aligned}
f(x,y) &= \frac{\beta^\alpha\delta^\gamma}{\Gamma(\alpha)\Gamma(\gamma)}x^{\alpha-\gamma-1}y^{\gamma-1}\exp\left(-\beta x-\frac{\delta y}{x}\right), \\
f_S(s) &= \sum_{k=0}^{\infty}\frac{\delta^{\frac{\alpha+k+\gamma-1}{2}}\exp\left(\frac{\delta}{2}\right)}{\Gamma(\alpha)}\frac{(-1)^k\beta^{\alpha+k}}{k!}s^{\alpha+k-1}W_{\frac{1-\gamma-\alpha-k}{2},\frac{\alpha+k-\gamma}{2}}(\delta), \\
g_{X,S}(s) &= \sum_{k=0}^{\infty}\frac{\delta^{\frac{\alpha+k+\gamma}{2}}\exp\left(\frac{\delta}{2}\right)}{\Gamma(\alpha)}\frac{(-1)^k\beta^{\alpha+k}}{k!}s^{\alpha+k}W_{\frac{-\gamma-\alpha-k}{2},\frac{\alpha+k-\gamma+1}{2}}(\delta)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $\delta>0$, $x>0$ and $y>0$.

Cheriyan (1941)'s bivariate gamma distribution in equation (8.31) of Balakrishnan and

Lai (2009):

$$
\begin{aligned}
f(x,y) &= C\exp(-x-y)\int_0^{\min(x,y)}(x-z)^{\theta_1}(y-z)^{\theta_2-1}z^{\theta_3-1}\exp(z)dz,\\
f_S(s) &= CB\left(\theta_1+1,\theta_2\right)\exp(-s)\\
&\quad\cdot\int_0^{s/2}x^{\theta_1+\theta_3}(s-x)^{\theta_2-1}\Phi_1\left(\theta_3,1-\theta_2,\theta_1+\theta_3+1,x,\frac{x}{1-x}\right)dx\\
&\quad+CB\left(\theta_2,\theta_3\right)\exp(-s)\\
&\quad\cdot\int_{s/2}^{1}x^{\theta_1}(s-x)^{\theta_2+\theta_3-1}\Phi_1\left(\theta_3,-\theta_1,\theta_2+\theta_3,s-x,\frac{s-x}{x}\right)dx,\\
g_{X,S}(s) &= CB\left(\theta_1+1,\theta_2\right)\exp(-s)\\
&\quad\cdot\int_0^{s/2}x^{\theta_1+\theta_3+1}(s-x)^{\theta_2-1}\Phi_1\left(\theta_3,1-\theta_2,\theta_1+\theta_3+1,x,\frac{x}{1-x}\right)dx\\
&\quad+CB\left(\theta_2,\theta_3\right)\exp(-s)\\
&\quad\cdot\int_{s/2}^{1}x^{\theta_1+1}(s-x)^{\theta_2+\theta_3-1}\Phi_1\left(\theta_3,-\theta_1,\theta_2+\theta_3,s-x,\frac{s-x}{x}\right)dx
\end{aligned}
$$

for $\theta_1>0$, $\theta_2>0$, $\theta_3>0$, $x>0$ and $y>0$.

Dussauchoy and Berland (1975)'s bivariate gamma distribution:

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty}C_k x^{\alpha_k-1}\left(y-\beta x\right)^{\theta_k-1}\exp\left[-\gamma_k x-\delta_k\left(y-\beta x\right)\right],\\
f_S(s) &= \sum_{k=0}^{\infty}\frac{C_k\exp\left(-\delta_k s\right)s^{\alpha_k+\theta_k-1}}{\alpha_k(1+\beta)^{\alpha_k}}\Phi_1\left(\alpha_k,1-\theta_k,\alpha_k+1,1+\beta,\gamma_k s-(1+\beta)\delta_k s\right),\\
g_{X,S}(s) &= \sum_{k=0}^{\infty}\frac{C_k s^{\alpha_k+\theta_k}\exp\left(-\delta_k s\right)}{\left(\alpha_k+1\right)\left(1+\beta\right)^{\alpha_k+1}}\Phi_1\left(\alpha_k+1,1-\theta_k,2+\alpha_k,1+\beta,\gamma_k s-(1+\beta)\delta_k s\right)
\end{aligned}
$$

for $\alpha_k>0$, $\beta>0$, $\gamma_k>0$, $\delta_k>0$ and $y>\beta x>0$.

Mixtures of independent two piece gamma distributions:

$$f(x,y)=\begin{cases}\sum_{k=1}^{\infty}C_k x^{\alpha_k-1}y^{\beta_k-1}\exp\left(-\gamma_k x-\delta_k y\right), & \text{if } x<y,\\ \sum_{k=1}^{\infty}D_k x^{p_k-1}y^{q_k-1}\exp\left(-r_k x-t_k y\right), & \text{if } x\geq y,\end{cases}$$

$$\begin{aligned}f_S(s)=&\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i}}s^{\beta_k-i-1}\exp\left(-\delta_k s\right)\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}\sum_{j=0}^{\alpha_k+i-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i-j}2^j j!}s^{\beta_k-i+j-1}\exp\left(-\frac{\gamma_k s}{2}-\frac{\delta_k s}{2}\right)\\&+\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!}{\left(t_k-r_k\right)^{p_k+i-j}2^j j!}s^{q_k-i+j-1}\exp\left(-\frac{r_k s}{2}-\frac{t_k s}{2}\right)\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!}{\left(t_k-r_k\right)^{p_k+i-j}j!}s^{q_k-i+j-1}\exp\left(-r_k s\right),\end{aligned}$$

$$\begin{aligned}F_S(s)=&\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!\gamma\left(\beta_k-i,\delta_k s\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i}\delta_k^{\beta_k-i}}\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}\sum_{j=0}^{\alpha_k+i-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!2^{\beta_k-i}\gamma\left(\beta_k-i+j,\frac{(\gamma_k+\delta_k)s}{2}\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i-j}j!\left(\gamma_k+\delta_k\right)^{\beta_k-i+j}}\\&+\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!2^{q_k-i}\gamma\left(q_k-i+j,\frac{(t_k+r_k)s}{2}\right)}{\left(t_k-r_k\right)^{p_k+i-j}\left(t_k+r_k\right)^{q_k-i+j}j!}\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!\gamma\left(q_k-i+j,r_k s\right)}{\left(t_k-r_k\right)^{p_k+i-j}r_k^{q_k-i+j}j!},\end{aligned}$$

$$\begin{aligned}E\left[S1_{\{S>b\}}\right]=&\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!\Gamma\left(\beta_k-i+1,\delta_k b\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i}\delta_k^{\beta_k-i+1}}\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}\sum_{j=0}^{\alpha_k+i-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i-1\right)!2^{\beta_k-i}\Gamma\left(\beta_k-i+j+1,\frac{(\gamma_k+\delta_k)b}{2}\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i-j}j!\left(\gamma_k+\delta_k\right)^{\beta_k-i+j+1}}\\&+\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!2^{q_k-i}\Gamma\left(q_k-i+j+1,\frac{(t_k+r_k)b}{2}\right)}{\left(t_k-r_k\right)^{p_k+i-j}\left(t_k+r_k\right)^{q_k-i+j+1}j!}\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i-1}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i-1\right)!\Gamma\left(q_k-i+j+1,r_k b\right)}{\left(t_k-r_k\right)^{p_k+i-j}r_k^{q_k-i+j+1}j!},\end{aligned}$$

$$\begin{aligned}g_{X,S}(s)=&\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i\right)!}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i+1}}s^{\beta_k-i-1}\exp\left(-\delta_k s\right)\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}\sum_{j=0}^{\alpha_k+i}C_k\binom{\beta_k-1}{i}\frac{(-1)^i\left(\alpha_k+i\right)!}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i-j+1}2^j j!}s^{\beta_k-i+j-1}\exp\left(-\frac{\gamma_k s}{2}-\frac{\delta_k s}{2}\right)\\&+\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i\right)!}{\left(t_k-r_k\right)^{p_k+i-j+1}2^j j!}s^{q_k-i+j-1}\exp\left(-\frac{r_k s}{2}-\frac{t_k s}{2}\right)\\&-\sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i}D_k\binom{q_k-1}{i}\frac{(-1)^i\left(p_k+i\right)!}{\left(t_k-r_k\right)^{p_k+i-j+1}j!}s^{q_k-i+j-1}\exp\left(-r_k s\right),\end{aligned}$$

$$
\begin{aligned}
E\left[X1_{\{S>b\}}\right] &= \sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1} C_k \binom{\beta_k-1}{i} \frac{(-1)^i\,(\alpha_k+i)!\Gamma\left(\beta_k-i,\delta_k b\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i+1}\delta_k^{\beta_k-i}} \\
&\quad - \sum_{k=1}^{\infty}\sum_{i=0}^{\beta_k-1}\sum_{j=0}^{\alpha_k+i} C_k \binom{\beta_k-1}{i} \frac{(-1)^i\,(\alpha_k+i)!2^{\beta_k-i}\Gamma\left(\beta_k-i+j,\frac{(\gamma_k+\delta_k)b}{2}\right)}{\left(\delta_k-\gamma_k\right)^{\alpha_k+i-j+1} j!\left(\gamma_k+\delta_k\right)^{\beta_k-i+j}} \\
&\quad + \sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i} D_k \binom{q_k-1}{i} \frac{(-1)^i\,(p_k+i)!2^{q_k-i}\Gamma\left(q_k-i+j,\frac{(t_k+r_k)b}{2}\right)}{\left(t_k-r_k\right)^{p_k+i-j+1}\left(t_k+r_k\right)^{q_k-i+j} j!} \\
&\quad - \sum_{k=1}^{\infty}\sum_{i=0}^{q_k-1}\sum_{j=0}^{p_k+i} D_k \binom{q_k-1}{i} \frac{(-1)^i\,(p_k+i)!\Gamma\left(q_k-i+j,r_k b\right)}{\left(t_k-r_k\right)^{p_k+i-j+1} r_k^{q_k-i+j} j!}
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $\delta_k > 0$, $p_k > 0$, $q_k > 0$, $r_k > 0$, $t_k > 0$, $x > 0$ and $y > 0$, provided that α_k, β_k, p_k and q_k are integers.

Beta Stacey distribution in equation (5.38) of Balakrishnan and Lai (2009):

$$
\begin{aligned}
f(x,y) &= Cx^{p-1}(y-x)^{q-1}y^r\exp\left(-\alpha y^{\beta}\right), \\
f_S(s) &= C\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}s^{\beta k+r+p+q-1}B(\beta k+r+1,p)_2F_1\left(1-q,p;\beta k+r+p+1;2\right), \\
F_S(s) &= C\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!(\beta k+r+p+q)}B(\beta k+r+1,p)_2F_1\left(1-q,p;\beta k+r+p+1;2\right)s^{\beta k+r+p+q}, \\
g_{X,S}(s) &= C\sum_{k=0}^{\infty}\frac{(-\alpha)^k}{k!}s^{\beta k+r+p+q}B(\beta k+r+1,p+1)_2F_1\left(1-q,p+1;\beta k+r+p+2;2\right)
\end{aligned}
$$

for $p > 0$, $q > 0$, $r > 0$, $\alpha > 0$, $\beta > 0$ and $y > x > 0$.

Mardia (1970)'s bivariate distribution in equation (5.77) of Balakrishnan and Lai (2009):

$$\begin{aligned}
f(x,y) &= C\left(\alpha x+1\right)^{p}\left(\beta y+1\right)^{q}\left(\gamma x+\delta y+1\right)^{r}, \\
f_S(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{r}{m}\alpha^{i}\beta^{j}\delta^{m}\gamma^{k-m} \\
&\quad \cdot B\left(i+k-m+1,j+m+1\right)s^{i+j+k+1}, \\
F_S(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{r}{m}\alpha^{i}\beta^{j}\delta^{m}\gamma^{k-m} \\
&\quad \cdot B\left(i+k-m+1,j+m+1\right)\frac{s^{i+j+k+2}}{i+j+k+2}, \\
g_{X,S}(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{r}{m}\alpha^{i}\beta^{j}\delta^{m}\gamma^{k-m} \\
&\quad \cdot B\left(i+k-m+2,j+m+1\right)s^{i+j+k+2}
\end{aligned}$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $\delta>0$, $p>0$, $q>0$, $r>0$, $x>0$ and $y>0$, provided that p, q and r are integers.

Mardia (1970)'s bivariate distribution in equation (5.78) of Balakrishnan and Lai (2009):

$$\begin{aligned}
f(x,y) &= Cx^{p}y^{q}(y-x-1)^{r}, \\
f_S(s) &= CB(q+1,p+1)s^{p+q+1}(s-1)^{r}{}_2F_1\left(-r,p+1;p+q+2;\frac{2s}{s-1}\right), \\
g_{X,S}(s) &= CB(q+1,p+2)s^{p+q+2}(s-1)^{r}{}_2F_1\left(-r,p+2;p+q+3;\frac{2s}{s-1}\right)
\end{aligned}$$

for $p>0$, $q>0$, $r>0$, $x>0$ and $y>x+1$.

Mardia (1970)'s bivariate gamma distribution in equation (5.81) of Balakrishnan and

Lai (2009):

$$
\begin{aligned}
f(x,y) &= C\left(\alpha x+1\right)^p\left(\beta x+\gamma y+1\right)^q\exp(-ry),\\
f_S(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{i+j}\binom{p}{i}\binom{q}{j}\binom{i+j}{k}\alpha^i(\beta-\gamma)^j(-1)^k k!r^{-k-1}s^{i+j-k}(1+\gamma s)^{q-j}\\
&\quad -C\sum_{i=0}^{p}\sum_{j=0}^{q}\binom{p}{i}\binom{q}{j}\alpha^i(\beta-\gamma)^j(-1)^{i+j}(i+j)!r^{-i-j-1}(1+\gamma s)^{q-j}\exp(-rs),\\
F_S(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{i+j}\sum_{m=0}^{q-j}\binom{p}{i}\binom{q}{j}\binom{i+j}{k}\binom{q-j}{m}\alpha^i(\beta-\gamma)^j\gamma^m(-1)^k k!r^{-k-1}\frac{s^{i+j-k+m+1}}{i+j-k+m+1}\\
&\quad -C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{m=0}^{q-j}\binom{p}{i}\binom{q}{j}\binom{q-j}{m}\alpha^i(\beta-\gamma)^j\gamma^m(-1)^{i+j}(i+j)!r^{-i-j-m-2}\gamma\left(m+1,rs\right),\\
g_{X,S}(s) &= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{i+j}\binom{p}{i}\binom{q}{j}\binom{i+j+1}{k}\alpha^i(\beta-\gamma)^j(-1)^k k!r^{-k-1}s^{i+j-k+1}(1+\gamma s)^{q-j}\\
&\quad -C\sum_{i=0}^{p}\sum_{j=0}^{q}\binom{p}{i}\binom{q}{j}\alpha^i(\beta-\gamma)^j(-1)^{i+j+1}(i+j+1)!r^{-i-j-2}(1+\gamma s)^{q-j}\exp(-rs)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $p>0$, $q>0$, $r>0$, $x>0$ and $y>0$, provided that p and q are integers.

Bivariate Liouville distribution on page 202 of Balakrishnan and Lai (2009):

$$
\begin{aligned}
&f(x,y)=Cx^{\alpha-1}y^{\beta-1}g(x+y),\\
&f_S(s)=CB(\alpha,\beta)s^{\alpha+\beta-1}g(s),\\
&F_S(s)=CB(\alpha,\beta)\int_0^s u^{\alpha+\beta-1}g(u)du,\\
&E\left[S1_{\{S>b\}}\right]=CB(\alpha,\beta)\int_b^\infty u^{\alpha+\beta}g(u)du,\\
&g_{X,S}(s)=CB(\alpha+1,\beta)s^{\alpha+\beta}g(s),\\
&E\left[X1_{\{S>b\}}\right]=CB(\alpha+1,\beta)\int_b^\infty u^{\alpha+\beta}g(u)du
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $x>0$ and $y>0$.

Bivariate Liouville distribution in equation (9.46) of Balakrishnan and Lai (2009):

$$
\begin{aligned}
&f(x,y) = Cx^{\alpha-1}y^{\beta-1}(x+y)^{-\gamma}g(x+y),\\
&f_S(s) = CB(\alpha,\beta)s^{\alpha+\beta-\gamma-1}g(s),\\
&F_S(s) = CB(\alpha,\beta)\int_0^s u^{\alpha+\beta-\gamma-1}g(u)du,\\
&E\left[S1_{\{S>b\}}\right] = CB(\alpha,\beta)\int_b^\infty u^{\alpha+\beta-\gamma}g(u)du,\\
&g_{X,S}(s) = CB(\alpha+1,\beta)s^{\alpha+\beta-\gamma}g(s),\\
&E\left[X1_{\{S>b\}}\right] = CB(\alpha+1,\beta)\int_b^\infty u^{\alpha+\beta-\gamma}g(u)du
\end{aligned}
$$

for $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $x > 0$ and $y > 0$.

Bivariate equilibrium distributions due to Unnikrishnan Nair and Sankaran (2014):

$$
\begin{aligned}
&f(x,y) = \mu^{-1}g\left(x+y\right),\\
&f_S(s) = \mu^{-1}sg\left(s\right),\\
&F_S(s) = \mu^{-1}\int_0^s ug\left(u\right)du,\\
&E\left[S1_{\{S>b\}}\right] = \mu^{-1}\int_b^\infty s^2g\left(s\right)ds,\\
&g_{X,S}(s) = 2^{-1}\mu^{-1}s^2g\left(s\right),\\
&E\left[X1_{\{S>b\}}\right] = 2^{-1}\mu^{-1}\int_b^\infty s^2g\left(s\right)ds,
\end{aligned}
$$

where $g(\cdot)$ is the probability density function of a univariate random variable X say and $\mu = E(X)$.

Chacko and Thomas (2007)'s bivariate Pareto distribution:

$$
\begin{aligned}
&f(x,y) = C(x+y)^{-\alpha},\\
&f_S(s) = Cs^{-\alpha}(s-2\beta),\\
&F_S(s) = \frac{C}{2-\alpha}\left[s^{2-\alpha}-(2\beta)^{2-\alpha}\right] - \frac{2\beta C}{1-\alpha}\left[s^{1-\alpha}-(2\beta)^{1-\alpha}\right],\\
&E\left[S1_{\{S>b\}}\right] = \frac{Cb^{3-\alpha}}{\alpha-3} - \frac{2\beta Cb^{2-\alpha}}{\alpha-2},\\
&g_{X,S}(s) = 2^{-1}Cs^{1-\alpha}(s-2\beta),\\
&E\left[X1_{\{S>b\}}\right] = \frac{Cb^{3-\alpha}}{2(\alpha-3)} - \frac{\beta Cb^{2-\alpha}}{\alpha-2}
\end{aligned}
$$

for $\alpha > 0$, $x > \beta > 0$ and $y > \beta > 0$.

Bivariate Pareto distribution in equation (10.68) of Balakrishnan and Lai (2009):

$$
\begin{aligned}
f(x,y) &= C\,(a+x+y)^{-\beta}\,,\\
f_S(s) &= Cs\,(a+s)^{-\beta}\,,\\
F_S(s) &= C\left[\frac{(a+s)^{2-\beta}}{2-\beta}-\frac{a(a+s)^{1-\beta}}{1-\beta}-\frac{a^{2-\beta}}{2-\beta}+\frac{a^{2-\beta}}{1-\beta}\right],\\
E\left[S1_{\{S>b\}}\right] &= C\left[\frac{(a+b)^{3-\beta}}{\beta-3}+\frac{2a(a+b)^{2-\beta}}{2-\beta}+\frac{a^2(a+b)^{1-\beta}}{\beta-1}\right],\\
g_{X,S}(s) &= 2^{-1}Cs^2\,(a+s)^{-\beta}\,,\\
E\left[X1_{\{S>b\}}\right] &= 2^{-1}C\left[\frac{(a+b)^{3-\beta}}{\beta-3}+\frac{2a(a+b)^{2-\beta}}{2-\beta}+\frac{a^2(a+b)^{1-\beta}}{\beta-1}\right]
\end{aligned}
$$

for $a>0$, $\beta>0$, $x>0$ and $y>0$.

Bivariate Pareto distribution with equal scale parameters:

$$
\begin{aligned}
f(x,y) &= Cx^{\alpha-1}y^{\beta-1}(1+px+py)^{-\gamma},\\
f_S(s) &= CB(\alpha,\beta)s^{\alpha+\beta-1}(1+ps)^{-\gamma},\\
F_S(s) &= \frac{CB(\alpha,\beta)}{\alpha+\beta}s^{\alpha+\beta}{}_2F_1\left(\gamma,\alpha+\beta;\alpha+\beta+1;-ps\right),\\
E\left[S1_{\{S>b\}}\right] &= CB(\alpha,\beta)\frac{b^{\alpha+\beta-\gamma+1}}{p^{\gamma}(\gamma-\alpha-\beta-1)}{}_2F_1\left(\gamma,\gamma-\alpha-\beta-1;\gamma-\alpha-\beta;-\frac{1}{pb}\right),\\
g_{X,S}(s) &= CB(\alpha+1,\beta)s^{\alpha+\beta}(1+ps)^{-\gamma},\\
E\left[X1_{\{S>b\}}\right] &= CB(\alpha+1,\beta)\frac{b^{\alpha+\beta-\gamma+1}}{p^{\gamma}(\gamma-\alpha-\beta-1)}{}_2F_1\left(\gamma,\gamma-\alpha-\beta-1;\gamma-\alpha-\beta;-\frac{1}{pb}\right)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $p>0$, $\gamma>0$, $x>0$ and $y>0$.

Bivariate Pareto distribution with unequal scale parameters:

$$
\begin{aligned}
f(x,y) &= Cx^{\alpha-1}y^{\beta-1}(1+px+qy)^{-\gamma},\\
f_S(s) &= CB(\alpha,\beta)s^{\alpha+\beta-1}(1+qs)^{-\gamma}{}_2F_1\left(\gamma,\alpha;\alpha+\beta;\frac{(q-p)s}{1+qs}\right),\\
F_S(s) &= CB(\alpha,\beta)s^{\alpha+\beta}\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k s^k}{(\alpha+\beta)_k k!(k+\alpha+\beta)}\\
&\quad\cdot{}_2F_1\left(k+\gamma,k+\alpha+\beta;k+\alpha+\beta+1;-qs\right),\\
E\left[S1_{\{S>b\}}\right] &= CB(\alpha,\beta)\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k}{(\alpha+\beta)_k k!}\frac{b^{\alpha+\beta-\gamma+1}}{q^{\gamma+k}(\gamma-\alpha-\beta-1)}\\
&\quad\cdot{}_2F_1\left(k+\gamma,\gamma-\alpha-\beta-1;\gamma-\alpha-\beta;-\frac{1}{qb}\right),\\
g_{X,S}(s) &= CB(\alpha+1,\beta)s^{\alpha+\beta}(1+qs)^{-\gamma}{}_2F_1\left(\gamma,\alpha+1;\alpha+\beta+1;\frac{(q-p)s}{1+qs}\right),\\
E\left[X1_{\{S>b\}}\right] &= CB(\alpha+1,\beta)\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha+1)_k(q-p)^k}{(\alpha+\beta+1)_k k!}\frac{b^{\alpha+\beta-\gamma+1}}{q^{\gamma+k}(\gamma-\alpha-\beta-1)}\\
&\quad\cdot{}_2F_1\left(k+\gamma,\gamma-\alpha-\beta-1;\gamma-\alpha-\beta;-\frac{1}{qb}\right)
\end{aligned}
$$

for $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $p > 0$, $q > 0$, $x > 0$ and $y > 0$.

Mixtures of independent Pareto distributions (Lee, 1981):

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty}\frac{C_k x^{\alpha_k-1}y^{\beta_k-1}}{(1+p_k x)^{\gamma_k}(1+q_k x)^{\delta_k}},\\
f_S(s) &= \sum_{k=0}^{\infty}C_k B(\alpha_k,\beta_k)s^{\alpha_k+\beta_k-1}F_1(\alpha_k,\gamma_k,\delta_k,\alpha_k+\beta_k;-p_k s,-q_k s),\\
g_{X,S}(s) &= \sum_{k=0}^{\infty}C_k B(\alpha_k+1,\beta_k)s^{\alpha_k+\beta_k}F_1(\alpha_k+1,\gamma_k,\delta_k,\alpha_k+\beta_k+1;-p_k s,-q_k s)
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $\delta_k > 0$, $p_k > 0$, $q_k > 0$, $x > 0$ and $y > 0$.

Mixtures of bivariate Pareto distributions with equal scale parameters (Jones, 2002;

El-Bassiouny and Jones, 2009; Nagar et al., 2009):

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty} C_k x^{\alpha_k - 1} y^{\beta_k - 1} \left(1 + p_k x + p_k y\right)^{-\gamma_k}, \\
f_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k, \beta_k\right) s^{\alpha_k + \beta_k - 1} \left(1 + p_k s\right)^{-\gamma_k}, \\
F_S(s) &= \sum_{k=0}^{\infty} \frac{C_k B\left(\alpha_k, \beta_k\right)}{\alpha_k + \beta_k} s^{\alpha_k + \beta_k} {}_2F_1\left(\gamma_k, \alpha_k + \beta_k; \alpha_k + \beta_k + 1; -p_k s\right), \\
E\left[S 1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k, \beta_k\right) \frac{b^{\alpha_k + \beta_k + 1 - \gamma_k}}{p_k^{\gamma_k}\left(\gamma_k - \alpha_k - \beta_k - 1\right)} {}_2F_1(\gamma_k, \gamma_k - \alpha_k - \beta_k - 1; \gamma_k - \alpha_k - \beta_k; -\frac{1}{p_k b}), \\
g_{X,S}(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k + 1, \beta_k\right) s^{\alpha_k + \beta_k} \left(1 + p_k s\right)^{-\gamma_k}, \\
E\left[X 1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k + 1, \beta_k\right) \frac{b^{\alpha_k + \beta_k + 1 - \gamma_k}}{p_k^{\gamma_k}\left(\gamma_k - \alpha_k - \beta_k - 1\right)} {}_2F_1(\gamma_k, \gamma_k - \alpha_k - \beta_k - 1; \gamma_k - \alpha_k - \beta_k; -\frac{1}{p_k b})
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $p_k > 0$, $\gamma_k > 0$, $x > 0$ and $y > 0$.

Mixtures of bivariate Pareto distributions with unequal scale parameters (Jones, 2002; El-Bassiouny and Jones, 2009; Nagar et al., 2009):

$$
\begin{aligned}
f(x,y) &= \sum_{k=0}^{\infty} C_k x^{\alpha_k - 1} y^{\beta_k - 1} \left(1 + p_k x + q_k y\right)^{-\gamma_k}, \\
f_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k, \beta_k\right) s^{\alpha_k + \beta_k - 1} \left(1 + q_k s\right)^{-\gamma_k} {}_2F_1\left(\gamma_k, \alpha_k; \alpha_k + \beta_k; \frac{\left(q_k - p_k\right) s}{1 + q_k s}\right), \\
F_S(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k, \beta_k\right) s^{\alpha_k + \beta_k} \sum_{i=0}^{\infty} \frac{\left(\gamma_k\right)_i \left(\alpha_k\right)_i \left(q_k - p_k\right)^i s^i}{\left(\alpha_k + \beta_k\right)_i i! \left(i + \alpha_k + \beta_k\right)} \\
&\quad \cdot {}_2F_1\left(i + \gamma_k, i + \alpha_k + \beta_k; i + \alpha_k + \beta_k + 1; -q_k s\right), \\
E\left[S 1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k, \beta_k\right) \sum_{i=0}^{\infty} \frac{\left(\gamma_k\right)_i \left(\alpha_k\right)_i \left(q_k - p_k\right)^i}{\left(\alpha_k + \beta_k\right)_i i! q_k^{i + \gamma_k}} \frac{b^{\alpha_k + \beta_k - \gamma_k + 1}}{\gamma_k - \alpha_k - \beta_k - 1} \\
&\quad \cdot {}_2F_1\left(\gamma_k + i, \gamma_k - \alpha_k - \beta_k - 1; \gamma_k - \alpha_k - \beta_k; -\frac{1}{q_k b}\right), \\
g_{X,S}(s) &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k + 1, \beta_k\right) s^{\alpha_k + \beta_k} \left(1 + q_k s\right)^{-\gamma_k} {}_2F_1\left(\gamma_k, \alpha_k + 1; \alpha_k + \beta_k + 1; \frac{\left(q_k - p_k\right) s}{1 + q_k s}\right), \\
E\left[X 1_{\{S>b\}}\right] &= \sum_{k=0}^{\infty} C_k B\left(\alpha_k + 1, \beta_k\right) \sum_{i=0}^{\infty} \frac{\left(\gamma_k\right)_i \left(\alpha_k + 1\right)_i \left(q_k - p_k\right)^i}{\left(\alpha_k + \beta_k + 1\right)_i i! q_k^{\gamma_k + i}} \frac{b^{\alpha_k + \beta_k - \gamma_k + 1}}{\gamma_k - \alpha_k - \beta_k - 1} \\
&\quad \cdot {}_2F_1\left(\gamma_k + i, \gamma_k - \alpha_k - \beta_k - 1; \gamma_k - \alpha_k - \beta_k; -\frac{1}{q_k b}\right)
\end{aligned}
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $p_k > 0$, $q_k > 0$, $x > 0$ and $y > 0$.

Generalized bivariate Pareto distribution:

$$
\begin{aligned}
f(x,y) &= \frac{Cx^{\alpha-1}y^{\beta-1}}{(1+px+qy+rxy)^{\delta}},\\
f_S(s) &= CB(\alpha,\beta)(-r)^{-\delta}(uv)^{-\delta}s^{\alpha+\beta-1}F_1\left(\alpha,\delta,\delta,\alpha+\beta;\frac{s}{u},\frac{s}{v}\right),\\
g_{X,S}(s) &= CB(\alpha+1,\beta)(-r)^{-\delta}(uv)^{-\delta}s^{\alpha+\beta}F_1\left(\alpha+1,\delta,\delta,\alpha+\beta+1;\frac{s}{u},\frac{s}{v}\right)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $\delta>0$, $p>0$, $q>0$, $r>0$, $x>0$ and $y>0$, where $u=\frac{p-q+rs}{2r}+\sqrt{\frac{1+qs}{r}+\frac{(p-q+rs)^2}{4r^2}}$ and $v=\frac{p-q+rs}{2r}-\sqrt{\frac{1+qs}{r}+\frac{(p-q+rs)^2}{4r^2}}$.

Lee and Cha (2014)'s bivariate distribution:

$$
\begin{aligned}
f(x,y) &= \begin{cases} C(x+\alpha)^{\gamma}(y+\beta)^{\delta}, & \text{if } x<y,\\ D(x+t)^{v}(y+u)^{w}, & \text{if } x>y,\end{cases}\\
f_S(s) &= 3^{-1}C\alpha^{\gamma}s^{-1}(s+\beta)^{\delta}F_1\left(1,-\gamma,-\delta,2;-\frac{s}{2\alpha},\frac{s}{2(\alpha+\beta)}\right)\\
&\quad +3^{-1}Ds^{-1}\left(\frac{s}{2}+t\right)^{v}\left(\frac{s}{2}+u\right)^{w}F_1\left(1,-v,-w,2;-\frac{s}{s+2t},\frac{s}{s+2u}\right),\\
g_{X,S}(s) &= 3^{-1}C\alpha^{1+\gamma}s^{-1}(s+\beta)^{\delta}F_1\left(1,-1-\gamma,-\delta,2;-\frac{s}{2\alpha},\frac{s}{2(\alpha+\beta)}\right)\\
&\quad -3^{-1}C\alpha^{1+\gamma}s^{-1}(s+\beta)^{\delta}F_1\left(1,-\gamma,-\delta,2;-\frac{s}{2\alpha},\frac{s}{2(\alpha+\beta)}\right)\\
&\quad +3^{-1}Ds^{-1}\left(\frac{s}{2}+t\right)^{1+v}\left(\frac{s}{2}+u\right)^{w}F_1\left(1,-1-v,-w,2;-\frac{s}{s+2t},\frac{s}{s+2u}\right)\\
&\quad -3^{-1}Dts^{-1}\left(\frac{s}{2}+t\right)^{v}\left(\frac{s}{2}+u\right)^{w}F_1\left(1,-v,-w,2;-\frac{s}{s+2t},\frac{s}{s+2u}\right)
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $t>0$, $u>0$, $\delta<0$, $v<0$, $x>0$ and $y>0$.

Truncated bivariate normal distribution:

$$
\begin{aligned}
f(x,y) &= C\exp\left(-\alpha x^2-\beta y^2-\gamma x-\delta y-\theta xy\right),\\
f_S(s) &= C\sqrt{\pi}p^{-1/2}\exp\left(-\beta s^2-\delta s+\frac{q^2}{4p}\right)\left[\Phi\left(\sqrt{2p}s-(2p)^{-1/2}q\right)-\Phi\left(-(2p)^{-1/2}q\right)\right],\\
g_{X,S}(s) &= 2^{-1}C\sqrt{\pi}p^{-3/2}q\exp\left(-\beta s^2-\delta s+\frac{q^2}{4p}\right)\left[\Phi\left(\sqrt{2p}s-(2p)^{-1/2}q\right)-\Phi\left(-(2p)^{-1/2}q\right)\right]\\
&\quad -2^{-1}Cp^{-1}\exp\left(-\beta s^2-\delta s+\frac{q^2}{4p}\right)\left[\exp\left(-ps^2+qs\right)-1\right]
\end{aligned}
$$

for $\alpha>0$, $\beta>0$, $-\infty<\gamma<\infty$, $-\infty<\delta<\infty$, $-\infty<\theta<\infty$, $x>0$ and $y>0$, where $p=\alpha+\beta-\theta$ and $q=2\beta s-\gamma+\delta-\theta s$.

Chapter 9

Moments using Copulas

9.1 Introduction

Sums of a random number of random variables arise in many areas of the finance and economics. A prominent example is the total insurance claim over a fixed period. In this example, the number $N(t) = n$ is the number of insurance claims in time interval $[0, t]$ and the variables $X = (X_1, X_2, \cdots, X_n)$ are the amounts claimed. A related variable $\{W_i\}$ represents the inter-arrival time between insurance claims. The model of the sums of a random number of random variables (or, the model of random sums) is very important to insurers in terms of making predictions of the expected total amount of claims from groups of individuals in different classes of risk. (Papush et. al., 2001). Accurate estimates of the random sums over a time period can help an insurer optimize the insurance contract to be more competetive and attract more potential clients.

There is a vast amount of literature on the distributions and moments of the random sums. Many of the papers suppose that the claims are independently and identically distributed conditioned on the number of claims. They suppose no relationship between the amounts claimed and the inter-arrival times between insurance claims. This may not be a realistic assumption. Recently, researchers have supposed that the amounts claimed and the inter-arrival times are dependent, but are marginally independently and identically distributed conditioned on the number of claims. Some papers based on this assumption are Albrecher and Teugels (2006), Asimit and Jones (2008a), Cossette et al. (2008), Marceau (2008), Ambagaspitiya (2009), Asimit and Badescu (2010) and Chueng et al. (2010). See also Asimit and Jones (2008b), Asimit et al. (2014a, 2014b), Asimit and Chen (2015) and Asimit et al. (2016).

A most recent of these papers is by Mao and Zhao (2013). They derived the first and second moments of the total insurance claim by supposing that the joint distribution of amounts claimed and the inter-arrival times is specified by a copula. They specialized their results to one of the simplest known copulas, the Farlie-Gumbel-Morgenstern (FGM) copula due to Morgenstern (1956), see also Nelsen (2006).

In this chapter, we extend Mao and Zhao (2013)'s results. We derive the general moment of the total insurance claim. Our results hold for a wide range of copulas not just the FGM copula. We also extend Mao and Zhao (2013)'s results to the case where the identical assumption does not hold.

In management science and related areas, higher order moments are of interest, not just the mean and variance. Examples include: portfolio selection (Harvey et al., 2010); value-at-risk forecasting (Polanski and Stoja, 2010); market risk assessment (Sihem and Slaheddine, 2014).

The contents of this chapter are organized as follows. Section 9.2 derives the general moment of the random sums, under the assumptions given in Mao and Zhao (2014). Section 9.3 extends the results of Section 9.2 to the case that the inter-arrival times between claims are independent but not identically distributed, the amounts claimed are independent but not identically distributed, and the copulas of the joint distributions are not identical. Section 9.4 performs a simulation study to show the practical values of the results in Section 9.2. We show in particular that the expressions in Section 9.2 are computationally less time consuming and computationally more accurate than results obtained by simulation.

Section 9.4 shows that the results in Sections 9.2 and 9.3 can really have practical appeal. Suppose an insurance company wants to determine the expected value and variance of the total claim over a year. Being able to compute them more accurately and in less time is of course crucial, so that important short-term and long-term decisions could be made in the company.

9.2 Main results

Here, we suppose $N(t)$ is a renewal process with inter arrival times W_i, and (X_i, W_i) are independent and identically distributed with common copula function C, common copula density c, common joint probability density function $f_{X,W}$, common joint cumulative distribution function $F_{X,W}$, common marginal probability density functions f_X, f_W, and common marginal cumulative distribution functions F_X, F_W.

From the Skar Theorem, we have

$$F_{(X,W)}(x,w) = C(F_X(x), F_W(w))$$

and

$$f_{(X,W)}(x,w) = \frac{\partial^2 C}{\partial\mu\partial\nu} \cdot f_X(x) \cdot f_W(w),$$

where $\mu = F_X(x)$ and $\nu = F_W(w)$.

Let $S_t = X_1 + \cdots + X_{N(t)}$. Theorem 1 derives the kth moment of S_t. It generalizes Theorems 3.1 and 4.1 in Mao and Zhao (2013). These theorems derived the first and second moments of S_t.

Theorem 1 *The kth moment of S_t can be expressed as*

$$\begin{aligned} E\left(S_t^k\right) = \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \Bigg\{ \int_{t-w_{i_1}-\cdots-w_{i_m}>0} & F^{*(n-m)}\left(t - w_{i_1} - \cdots - w_{i_m}\right) \\ & \cdot \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}\left(w_{i_j}\right) f_W\left(w_{i_j}\right) dw_{i_j} \right] \\ - \int_{y=0}^{t} \int_{t-y-w_{i_1}-\cdots-w_{i_m}>0} & F^{*(n-m)}\left(t - y - w_{i_1} - \cdots - w_{i_m}\right) \\ & \cdot \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}\left(w_{i_j}\right) f_W\left(w_{i_j}\right) dw_{i_j} \right] f_W(y) dy \Bigg\}, \end{aligned}$$

where

$$g_{X^k}(w) = \int_0^{+\infty} x^k f_X(x) c\left(F_X(x), F_W(w)\right) dx, \tag{9.1}$$

and $k_{i_1}, \ldots, k_{i_m}$ denote the positive values among $k_1, \ldots, k_n$ and $w_{i_1}, \ldots, w_{i_m}$ are the corresponding ws.

Proof: Set $B_n = \{N(t) = n\} = \{W_1 + \cdots + W_n \le t\} - \{W_1 + \cdots + W_{n+1} \le t\}$ and $b_n = \{w_1 + \cdots + w_n \le t\} - \{w_1 + \cdots + w_{n+1} \le t\}$. Let I_A denote the indicator function, that is $I_A = 1$ if A is true and $I_A = 0$ if A is false. By the multinomial theorem,

$$
\begin{aligned}
E\left[S_t^k\right] &= E\left[\left(X_1+\cdots+X_{N(t)}\right)^k\right] \\
&= E\left\{E\left[\left(X_1+\cdots+X_{N(t)}\right)^k \middle| N(t)\right]\right\} \\
&= \sum_{n=1}^{\infty} E\left[\left(X_1+\cdots+X_n\right)^k \middle| N(t)=n\right] \Pr\left(N(t)=n\right) \\
&= \sum_{n=1}^{\infty} E\left[\sum_{k_1+\cdots+k_n=k} \binom{k}{k_1 \cdots k_n} \prod_{t=1}^{n} X_t^{k_t} \middle| N(t)=n\right] \Pr\left(N(t)=n\right) \\
&= \sum_{n=1}^{\infty} \sum_{k_1+\cdots+k_n=k} \binom{k}{k_1 \cdots k_n} E\left[\prod_{t=1}^{n} X_t^{k_t} \middle| N(t)=n\right] \Pr\left(N(t)=n\right) \\
&= \sum_{n=1}^{\infty} \sum_{k_1+\cdots+k_n=k} \binom{k}{k_1 \cdots k_n} E\left[I_{B_n} \prod_{t=1}^{n} X_t^{k_t}\right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(k,n)} \binom{n}{m} \sum_{k_{i_1}+k_{i_2}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0} \int \cdots \int_{b_n} \left\{\prod_{t=1}^{m} \int_0^{\infty} x_{i_t}^{k_{i_t}} f_{X,W}\left(x_{i_t}, w_{i_t}\right) dx_{i_t}\right\} \\
&\qquad \cdot \left[\prod_{j=1, j\neq i_1,\ldots,i_m}^{n+1} x_j^0 \cdot f_W\left(w_j\right) dw_j\right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(k,n)} \binom{n}{m} \sum_{k_{i_1}+k_{i_2}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0} \int \cdots \int_{b_n} \left\{\prod_{t=1}^{m} \int_0^{\infty} x_{i_t}^{k_{i_t}} c\left(F_X\left(x_{i_t}\right), F_W\left(w_{i_t}\right)\right)\right. \\
&\qquad \left. \cdot f_X\left(x_{i_t}\right) dx_{i_t}\right\} \cdot \left[\prod_{j=1}^{n+1} f_W\left(w_j\right) dw_j\right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(k,n)} \binom{n}{m} \sum_{k_{i_1}+k_{i_2}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0} \int \cdots \int_{b_n} \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\} \\
&\qquad \cdot \left[\prod_{j=1}^{n+1} f_W\left(w_j\right) dw_j\right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(k,n)} \binom{n}{m} \sum_{k_{i_1}+k_{i_2}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0} \int_{w_1+\cdots+w_n\leq t} \cdots \int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\} \left[\prod_{j=1}^{n} f_W\left(w_j\right) dw_j\right] \\
&\quad - \sum_{n=1}^{\infty} \sum_{m=1}^{min(k,n)} \binom{n}{m} \sum_{k_{i_1}+k_{i_2}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0} \int_{w_1+\cdots+w_{n+1}\leq t} \cdots \int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\} \\
&\qquad \cdot \left[\prod_{j=1}^{n+1} f_W\left(w_j\right) dw_j\right].
\end{aligned} \tag{9.2}
$$

The two integrals in (9.2) reduce to

$$
\begin{aligned}
(1) \quad : \quad & \int_{w_1+\cdots+w_n\le t}\cdots\int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\}\left[\prod_{j=1}^{n} f_W\left(w_j\right)dw_j\right] \\
&= \int_{w_1+\cdots+w_n\le t}\cdots\int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right) f_W\left(w_{i_j}\right)\right\}\left[\prod_{j=1, j\neq i_1,\ldots,i_m}^{n} f_W\left(w_j\right)\right]dw_j \\
&= \int_{w_{i_1}+\cdots+w_{i_m}\le t}\cdots\int F^{*(n-m)}\left(t-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right) f_W\left(w_{i_j}\right)dw_{i_j}\right] \\
&= \int_{t-w_{i_1}-\cdots-w_{i_m}\ge 0} F^{*(n-m)}\left(t-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right) f_W\left(w_{i_j}\right)dw_{i_j}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
(2) \quad : \quad & \int_{w_1+\cdots+w_{n+1}\le t}\cdots\int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\}\cdot\left[\prod_{j=1}^{n+1} f_W\left(w_j\right)dw_j\right] \\
&= \int_{y=0}^{t}\int_{w_1+\cdots+w_{n+1}\le t-y}\cdots\int \left\{\prod_{t=1}^{m} g_{X^{k_{i_t}}}\left(w_{i_t}\right)\right\}\cdot\left[\prod_{j=1}^{n} f_W\left(w_j\right)dw_j\right] f_W\left(y\right)dy \\
&= \int_{y=0}^{t}\int_{t-y-w_{i_1}-\cdots-w_{i_m}\ge 0} F^{*(n-m)}\left(t-y-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right) f_W\left(w_{i_j}\right)dw_{i_j}\right] f_W(y)dy,
\end{aligned}
$$

where $F^{*(n-m)}$, $m \le n$ is the $(n-m)-$fold convolution of F, $F^{*0} \equiv 1$. Hence, the result. The author also have checked the expressions of Theorem 1 when $k = 1$ and 2, which are identical with the derivations by Mao and Zhao (2014).

Corollaries 1 and 2 specialize Theorem 1 for $k = 3, 4$ and the author also gives the corresponding examples when n is a given real number. These corollaries can be used to compute among others the skewness and kurtosis of S_t.

Corollary 1 *The third moment of S_t can be expressed as*

$$
\begin{aligned}
E\left(S_t^3\right) &= \sum_{n=1}^{\infty}\sum_{m=1}^{3}\binom{n}{m}\sum_{k_{i_1}+\cdots+k_{i_m}=k}\binom{3}{k_{i_1},\cdots,k_{i_m},0,\cdots,0}\Bigg\{\int\cdots\int_{t-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{x^{k_{i_j}}}(w_{i_j})f_W(w_{i_j})dw_{i_j}\right] \\
&\qquad -\int_{y=0}^{t}\int_{t-y-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t-y-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{x^{k_{i_j}}}(w_{i_j})f_W(w_{i_j})dw_{i_j}\right]\Bigg\} \\
&= \sum_{n=1}^{\infty}\binom{n}{1}\binom{3}{3,0,\cdots,0}\left\{\int_{w=0}^{t}F^{*(n-1)}(t-w)\,g_{X^3}(w)\,f_W(w)\,dw-\int_{y=0}^{t}\int_{w=0}^{t-y}F^{*(n-1)}(t-y-w)\,g_{X^3}(w)\,f_W(w)\,dwdy\right\} \\
&\quad+\sum_{n=2}^{\infty}\binom{n}{2}\binom{3}{1,2,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_{X^1}(w_1)\,f_W(w_1)\,g_{X^2}(w_2)\,f_W(w_2)\,dw_2dw_1 \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_{X^1}(w_1)\,f_W(w_1)\,g_{X^2}(w_2)\,f_W(w_2)\,f_W(y)dw_2dw_1dy\Bigg\} \\
&\quad+\sum_{n=2}^{\infty}\binom{n}{2}\binom{3}{2,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_{X^2}(w_1)\,f_W(w_1)\,g_{X^1}(w_2)\,f_W(w_2)\,dw_2dw_1 \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_{X^2}(w_1)\,f_W(w_1)\,g_{X^1}(w_2)\,f_W(w_2)\,f_W(y)dw_2dw_1dy\Bigg\} \\
&\quad+\sum_{n=3}^{\infty}\binom{n}{3}\binom{3}{1,1,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3}^{t-w_1-w_2}F^{*(n-3)}(t-w_1-w_2-w_3)\prod_{i=1}^{3}\left[g_X(w_i)f_W(w_i)dw_i\right] \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3}^{t-y-w_1-w_2}F^{*(n-3)}(t-y-w_1-w_2-w_3)\prod_{i=1}^{3}\left[g_X(w_i)f_W(w_i)dw_i\right]f_W(y)dy\Bigg\} \\
&= \sum_{n=1}^{\infty}n\left\{\int_{w=0}^{t}F^{*(n-1)}(t-w)\,g_{X^3}(w)\,f_W(w)\,dw-\int_{y=0}^{t}\int_{w=0}^{t-y}F^{*(n-1)}(t-y-w)\,g_{X^3}(w)\,f_W(w)\,dwdy\right\} \\
&\quad+\sum_{n=2}^{\infty}\frac{3}{2}n(n-1)\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_{X^1}(w_1)\,f_W(w_1)\,g_{X^2}(w_2)\,f_W(w_2)\,dw_2dw_1 \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_{X^1}(w_1)\,f_W(w_1)\,g_{X^2}(w_2)\,f_W(w_2)\,f_W(y)dw_2dw_1dy \\
&\qquad+\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_{X^2}(w_1)\,f_W(w_1)\,g_{X^1}(w_2)\,f_W(w_2)\,dw_2dw_1 \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_{X^2}(w_1)\,f_W(w_1)\,g_{X^1}(w_2)\,f_W(w_2)\,f_W(y)dw_2dw_1dy\Bigg\} \\
&\quad+\sum_{n=3}^{\infty}n(n-1)(n-2)\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3}^{t-w_1-w_2}F^{*(n-3)}(t-w_1-w_2-w_3)\prod_{i=1}^{3}\left[g_X(w_i)f_W(w_i)dw_i\right] \\
&\qquad-\int_{y=0}^{t}\int_{w_1=0}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3}^{t-y-w_1-w_2}F^{*(n-3)}(t-y-w_1-w_2-w_3)\prod_{i=1}^{3}\left[g_X(w_i)f_W(w_i)dw_i\right]f_W(y)dy\Bigg\}
\end{aligned}
\tag{9.3}
$$

Example 1 *When* $n = 3$, $S_t^3 = (X_1 + X_2 + X_3)^3$. $E\left[S_t^3\right]$ *can be expressed as*

$$
\begin{aligned}
&E\left(S_t^3\right) = 3\int_0^t F^{*(2)}\left(t-w\right) g_{X^3}\left(w\right) f_W\left(w\right) dw \\
&+9\int_0^t \int_0^{t-w_1} F^{*(1)}\left(t-w_1-w_2\right) g_X\left(w_1\right) g_{X^2}\left(w_2\right) f_W\left(w_1\right) f_W\left(w_2\right) dw_2 dw_1 \\
&+9\int_0^t \int_0^{t-w_1} F^{*(1)}\left(t-w_1-w_2\right) g_{X^2}\left(w_1\right) g_X\left(w_2\right) f_W\left(w_1\right) f_W\left(w_2\right) dw_2 dw_1 \\
&+6\int_0^t \int_0^{t-w_1} \int_0^{t-w_1-w_2} F^{*(0)}\left(t-w_1-w_2\right) \prod_{j=1}^{3}\left[g_X\left(w_j\right) f_W\left(w_j\right)\right] dw_3 dw_2 dw_1 \\
&-3\int_0^t \int_0^{t-y} F^{*(2)}\left(t-y-w\right) g_{X^3}\left(w\right) f_W\left(w\right) f_W(y) dw dy \\
&-9\int_0^t \int_0^{t-y} \int_0^{t-y-w_1} F^{*(1)}\left(t-y-w_1-w_2\right) g_X\left(w_1\right) g_{X^2}\left(w_2\right) \\
&\qquad \cdot f_W\left(w_1\right) f_W\left(w_2\right) f_W(y) dw_2 dw_1 dy \\
&-9\int_0^t \int_0^{t-y} \int_0^{t-y-w_1} F^{*(1)}\left(t-y-w_1-w_2\right) g_{X^2}\left(w_1\right) g_X\left(w_2\right) \\
&\qquad \cdot f_W\left(w_1\right) f_W\left(w_2\right) f_W(y) dw_2 dw_1 dy \\
&-6\int_0^t \int_0^{t-y} \int_0^{t-y-w_1} \int_0^{t-y-w_1-w_2} F^{*(0)}\left(t-y-w_1-w_2-w_3\right) \\
&\qquad \cdot \prod_{j=1}^{3}\left[g_X\left(w_j\right) f_W\left(w_j\right)\right] f_W(y) dw_3 dw_2 dw_1 dy.
\end{aligned}
$$

Corollary 2 *The fourth moment of S_t can be expressed as*

$$
\begin{aligned}
E\left(S_t^4\right) = & \sum_{n=1}^{\infty}\sum_{m=1}^{4}\binom{n}{m}\sum_{k_{i_1}+\cdots+k_{i_m}=k}\binom{4}{k_{i_1},\cdots,k_{i_m},0,\cdots,0}\Bigg\{\int\cdots\int_{t-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right)f_W\left(w_{i_j}\right)dw_{i_j}\right] \\
& -\int_{y=0}^{t}\int_{t-y-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t-y-w_{i_1}-\cdots-w_{i_m}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right)f_W\left(w_{i_j}\right)dw_{i_j}\right]f_W(y)dy\Bigg\} \\
= & \sum_{n=1}^{\infty}\binom{n}{1}\binom{4}{4,0,\cdots,0}\Bigg\{\int_{w=0}^{t}F^{*(n-1)}(t-w)\,g_{X^4}(w)\,f_W(w)\,dw-\int_{y=0}^{t}\int_{w=0}^{t-y}F^{*(n-1)}(t-y-w)\,g_{X^4}(w)\,f_W(w)\,f_W(y)\,dwdy\Bigg\} \\
& +\sum_{n=2}^{\infty}\binom{n}{2}\binom{4}{2,2,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\prod_{i=1}^{2}\left[g_{X^2}(w_i)\,f_W(w_i)\,dw_i\right] \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\prod_{i=1}^{2}\left[g_{X^2}(w_i)\,f_W(w_i)\,dw_i\right]f_W(y)\,dy\Bigg\} \\
& +\sum_{n=2}^{\infty}\binom{n}{2}\binom{4}{1,3,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_X(w_1)\,g_{X^3}(w_2)\,f_W(w_1)\,f_W(w_2)\,dw_1dw_2 \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_X(w_1)\,g_{X^3}(w_2)\,f_W(w_1)\,f_W(w_2)\,f_W(y)\,dw_1dw_2dy\Bigg\} \\
& +\sum_{n=2}^{\infty}\binom{n}{2}\binom{4}{3,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}F^{*(n-2)}(t-w_1-w_2)\,g_{X^3}(w_1)\,g_X(w_2)\,f_W(w_1)\,f_W(w_2)\,dw_1dw_2 \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}F^{*(n-2)}(t-y-w_1-w_2)\,g_{X^3}(w_1)\,g_X(w_2)\,f_W(w_1)\,f_W(w_2)\,f_W(y)\,dw_1dw_2dy\Bigg\} \\
& +\sum_{n=3}^{\infty}\binom{n}{3}\binom{4}{1,1,2,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-w_1-w_2-w_3)\,g_X(w_1)\,g_X(w_2)\,g_{X^2}(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1 \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-y-w_1-w_2-w_3)\,g_X(w_1)\,g_X(w_2)\,g_{X^2}(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1f_W(y)\,dy\Bigg\} \\
& +\sum_{n=3}^{\infty}\binom{n}{3}\binom{4}{1,2,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-w_1-w_2-w_3)\,g_X(w_1)\,g_{X^2}(w_2)\,g_X(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1 \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-y-w_1-w_2-w_3)\,g_X(w_1)\,g_{X^2}(w_2)\,g_X(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1f_W(y)\,dy\Bigg\} \\
& +\sum_{n=3}^{\infty}\binom{n}{3}\binom{4}{2,1,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-w_1-w_2-w_3)\,g_{X^2}(w_1)\,g_X(w_2)\,g_X(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1 \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3=0}^{t-w_1-w_2}F^{*(n-3)}(t-y-w_1-w_2-w_3)\,g_{X^2}(w_1)\,g_X(w_2)\,g_X(w_3)\prod_{i=1}^{3}f_W(w_i)\,dw_3dw_2dw_1f_W(y)\,dy\Bigg\} \\
& +\sum_{n=4}^{\infty}\binom{n}{4}\binom{4}{1,1,1,1,0,\cdots,0}\Bigg\{\int_{w_1=0}^{t}\int_{w_2=0}^{t-w_1}\int_{w_3=0}^{t-w_1-w_2}\int_{w_4=0}^{t-w_1-w_2-w_3}F^{*(n-4)}(t-w_1-w_2-w_3-w_4)\prod_{i=1}^{4}\left[g_X(w_i)\,f_W(w_i)\,dw_i\right] \\
& -\int_{y=0}^{t}\int_{w_1}^{t-y}\int_{w_2=0}^{t-y-w_1}\int_{w_3=0}^{t-w_1-w_2}\int_{w_4=0}^{t-w_1-w_2-w_3}F^{*(n-4)}(t-y-w_1-w_2-w_3-w_4)\prod_{i=1}^{4}\left[g_X(w_i)\,f_W(w_i)\,dw_i\right]f_W(y)\,dy\Bigg\}
\end{aligned}
\tag{9.4}
$$

Example 2 *When $n = 4$, $S_t^4 = (X_1 + X_2 + X_3 + X_4)^4$. $E\left[S_t^4\right]$ can be expressed as*

$$
\begin{aligned}
&E\left(S_t^4\right) = 4\int_0^t F^{*(3)}\left(t-w\right)g_{X^4}\left(w\right)f_W\left(w\right)dw \\
&+36\int_0^t\int_0^{t-w_1} F^{*(2)}\left(t-w_1-w_2\right)g_{X^2}\left(w_1\right)g_{X^2}\left(w_2\right)f_W\left(w_1\right)f_W\left(w_2\right)dw_2dw_1 \\
&+24\int_0^t\int_0^{t-w_1} F^{*(2)}\left(t-w_1-w_2\right)g_X\left(w_1\right)g_{X^3}\left(w_2\right)f_W\left(w_1\right)f_W\left(w_2\right)dw_2dw_1 \\
&+24\int_0^t\int_0^{t-w_1} F^{*(2)}\left(t-w_1-w_2\right)g_{X^3}\left(w_1\right)g_X\left(w_2\right)f_W\left(w_1\right)f_W\left(w_2\right)dw_2dw_1 \\
&+48\int_0^t\int_0^{t-w_1}\int_0^{t-w_1-w_2} F^{*(1)}\left(t-w_1-w_2-w_3\right)g_{X^2}\left(w_1\right)g_X\left(w_2\right)g_X\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)dw_3dw_2dw_1 \\
&+48\int_0^t\int_0^{t-w_1}\int_0^{t-w_1-w_2} F^{*(1)}\left(t-w_1-w_2-w_3\right)g_X\left(w_1\right)g_{X^2}\left(w_2\right)g_X\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)dw_3dw_2dw_1 \\
&+48\int_0^t\int_0^{t-w_1}\int_0^{t-w_1-w_2} F^{*(1)}\left(t-w_1-w_2-w_3\right)g_X\left(w_1\right)g_X\left(w_2\right)g_{X^2}\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)dw_3dw_2dw_1 \\
&+24\int_0^t\int_0^{t-w_1}\int_0^{t-w_1-w_2}\int_0^{t-w_1-w_2-w_3} F^{*(0)}\left(t-w_1-w_2-w_3-w_4\right) \\
&\qquad\cdot\prod_{j=1}^{4}\left[g_X\left(w_j\right)f_W\left(w_j\right)\right]dw_4dw_3dw_2dw_1 \\
&-4\int_0^t\int_0^{t-y} F^{*(3)}\left(t-y-w\right)g_{X^4}\left(w\right)f_W\left(w\right)f_W(y)dwdy \\
&-36\int_0^t\int_0^{t-y}\int_0^{t-y-w_1} F^{*(2)}\left(t-y-w_1-w_2\right)g_{X^2}\left(w_1\right)g_{X^2}\left(w_2\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W(y)dw_2dw_1dy \\
&-24\int_0^t\int_0^{t-y}\int_0^{t-y-w_1} F^{*(2)}\left(t-y-w_1-w_2\right)g_X\left(w_1\right)g_{X^3}\left(w_2\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W(y)dw_2dw_1dy \\
&-24\int_0^t\int_0^{t-y}\int_0^{t-y-w_1} F^{*(2)}\left(t-y-w_1-w_2\right)g_{X^3}\left(w_1\right)g_X\left(w_2\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W(y)dw_2dw_1dy \\
&-48\int_0^t\int_0^{t-y}\int_0^{t-y-w_1}\int_0^{t-y-w_1-w_2} F^{*(1)}\left(t-y-w_1-w_2-w_3\right) \\
&\qquad\cdot g_{X^2}\left(w_1\right)g_X\left(w_2\right)g_X\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)f_W(y)dw_3dw_2dw_1dy \\
&-48\int_0^t\int_0^{t-y}\int_0^{t-y-w_1}\int_0^{t-y-w_1-w_2} F^{*(1)}\left(t-y-w_1-w_2-w_3\right) \\
&\qquad\cdot g_X\left(w_1\right)g_{X^2}\left(w_2\right)g_X\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)f_W(y)dw_3dw_2dw_1dy \\
&-48\int_0^t\int_0^{t-y}\int_0^{t-y-w_1}\int_0^{t-y-w_1-w_2} F^{*(1)}\left(t-y-w_1-w_2-w_3\right) \\
&\qquad\cdot g_X\left(w_1\right)g_X\left(w_2\right)g_{X^2}\left(w_3\right) \\
&\qquad\cdot f_W\left(w_1\right)f_W\left(w_2\right)f_W\left(w_3\right)f_W(y)dw_3dw_2dw_1dy
\end{aligned}
$$

$$-24\int_0^t\int_0^{t-y}\int_0^{t-y-w_1}\int_0^{t-y-w_1-w_2}\int_0^{t-y-w_1-w_2-w_3} F^{*(0)}\left(t-y-w_1-w_2-w_3-w_4\right)$$
$$\cdot\prod_{j=1}^{4}\left[g_X\left(w_j\right)f_W\left(w_j\right)\right]f_W(y)dw_4dw_3dw_2dw_1dy.$$

Corollary 3 specializes Theorem 1 to the case that $N(t)$ is a Poisson process with rate parameter λ.

Corollary 3 *Suppose $\{N(t)\}$ is a Poisson process with rate parameter λ. Then the kth moment of S_t can be expressed as*

$$E\left(S_t^k\right)=\sum_{n=1}^{\infty}\sum_{m=1}^{min(n,k)}\binom{n}{m}\sum_{k_{i_1}+\cdots+k_{i_m}=k}\binom{k}{k_{i_1},\cdots,k_{i_m},0,\cdots,0}\frac{\lambda^m}{\Gamma(n-m)}$$
$$\cdot\left\{\int_{t-w_{i_1}-\cdots-w_{i_m}>0}\gamma\left(n-m,\lambda\left(t-w_{i_1}-\cdots-w_{i_m}\right)\right)\right.$$
$$\cdot\exp\left(-\lambda\sum_{j=1}^{m}w_{i_j}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right)dw_{i_j}\right]$$
$$-\lambda\int_0^t\int_{t-y-w_{i_1}-\cdots-w_{i_m}>0}\gamma\left(n-m,\lambda\left(t-y-w_{i_1}-\cdots-w_{i_m}\right)\right)$$
$$\left.\cdot\exp\left(-\lambda y-\lambda\sum_{j=1}^{m}w_{i_j}\right)\prod_{j=1}^{m}\left[g_{X^{k_{i_j}}}\left(w_{i_j}\right)dw_{i_j}\right]dy\right\},$$

where $g_{X^k}(\cdot)$ is given by (9.1), $\gamma(a,x)$ denotes the incomplete gamma function defined by

$$\gamma(a,x)=\int_0^x t^{a-1}\exp(-t)dt,$$

and $k_{i_1},\ldots,k_{i_m}$ denote the positive values among $k_1,\ldots,k_n$ and $w_{i_1},\ldots,w_{i_m}$ are the corresponding ws.

Nadarajah (2015) showed that a wide range of bivariate copulas (including the FGM copula) can be expressed as

$$c(u,v)=\sum_{i=1}^{n}\alpha_i a_i b_i u^{a_i-1}v^{b_i-1} \tag{9.5}$$

for $n\geq 1$ an integer and $\{(\alpha_i,a_i,b_i):i\geq 1\}$ some real numbers. Using (9.5), we can write

$$g_{X^k}(w)=\sum_{i=1}^{n}\alpha_i a_i b_i F_W^{b_i-1}(w)\int_0^{+\infty}x^k f_X(x)F_X^{a_i-1}(x)dx. \tag{9.6}$$

If a_i are positive integers and if $X_1, \ldots, X_{a_i}$ are independent and identical copies of X, then based on $E[X_{n:n}^k] = n \int x^k F^{n-1}(x) f(x)\, dx$, (9.6) can be further rewritten as

$$g_{X^k}(w) = \sum_{i=1}^{n} \alpha_i b_i F_W^{b_i - 1}(w) E\left[Z_{a_i:a_i}^k\right],$$

where $Z_{a_i:a_i} = \max(X_1, \ldots, X_{a_i})$, which represents the maximum order statistic.

9.3 An extension

Theorem 2 extends Theorem 1 to the case that $(X_1, W_1), (X_2, W_2), \ldots$ are independent but not identical. We suppose that (X_i, W_i) has copula function $C^{(i)}$, copula density $c^{(i)}$, joint probability density function f_{X_i,W_i}, joint cumulative distribution function F_{X_i,W_i}, marginal probability density functions f_{X_i}, f_{W_i}, and marginal cumulative distribution functions F_{X_i}, F_{W_i}.

Theorem 2 *The kth moment of S_t can be expressed as*

$$\begin{aligned} E\left(S_t^k\right) = \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1}+\cdots+k_{i_m}=k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \Bigg\{ \int_{t-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t - w_{i_1} - \cdots - w_{i_m}\right) \\ \cdot \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}^{(i_j)}\left(w_{i_j}\right) f_{W_{i_j}}\left(w_{i_j}\right) dw_{i_j} \right] \\ - \int_0^t \int_{t-y-w_{i_1}-\cdots-w_{i_m}>0} F^{*(n-m)}\left(t - y - w_{i_1} - \cdots - w_{i_m}\right) \\ \cdot \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}^{(i_j)}\left(w_{i_j}\right) f_{W_{i_j}}\left(w_{i_j}\right) dw_{i_j} \right] f_{W_{n+1}}(y) dy \Bigg\}, \end{aligned}$$

where

$$g_{X^k}^{(i)}(w) = \int_0^{+\infty} x^k f_{X_i}(x) c^{(i)}\left(F_{X_i}(x), F_{W_i}(w)\right) dx,$$

and $k_{i_1}, \ldots, k_{i_m}$ denote the positive values among $k_1, \ldots, k_n$ and $w_{i_1}, \ldots, w_{i_m}$ are the corresponding ws.

Proof: Define B_n, b_n and I_A as in the proof of Theorem 1. By the multinomial theorem,

$$
\begin{aligned}
E\left[S_t^k\right] &= E\left[\left(X_1 + \cdots + X_{N(t)}\right)^k\right] \\
&= \sum_{n=1}^{\infty} E\left[(X_1 + \cdots + X_n)^k \middle| N(t) = n\right] \Pr\left(N(t) = n\right) \\
&= \sum_{n=1}^{\infty} \sum_{k_1 + \cdots + k_n = k} \binom{k}{k_1 \cdots k_n} E\left[\prod_{t=1}^{n} X_t^{k_t} \middle| N(t) = n\right] \Pr\left(N(t) = n\right) \\
&= \sum_{n=1}^{\infty} \sum_{k_1 + \cdots + k_n = k} \binom{k}{k_1 \cdots k_n} E\left[I_{B_n} \prod_{t=1}^{n} X_t^{k_t}\right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1} + \cdots + k_{i_m} = k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \int \cdots \int_{b_n} \left\{ \prod_{t=1}^{m} \int x_{i_t}^{k_{i_t}} f_{X_{i_t}, W_{i_t}}\left(x_{i_t}, w_{i_t}\right) dx_{i_t} \right\} \\
&\quad \cdot \left[\prod_{j=1, j \neq i_1, \ldots, i_m}^{n+1} f_{W_j}\left(w_j\right) dw_j \right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1} + \cdots + k_{i_m} = k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \int \cdots \int_{b_n} \left\{ \prod_{t=1}^{m} \int x_{i_t}^{k_{i_t}} c^{(i_t)}\left(F_{X_{i_t}}\left(x_{i_t}\right), F_{W_{i_t}}\left(w_{i_t}\right)\right) f_{X_{i_t}}\left(x_{i_t}\right) dx_{i_t} \right\} \\
&\quad \cdot \left[\prod_{j=1}^{n+1} f_{W_j}\left(w_j\right) dw_j \right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1} + \cdots + k_{i_m} = k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \int \cdots \int_{b_n} \left\{ \prod_{t=1}^{m} g_{X^{k_{i_t}}}^{(i_t)}\left(w_{i_t}\right) \right\} \left[\prod_{j=1}^{n+1} f_{W_j}\left(w_j\right) dw_j \right] \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1} + \cdots + k_{i_m} = k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \int_{w_1 + \cdots + w_n \leq t} \left\{ \prod_{t=1}^{m} g_{X^{k_{i_t}}}^{(i_t)}\left(w_{i_t}\right) \right\} \left[\prod_{j=1}^{n+1} f_{W_j}\left(w_j\right) dw_j \right] \\
&\quad - \sum_{n=1}^{\infty} \sum_{m=1}^{min(n,k)} \binom{n}{m} \sum_{k_{i_1} + \cdots + k_{i_m} = k} \binom{k}{k_{i_1}, \cdots, k_{i_m}, 0, \cdots, 0} \int_{w_1 + \cdots + w_{n+1} \leq t} \left\{ \prod_{t=1}^{m} g_{X^{k_{i_t}}}^{(i_t)}\left(w_{i_t}\right) \right\} \left[\prod_{j=1}^{n+1} f_{W_j}\left(w_j\right) dw_j \right].
\end{aligned}
\tag{9.7}
$$

The two integrals in (9.7) reduce to

$$
\int_{t - w_{i_1} - \cdots - w_{i_m} > 0} F^{*(n-m)}\left(t - w_{i_1} - \cdots - w_{i_m}\right) \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}^{(i_j)}\left(w_{i_j}\right) f_{W_{i_j}}\left(w_{i_j}\right) dw_{i_j} \right]
$$

and

$$
\int_0^t \int_{t - y - w_{i_1} - \cdots - w_{i_m} > 0} F^{*(n-m)}\left(t - y - w_{i_1} - \cdots - w_{i_m}\right) \prod_{j=1}^{m} \left[g_{X^{k_{i_j}}}^{(i_j)}\left(w_{i_j}\right) f_{W_{i_j}}\left(w_{i_j}\right) dw_{i_j} \right] f_{W_{n+1}}(y) dy \Bigg\}.
$$

Hence, the result.

9.4 Simulation

Here, we illustrate computational efficiency of the expressions derived in Section 9.2. Computational efficiency is assessed in terms of time and accuracy.

Suppose $N(t)$ is a Poisson process with rate parameter λ, X_i are independent and identical exponential random variables with rate parameter λ and (X_i, W_i) have the common copula function $C(u, v) = uv + \theta uv(1-u)(1-v)$, the FGM copula function. We computed $E\left(S_t^k\right)$ by simulation and using Theorem 1 for every $k = 1, 2, \ldots, 100$. Note that the $k's$ value here is only chosen as an illustration in the simulation. The simulation was performed as follows:

1. simulate N from a Poisson distribution with parameter λ;
2. simulate $(U_1, V_1), \ldots, (U_N, V_N)$ from the FGM copula;
3. set $X_i = -\frac{1}{\lambda} \log(1 - U_i)$ and $W_i = -\frac{1}{\lambda} \log(1 - V_i)$ for $i = 1, \ldots, N$;
4. compute $S_t^k = (X_1 + \cdot + X_N)^k$;
5. repeat steps 1 to 4 for one million times;
6. compute the average of the values of S_t^k.

$E\left(S_t^k\right)$ was also computed using Theorem 1 for every $k = 1, 2, \ldots, 100$. The Maple software was used. Maple like other algebraic manipulation packages allows for arbitrary precision, so the accuracy of computed values was not an issue. That is, the values computed using Theorem 1 can be considered as exact.

The central processing unit times taken in seconds to compute $E\left(S_t^k\right)$ are plotted in Figure 9.1. The differences between the simulated and exact values of $E\left(S_t^k\right)$ are plotted in Figure 9.2. We have taken $\lambda = 1$, $\theta = -0.6, -0.2, 0.2, 0.6$ and $k = 1, \ldots, 100$.

We see that our expression in Theorem 1 is computationally more efficient for all values of k. The central processing unit times for the expression in Theorem 1 appear about two times smaller. The simulated values are computationally less accurate in addition to being computationally more time consuming. There is no evidence that the computational times or the computational accuracy change significantly with respect to k or θ.

For reasons of illustration, we have taken the copula to be the FGM copula. But the results in Figures 9.1 and 9.2 held for a wide range of copulas of the form (9.5) and for a wide range of parameter values. In particular, the central processing unit times for the expression in Theorem 1 always appeared about two times smaller. The simulated values were always computationally less accurate in addition to being computationally more time consuming.

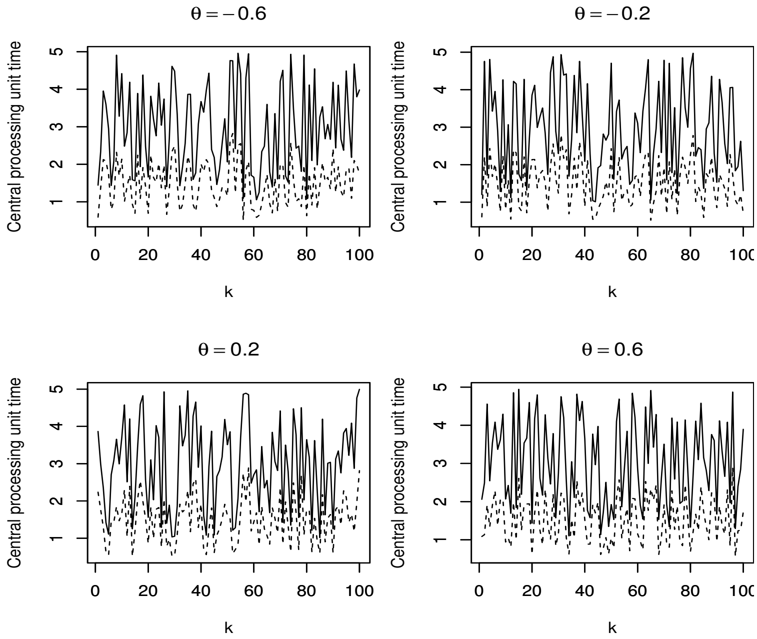

Figure 9.1: Central processing unit times in seconds taken to compute $E\left(S_t^k\right)$ by simulation (solid line) and by using Theorem 1 (broken line).

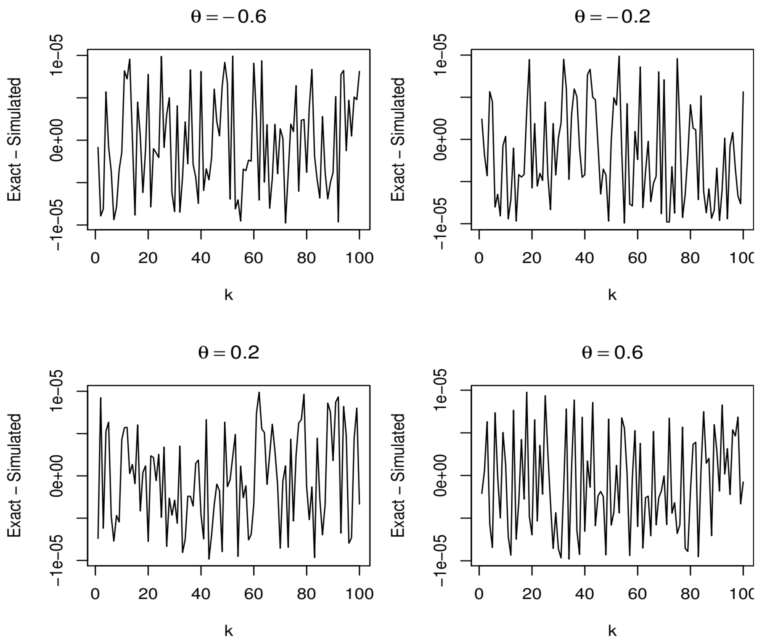

Figure 9.2: The exact minus the simulated values of $E\left(S_t^k\right)$ versus k.

Chapter 10

Conclusions

10.1 Conclusions

As metioned in the Introduction (Chapter 1), this thesis is a collection of selected works I have contributed to the field of distribution theory. The main chapters are presented in Chapter 2 through to Chapter 9.

Chapter 2 introduces and provides the reviews of the Generalized Hyperbolic distribution and its most popular univariate relatives. Chapter 3 and Chapter 4 construct the smallest order statistic study on the Pareto distribution and the Weibull distribution, repectively. We succeed in providing the simplified conditions for the comparison of smallest order statistics of two datasets from the same type of distribution but with different parameters. A real data application is also presented in each chapter. Chapter 5 can be thought of as a novel work where we construct seven new classes of discrete bivariate models. Even though, some of the newly proposed models outperform the models proposed by Lee and Cha previously in fitting the same data provided by them, some of the future works still need to done on those proposed models which are disscussed in later section. In Chapter 6, we contribute new and original expressions for the distribution of amuplitude, its moments and the distribution of phase for a wide range of bivariate distributions, not only the Gaussian models. Another theoretical contribution to the area of statistical distributions is presented in Chapter 7. We give the explicit expressions for the characteristic function of the product of two independent random variables, in which one follows the standard normal distribution, and the other follows one from a selection of nearly forty distributions. A potential practical area can be in signal transfer. In Chapter 8, we present the closed expressions for the distribution of aggregate risks, its TVaR and TVaR based allocation for a comprehensive collection of light-tailed, semi-heavy tailed, heavy tailed bivariate distributions which are widely used in finance. In the last chapter, we construct and present the derivation of the general form of the moment of the random sums using copulas. We let the number of claims, $N(t)$, be a renewal process, and let the claims X and time-interval, W, be independently and identically distributed. We also extend the case to non-identical random variables and to a wide ranges of different copulas.

Four of the main chapters have been published in refereed journals:

- Chapter 3 has been published as
  Nadarajah, S., Jiang, X. and Chu, J., (2017). Comparisons of smallest order statistics from Pareto distributions with different scale and shape parameters. Annals of Operations Research, pp.1-19.
- Chapter 5 has been published as
  Jiang, X., Chu, J. and Nadarajah, S., (2017). New classes of discrete bivariate distributions with application to football data. Communications in Statistics-Theory and Methods, 46(16), pp.8069-8085.
- Chapter 6 has been published as
  Nadarajah, S., Chu, J. and Jiang, X., (2016). Distributions of amplitude and phase for bivariate distributions. AEU-International Journal of Electronics and Communications, 70(9), pp.1249-1258.
- Chapter 8 has been published as
  Nadarajah, S., Chu, J. and Jiang, X., (2017). Aggregation and capital allocation formulas for bivariate distributions. Probability in the Engineering and Informational Sciences, pp.1-11.

10.2 Future work

Listed below are some aspects of future work which are based on the work presented in this thesis.

In Chapter 2, we have reviewed the univariate Generalized Hyperbolic (GH) and other twenty related distributions, as well as the related packages in the programming language, R. A future work is to produce a more comprehensive review including univariate, bivariate, multivariate, complex variate and matrix variate GH distributions, and the related packages and modules in different programming languages.

In Chapter 4, we have presented the comparisons of smallest order statistics from the standard and a lower truncated Weibull distribution, respectively. A future work is to build the comparisons of maximum order statistics or kth order statistics from Weibull distributions with different scale and shape parameters.

Predictions of the results of football matches can be considered one of the classical applications for discrete bivariate distributions. We have introduced seven new families of discrete bivariate distributions in Chapter 5. Some of the newly proposed models also provided better fits to the football data than those proposed by Lee and Cha (2015). However, none of them present a goodness-of-fit to the data. Moreover, it is better to have the same model family for each team with different parameter values. Besides, the dataset spans too many years. Over this time, each team may be expected to change substantially. It is suggested to model the time variation in team attack and defensive strengths.

In Chapter 7, we have derived a number of characteristic functions of the product of two independent but non-identical random variables. A future work can be to relax the assumption that one follows the standard Normal distribution to produce more combinations of products. Another future work could be to extend the characteristic function to the products of two dependent random variables.

Bibliography

[1] Cossette, H., Marceau, E., and Perreault, S. (2015). On two families of bivariate distributions with exponential marginals: Aggregation and capital allocation. Insurance: Mathematics and Economics, 64, 214-224.

[2] Lee, H., and Cha, J. H. (2015). On two general classes of discrete bivariate distributions. The American Statistician, 69(3), 221-230.

[3] Mao, Z. C., and Zhao, X. (2013). The expectation and variance of dependent random sums using copulas. IMA Journal of Management Mathematics, 25(4), 421-433.

[4] Nadarajah, S. (2015). Expansions for bivariate copulas. Statistics and Probability Letters, 100, 77-84.

[5] Schoenecker, S., and Luginbuhl, T. (2016). Characteristic Functions of the Product of Two Gaussian Random Variables and the Product of a Gaussian and a Gamma Random Variable. IEEE Signal Processing Letters, 23(5), 644-647.

[6] Torrado, N. (2015). Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters. Journal of the Korean Statistical Society, 44(1), 68-76.

[7] Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew Student's t-distribution. Journal of Financial Econometrics, 4, 275-309.

[8] Agrawal, R.(2006). Study of ultrasonic echo envelope based on Nakagami-inverse Gaussian distribution. Ultrasound in Medicine and Biology, 32(3), pp.371-376.

[9] Agrawal, R. and Karmeshu (2006). Ultrasonic backscattering in tissue: characterization through Nakagami-generalized inverse Gaussian distribution. Computers in Biology and Medicine, 37, 166-172.

[10] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

[11] Alexandrov, M. D. and Lacis, A. A. (2000). A new three-parameter cloud/aerosol particle size distribution based on the generalized inverse Gaussian density function. Applied Mathematics and Computation, 116, 153-165.

[12] Avramidis, A. N., L'Ecuyer, P. (2006). Efficient Monte Carlo and quasi–Monte Carlo option pricing under the variance gamma model. Management Science, 52(12), 1930-1944.

[13] Baciu, O. A. (2015). Generalized hyperbolic distributions: Empirical evidence on Bucharest stock exchange. Review of Finance and Banking, 7, 2067-3825.

[14] Bagnold, R. A. (1941). The Physics of Blown Sand and Desert Dunes. Methuen, London.

[15] Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 353, 401-409.

[16] Barndorff-Nielsen, O. (1978). Hyperbolic distributions and distributions on hyperbolae. Scandinavian Journal of Statistics, 5, 151-157.

[17] Barndorff-Nielsen, O. and Stelzer, R. (2004). Absolute moments of generalized hyperbolic distributions and approximate scaling of normal inverse Gaussian Lévy process. Scandinavian Journal of Statistics, 32, 617-637.

[18] Berger, T., Levy, W. B. and Xing, J. (2011). Energy efficient neurons with generalized inverse Gaussian inter spike interval durations. In: Proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing, pp. 1737-1742.

[19] Bhatia, J. C., Domnick, J., Durst, F. and Tropea, C. (1988). Phase-doppler-anemometry and the log-hyperbolic distribution applied to liquid sprays. Particle and Particle Systems Characterization, 5, 153-164.

[20] Blaesild, P. (1981). The two-dimensional hyperbolic distribution and related distributions, with an application to Johannsen's bean data. Biometrika, 68, 251-263.

[21] Browne, R. P. and McNicholas, P. D. (2015). A mixture of generalized hyperbolic distributions. Canadian Journal of Statistics, 43, 176-198.

[22] Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods and Research, 33, 261-304.

[23] Chinhamu, K., Huang, C. K. and Chikobvu, D. (2015). Evaluating risk in gold prices with generalized hyperbolic and stable distributions. In: Proceedings of the 57th Annual Conference of SASA, pp. 17–24.

[24] de Moivre, A. (1738). The Doctrine of Chances. ISBN 0821821032.

[25] Eriksson, A., Ghysels, E., Wang, F. (2009). The normal inverse gaussian distribution and the pricing of derivatives. Journal of Derivatives, 16(3), 23.

[26] Fajardo, J. and Farias, A. (2002). Generalized hyperbolic distributions and Brazilian data. Banco Central do Brasil Working Paper No. 52

[27] Fang, Y. (2011). Asymptotic equivalence between cross-validations and Akaike Information Criteria in mixed-effects models. Journal of Data Science, 9, 15-21.

[28] Gauss, C. F. (1809). Theoria motvs corporvm coelestivm in sectionibvs conicis Solem ambientivm (in Latin).

[29] Gómez-Déniz, E., Calderín-Ojeda, E. and Sarabia, J. M. (2013). Gamma-generalized inverse Gaussian class of distributions with applications. Communications in Statistics - Theory and Methods, 42, 919-933.

[30] Gómez-Déniz, E., Vázquez-Polo, F. J., García, V. and Calderín-Ojeda, E. (2015). The exponential reciprocal generalized inverse Gaussian distribution: A new model to describe loss distributions. In: Current Topics on Risk Analysis, editors M. Guillén, A. A. Juan, H. Ramalhinho, I. Serra and C. Serrat, pp. 377-384.

[31] Gosset, W. S. (1908). The probable error of a mean. Biometrika, 6, 1-25.

[32] Gradshteyn, I. S. and Ryzhik, I. M. (2000). Table of Integrals, Series, and Products, sixth edition. Academic Press, San Diego.

[33] Gupta, A. and Karmeshu (2015). Study of compound generalized Nakagami-generalized inverse Gaussian distribution and related densities: Application to ultrasound imaging. Computational Statistics, 30, 81-96.

[34] Gupta, R. C. and Viles, W. (2011). Roller-coaster failure rates and mean residual life functions with application to the extended generalized inverse Gaussian model. Probability in the Engineering and Informational Sciences, 25, 103-118.

[35] Hu, W., Kercheval, A. N. (2010). Portfolio optimization for student t and skewed t returns. Quantitative Finance, 10(1), 91-105.

[36] Huang, C. K., Chinhamu, K., Huang, C. S. and Hammujuddy, J. (2014). Generalized hyperbolic distributions and value-at-risk estimation for the South African mining index. The International Business and Economics Research Journal, 13, 319-

[37] Hurst, S.R., Platen, E. and Rachev, S.R. (1997). Subordinated Markov models: a comparison. Financial Eng.Japanese Markets 4, 97-124.

[38] Iyengar, S. and Liao, Q. (1997). Modeling neural activity using the generalized inverse Gaussian distribution. Biological Cybernetics, 77, 289-295.

[39] Jones, M. C., Faddy, M. J. (2003). A skew extension of the t-distribution, with applications. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65(1), 159-174

[40] Jϕrgensen, B. (2012). Statistical properties of the generalized inverse Gaussian distribution (Vol. 9). Springer Science Business Media.

[41] Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation of the normal–inverse Gaussian distribution. Statistics probability letters, 57(1), 43-52.

[42] Karlis, D., Lilleställ, J. (2004). Bayesian estimation of NIG models via Markov chain Monte Carlo methods. Applied Stochastic Models in Business and Industry, 20(4), 323-338.

[43] Klebanov, L. B. and Rachev, S. T. (2015). ν-generalized hyperbolic distributions. arXiv:1502.02486

[44] Korolev, V. Yu., Korchagin, A. Yu. and Sokolov, I. A. (2015). Modeling of statistical regularities in financial markets by generalized variance gamma distributions. Inform. Primen., 9, 14-28.

[45] Lemonte, A. J. and Cordeiro, G. M. (2011). The exponentiated generalized inverse Gaussian distribution, Statistics and Probability Letters, 81, 506-517.

[46] Li, H. Q. and Wu, L. C. (2007). The statistical analysis of the log-return series of the Chinese stock prices: An application of the generalized hyperbolic distributions. Journal of Data Analysis, 2, 161-178.

[47] Madan, D. B. and Seneta, E. (1990). The variance gamma model for share market returns. The Journal of Business, 63, 511-524.

[48] Madan, D. B., Carr, P. P., Chang, E. C. (1998). The variance gamma process and option pricing. Review of Finance, 2(1), 79-105.

[49] Nakajima, J., Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution. Computational Statistics Data Analysis, 56(11), 3690-3704.

[50] Nwobi, F. N. (2014). Modeling electricity price returns using generalized hyperbolic distributions. Communications in Mathematical Finance, 3, 33-50.

[51] Predota, M. (2005). On European and Asian option pricing in the generalized hyperbolic model. European Journal of Applied Mathematics, 16, 111-144.

[52] R Development Core Team (2016). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

[53] Rege, S. and de Menezes, A. G. (2012). Comparing the generalised hyperbolic and the normal inverse Gaussian distributions for the daily returns of the PSI20. CEEAplA Working Paper No. 05/2012.

[54] Saboor, A., Khan, A. and Mubeen, S. (2014). A Ψ_2 hypergeometric generalized inverse Gaussian distribution. Bothalia Journal, 44, 18-25.

[55] Saboor, A., Provost, S. B. and Ahmad, M. (2015). A confluent hypergeometric generalized inverse Gaussian distribution. Pakistan Journal of Statistics, 31, 21-32.

[56] Schwarz, G. E. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464.

[57] Scott, D. J., Wurtz, D., Dong, C. and Tran, T. T. (2011). Moments of the generalized hyperbolic distribution. Computational Statistics, 26, 459-476.

[58] Shakil, M., Kibria, B. M. G. and Singh, J. N. (2010). A new family of distributions based on the generalized Pearson differential equation with some applications. Austrian Journal of Statistics, 39, 259-278.

[59] Shankar, P.M. (2001). Ultrasonic tissue characterization using a generalized Nakagami model. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 48(6), pp.1716-1720.

[60] Sichel, H. S., Dohm, C. E. and Kleingeld, W. J. (1997). The logarithmic generalized inverse Gaussian distribution. South African Statistical Journal, 31, 125-149.

[61] Silva, R. S., Lopes, H. F. and Migon, H. S. (2006). The extended generalized inverse Gaussian distribution for log-linear and stochastic volatility models. Brazilian Journal of Probability and Statistics, 20, 67-91.

[62] Socgnia, V. K. and Wilcox, D. (2014). A comparison of generalized hyperbolic distribution models for equity returns. Journal of Applied Mathematics, Article ID 263465, 15 pages.

[63] Stentoft, L. (2008). American option pricing using GARCH models and the normal inverse Gaussian distribution. Journal of Financial Econometrics, 6(4), 540-582.

[64] Venter, J. H., de Jongh, P. J. (2002). Risk estimation using the normal inverse Gaussian distribution. Journal of Risk, 4, 1-24.

[65] Weiss, M. (1984). A note on the role of generalized inverse Gaussian distributions of circulatory transit times in pharmacokinetics. Journal of Mathematical Biology, 20, 95-102.

[66] Weron, R. (1999). Pricing options on dividend paying instruments under the generalized hyperbolic model. Research Report HSC/99/1

[67] Zaks, L. M. and Korolev, V. Yu. (2013). Variance-generalized-gamma-distributions as limit laws for random sums. Inform. Primen., 7, 105-115.

[68] Zhu, D., Galbraith, J. W. (2011). Modeling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. Journal of Empirical Finance, 18(4), 765-778.

[69] Asimit, A. V., Bignozzi, V., Cheung, K. C., Hu, J., Kim, E. S. (2017). Robust and Pareto optimality of insurance contracts. European Journal of Operational Research, 262(2), 720-732.

[70] Bae, S. J., Kim, S. J., Lee, J. H., Song, I., Kim, N. I., Seo, Y., Kim, K. B., Lee, N. and Park, J. Y. (2014). Degradation pattern prediction of a polymer electrolyte membrane fuel cell stack with series reliability structure via durability data of single cells. Applied Energy, 131, 48-55.

[71] Balakrishnan, N. and Zhao, P. (2013). Ordering properties of order statistics from heterogeneous populations: A review with an emphasis on some recent developments (with discussion). Probability in the Engineering and Informational Sciences, 27, 403-467.

[72] Beirlant, J., Maribe, G., Verster, A. (2018). Penalized bias reduction in extreme value estimation for censored Pareto-type data, and long-tailed insurance applications. Insurance: Mathematics and Economics, 78, 114-122.

[73] Cai, J., Liu, H., Wang, R. (2017). Pareto-optimal reinsurance arrangements under general model settings. Insurance: Mathematics and Economics, 77, 24-37.

[74] Chevalier, J., Goolsbee, A. (2003). Measuring prices and price competition online: Amazon. com and BarnesandNoble. com. Quantitative marketing and Economics, 1(2), 203-222.

[75] Fang, L. and Tang, W. (2014). On the right spread ordering of series systems with two heterogeneous Weibull components. Journal of Inequalities and Applications, doi: 10.1186/1029-242X-2014-190

[76] Grahovac, D. (2018). Densities of ruin-related quantities in the Cramér-Lundberg model with Pareto claims. Methodology and computing in applied probability, 20(1), 273-288.

[77] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate Distributions, volume 1. John Wiley and Sons, New York.

[78] Khaledi, B. -E. and Kochar, S. C. (2006). Weibull distribution: Some stochastic comparisons results. Journal of Statistical Planning and Inference, 136, 3121-3129.

[79] Li, C. and Li, X. (2015). Likelihood ratio order of sample minimum from heterogeneous Weibull random variables. Statistics and Probability Letters, 97, 46-53.

[80] Liu, J., Chen, W., Zhang, Y. J. and Cao, Z. (2013). A utility maximization framework for fair and efficient multicasting in multicarrier wireless cellular networks. IEEE/ACM Transactions on Networking, 21, 110-120

[81] Morales, M. (2005). On an approximation for the surplus process using extreme value theory: Applications in ruin theory and reinsurance pricing. Northe American Journal, 8: 46-66.

[82] R Development Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

[83] Ramos Romero, H. M. and Sordo Diaz, M. A. (2001). The proportional likelihood ratio order and applications. Questiio, 25, 211-223.

[84] Shaked, M. and Shanthikumar, J. G. (2007). Stochastic Orders. Springer Verlag, New York.

[85] Torrado, N. (2015). Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters. Journal of the Korean Statistical Society, 44, 68-76.

[86] Warin, T. and Leiter, D. (2012). Homogenous goods markets: An empirical study of price dispersion on the internet. International Journal of Economics and Business Research, 4, 514-529.

[87] Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. Econometrica, 64, 1183-1193.

[88] Balakrishnan, N. and Zhao, P. (2013). Ordering properties of order statistics from heterogeneous populations: A review with an emphasis on some recent developments (with discussion). Probability in the Engineering and Informational Sciences, 27, 403-467.

[89] Barrett, G. F. and Donald, S. G. (2003). Consistent tests for stochastic dominance. Econometrica, 71, 71-104.

[90] Fang, L. and Tang, W. (2014). On the right spread ordering of series systems with two heterogeneous Weibull components. Journal of Inequalities and Applications, doi: 10.1186/1029-242X-2014-190

[91] Khaledi, B. -E. and Kochar, S. C. (2006). Weibull distribution: Some stochastic comparisons results. Journal of Statistical Planning and Inference, 136, 3121-3129.

[92] Li, C. and Li, X. (2015). Likelihood ratio order of sample minimum from heterogeneous Weibull random variables. Statistics and Probability Letters, 97, 46-53.

[93] Murthy, D. N. P., Xie, M. and Jiang, R. (2003). Weibull Models. John Wiley and Sons, New York.

[94] R Development Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

[95] Shaked, M. and Shanthikumar, J. G. (2007). Stochastic Orders. Springer Verlag, New York.

[96] Torrado, N. (2015). Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters. Journal of the Korean Statistical Society, 44, 68-76.

[97] Ahooyi, T. M., Arbogast, J. E., Oktem, U. G., Seider, W. D. and Soroush, M. (2014). Estimation of complete discrete multivariate probability distributions from scarce data with application

to risk assessment and fault detection. Industrial and Engineering Chemistry Research, 53, 7538-7547.

[98] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

[99] Antony, D. and Jayakumar, K. (2013). Bivariate discrete Linnik distribution. Statistica.

[100] Bairamov, I. and Gultekin, O. E. (2010). Discrete distributions connected with the bivariate binomial. Hacettepe Journal of Mathematics and Statistics, 39, 109-120.

[101] Becker, N. G. and Utev, S. (2002). Multivariate discrete distributions with a product-type dependence. Journal of Multivariate Analysis, 83, 509-524.

[102] Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods and Research, 33, 261-304.

[103] Cacoullos, T. and Papageorgiou, H. (2012). On bivariate discrete distributions generated by compounding. In: Statistical Distributions in Scientific Work, volume 4, Models, Structures, and Characterizations, editors C. Taillie, G. P. Patil and B. A. Baldessari.

[104] Fang, Y. (2011). Asymptotic equivalence between cross-validations and Akaike Information Criteria in mixed-effects models. Journal of Data Science, 9, 15-21.

[105] Garrido, L. R., Birchmeier, A. N., Munilla, S. and Cantet, R. J. C. (2008). Estimation of effective population size using bivariate discrete distributions for modeling family size in beef cattle. Livestock Science, 117, 43-51.

[106] Gomez-Deniz, E., Sarabia, J. M. and Balakrishnan, N. (2012). A multivariate discrete Poisson-Lindley distribution: Extensions and actuarial applications. ASTIN Bulletin, 42, 655-678.

[107] He, Q. -M. and Ren, J. (2015). Parameter estimation of discrete multivariate phase-type distributions. Methodology and Computing in Applied Probability, doi: 10.1007/s11009-015-9442-y

[108] Holgate, B. (1964). Estimation for the bivariate Poisson distribution. Biometrika, 51, 241-245.

[109] Hoshino, N. (2006). A discrete multivariate distribution resulting from the law of small numbers. Journal of Applied Probability, 43, 852-866.

[110] Ip, E. H. and Wang, Y. J. (2009). Canonical representation of conditionally specified multivariate discrete distributions. Journal of Multivariate Analysis, 100, 1282-1290.

[111] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997). Discrete Multivariate Distributions. John Wiley and Sons, New York.

[112] Kumar, C. S. (2007). A unified approach to bivariate discrete distributions. Metrika, 67, 113-123.

[113] Lee, H. and Cha, J. H. (2015). On two general classes of discrete bivariate distributions. The American Statistician, 69, 221-230.

[114] McHale, I. and Scarf, P. (2007). Modelling soccer matches using bivariate discrete distributions with general dependence structure. Statistica Neerlandica, 61, 432-445.

[115] McHale, I. and Scarf, P. (2011). Modelling the dependence of goals scored by opposing teams in international soccer matches. Statistical Modelling, 11, 219-236.

[116] Miethe, T. D., Hart, T. C. and Regoeczi, W. C. (2008). The conjunctive analysis of case configurations: An exploratory method for discrete multivariate analyses of crime data. Journal of Quantitative Criminology, 24, 227-241.

[117] Nikoloulopoulos, A. K. (2013). Copula-based models for multivariate discrete response data. In: Copulae in Mathematical and Quantitative Finance, volume 213 of the series Lecture Notes in Statistics, pp. 231-249.

[118] Nikulin, M. S., Smirnovc, T. I. and Voinov, V. G. (2002). Multivariate discrete distributions induced by an urn scheme, linear diophantine equations, unbiased estimating, testing and applications. Journal of Statistical Planning and Inference, 101, 255-266.

[119] Ong, S. H. and Ng, C. M. (2013). A bivariate generalization of the noncentral negative binomial distribution. Communications in Statistics—Simulation and Computation, 42, 570-585.

[120] Ozel, G. (2011a). A bivariate compound Poisson model for the occurrence of foreshocks and aftershock sequences in Turkey. Environmetrics, 22, 847-856.

[121] Ozel, G. (2011b). On certain properties of a class of bivariate compound Poisson distributions and an application on earthquake data. Revista Colombiana de Estadistica, 34, 545-566.

[122] Ozel, G. (2013). On the moment characteristics of univariate compound Poisson and bivariate compound Poisson processes with applications. Revista Colombiana Estadistica, 36, 59-77.

[123] Panagiotelis, A., Czado, C. and Joe, H. (2012). Pair copula constructions for multivariate discrete data. Journal of the American Statistical Association, 107, 1063-1072.

[124] Piperigou, V. E. and Papageorgiou, H. (2003). On truncated bivariate discrete distributions: A unified treatment. Metrika, 58, 221-233.

[125] R Development Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

[126] Reilly, C. H. and Sapkota, N. (2015). A family of composite discrete bivariate distributions with uniform marginals for simulating realistic and challenging optimization-problem instances. European Journal of Operational Research, 241, 642-652.

[127] Schwarz, G. E. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464.

[128] Tonda, T. (2005). A class of multivariate discrete distributions based on an approximate density in GLMM. Hiroshima Mathematical Journal, 35, 327-349.

[129] Aalo, V., Efthymoglou, G. and Chayawan, C. (2007). On the envelope and phase distributions for correlated Gaussian quadratures. IEEE Communications Letters, 11, 985-987.

[130] Aravkin, A., Burke, J. V., Chiuso, A. and Pillonetto, G. (2011). Convex vs nonconvex approaches for sparse estimation: Lasso, multiple kernel learning and hyperparameter Lasso. In: Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, pp. 156-161.

[131] Arnold, B. C. and Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and / or selective reporting. Test, 11, 1-54.

[132] Arnold, B. C. Castillo, E. and Sarabia, J. M. (2002). Conditionally specified multivariate skewed distributions. Sankhyā, 64, 206-226.

[133] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. Journal of the Royal Statistical Society B, 65, 367-389.

[134] Balakrishnan, N. and Lai, C. -D. (2009). Continuous Bivariate Distributions. Springer Verlag, New York.

[135] Benedetto, S. and Steila, O. (1977). Analysis of an hybrid amplitude and phase modulation scheme for analog data-transmission. AEU-International Journal of Electronics and Communications, 31, 15-20.

[136] Brummer, G., Rafique, R. and Ohki, T. A. (2011). Phase and amplitude modulator for microwave pulse generation. IEEE Transactions on Applied Superconductivity, 21, 583-586.

[137] Chen, Y. -M. and Ueng, Y. -L. (2013). Noncoherent amplitude/phase modulated transmission schemes for Rayleigh block fading channels. IEEE Transactions on Communications, 61, 217-227.

[138] Coluccia, A. (2013). On the expected value and higher-order moments of the Euclidean norm for elliptical normal variates. IEEE Communications Letters, 17, 2364-2367.

[139] Dharmawansa, P., Rajatheva, N. and Tellambura, C. (2009). Envelope and phase distribution of two correlated Gaussian variables. IEEE Transactions on Communications, 57, 915-921.

[140] Dickey, J. M. (1968). Three multidimensional integral identities with Bayesian applications. Annals of Mathematical Statistics, 39, 1615-1627.

[141] Dowhuszko, A. and Hamalainen, J. (2015). Performance of transmit beamforming codebooks with separate amplitude and phase quantization. IEEE Signal Processing Letters, 22, 813-817.

[142] Eltoft, T., Kim, T. and Lee, T. -W. (2006). On the multivariate Laplace distribution. IEEE Signal Processing Letters, 13, 300-303.

[143] Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. Biometrika, 47, 307-323.

[144] Fathi-Vajargah, B. and Hasanalipour, P. (2013). Introducing a novel bivariate generalized skew-symmetric normal distribution. Journal of Mathematics and Computer Science, 7, 266-271.

[145] Gradshteyn, I. S. and Ryzhik, I. M. (2014). Tables of Integrals, Series and Products, eighth edition. Academic Press, New York.

[146] Gumbel, E. J. (1960). Bivariate exponential distributions. Journal of the American Statistical Association, 55, 698-707.

[147] Herben, M. H. A. J. and Maanders, E. J. (1979). Radiation-pattern in amplitude and phase of an offset fed paraboloidal reflector antenna. AEU-International Journal of Electronics and Communications, 33, 413-414.

[148] Kersten, P. R. and Anfinsen, S. N. (2012). A flexible and computationally efficient density model for the multilook polarimetric covariance matrix. In: Proceedings of the 9th European Conference on Synthetic Aperture Radar, pp. 760-763.

[149] Kotz, S., Kozubowski, T. and Podgorski, K. (2001). The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance. Birkhäuser Boston, Boston.

[150] Kotz, S. and Nadarajah, S. (2004). Multivariate T-Distributions and Their Applications. Cambridge University Press, New York.

[151] Krishnamurthy, S., Bliss, D. Richmond, C. and Tarokh, V. (2015). Peak sidelobe level Gumbel distribution for arrays of randomly placed antennas. In: Proceedings of the 2015 IEEE Radar Conference, pp. 1671-1676.

[152] Li, J. and Manikopoulos, C. (2004). Novel statistical network model: The hyperbolic distribution. IEE Proceedings-Communications, 151, 539-548.

[153] Li, W., Wang, W. T. and Zhu, N. H. (2014). Broadband microwave photonic splitter with arbitrary amplitude ratio and phase shift. IEEE Photonics Journal, 6, Article Number 5501507.

[154] Louzada, F., Ara, A. and Fernandes, G. (2016). The bivariate alpha-skew-normal distribution. Communications in Statistics - Theory and Methods, doi: 10.1080/03610926.2015.1024865

[155] Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. Mitteilungsblatt für Mathematische Statistik, 8, 234-235.

[156] Nguyen, T. M. and Wu, Q. M. J. (2012). Robust Student's t mixture model with spatial constraints and its application in medical image segmentation. IEEE Transactions on Medical Imaging, 31, 103-116.

[157] Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). Integrals and Series, volumes 1, 2 and 3. Gordon and Breach Science Publishers, Amsterdam.

[158] R Development Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[159] Sarmanov, I. O. (1966). Generalized normal correlation and two dimensional Fréchet classes. Doklady Mathematics, 7, 596-599.

[160] Tran, T. Q. and Sharma, S. K. (2012). Radiation characteristics of a multimode concentric circular microstrip patch antenna by controlling amplitude and phase of modes. IEEE Transactions on Antennas and Propagation, 60, 1601-1605.

[161] Vijaykumar, M. and Vasudevan, V. (2016). A skew-normal canonical model for statistical static timing analysis. IEEE Transactions on Very Large Scale Integration Systems, doi: 10.1109/TVLSI.2015.2501370

[162] Wan, T., Canagarajah, N. and Achim, A. (2011). Segmentation of noisy colour images using Cauchy distribution in the complex wavelet domain. IET Image Processing, 5, 159-170.

[163] Xu, K., Wu, X. R., Sung, J. Y., Cheng, Z. Z., Chow, C. W., Song, Q. H. and Tsang, H. K. (2016). Amplitude and phase modulation of UWB monocycle pulses on a silicon photonic chip. IEEE Photonics Technology Letters, 28, 248-251.

[164] Yacoub, M. D. (2010). Nakagami-Phase-Envelope joint distribution: A new model. IEEE Transactions on Vehicular Technology, 59, 1552-1557.

[165] Yi, X., Huang, T. X. H. and Minasian, R. A. (2011). Photonic beamforming based on programmable phase shifters with amplitude and phase control. IEEE Photonics Technology Letters, 23, 1286-1288.

[166] Zhang, F., Lin, W., Chen, Z. and Ngan, K. N. (2013). Additive log-logistic model for networked video quality assessment. IEEE Transactions on Image Processing, 22, 1536-1547.

[167] Zhang, Y. and Ruan, X. (2015). AC-AC converter with controllable phase and amplitude. IEEE Transactions on Power Electronics, 29, 6235-6244.

[168] Albrecht, H. (1990). Search for hadronic $b \to u$ decays. Physics Letters B, 241, 278-282.

[169] Azzalini, A. (1985). A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12, 171-178.

[170] Bain, L. J. (1974). Analysis for the linear failure-rate life-testing distribution. Technometrics, 16, 551-559.

[171] Borghi, O. (1965). Sobre una distribución de frecuencias. Trabajos de Estadistica, 16, 171-192.

[172] Cohen, L. (1996). A general approach for obtaining joint representations in signal analysis. I. Characteristic function operator method. IEEE Transactions on Signal Processing, 44(5), 1080-1090.

[173] de Moivre, A. (1738). The Doctrine of Chances. ISBN 0821821032.

[174] Gauss, C. F. (1809). Theoria motvs corporvm coelestivm in sectionibvs conicis Solem ambientivm (in Latin).

[175] Gosset, W. S. (1908). The probable error of a mean. Biometrika, 6, 1-25.

[176] Gradshteyn, I. S. and Ryzhik, I. M. (2000). Table of Integrals, Series, and Products, sixth edition. Academic Press, San Diego.

[177] Hall, P. (1927). The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable. Biometrika, 19, 240-245.

[178] Hamming, R. W. (1970). On the distribution of numbers. The Bell System Technical Journal, 49, 1609-1625.

[179] Holla, M. S. and Bhattacharya, S. K. (1968). On a compound Gaussian distribution. Annals of the Institute of Statistical Mathematics, 20, 331-336.

[180] Irwin, J. O. (1927). On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's type II. Biometrika, 19, 225-239.

[181] Jiang, G. J., Knight, J. L. (2002). Estimation of continuous-time processes via the empirical characteristic function. Journal of Business Economic Statistics, 20(2), 198-212.

[182] Kozubowski, T. J. and Podgorski, K. (2000). A multivariate and asymmetric generalization of Laplace distribution. Computational Statistics, 15, 531-540.

[183] Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46, 79-88.

[184] Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: Different methods of estimations. Computational Statistics and Data Analysis, 49, 187-200.

[185] Laplace, P. -S. (1774). Mémoire sur la probabilité des causes par les évènements. Mémoires de l'Academie Royale des Sciences Presentés par Divers Savan, 6, 621-656.

[186] Leone, F. C., Nottingham, R. B. and Nelson, L. S. (1961). The folded normal distribution. Technometrics, 3, 543-550.

[187] Madan, D. B. and Seneta, E. (1990). The variance gamma (V.G.) model for share market returns. Journal of Business, 63, 511-524.

[188] Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. Part I. On the motions and collisions of perfectly elastic spheres. Philosophical Magazine, 19, 19-32.

[189] McGill, W. J. (1962). Random fluctuations of response rate. Psychometrika, 27, 3-17.

[190] Moffat, A. F. J. (1969). A theoretical investigation of focal stellar images in the photographic emulsion and application to photographic photometry. Astronomy and Astrophysics, 3, 455-461.

[191] Nakagami, M. (1960). The m-distribution, a general formula of intensity of rapid fading. In: Statistical Methods in Radio Wave Propagation, editor W. C. Hoffman, pp. 3-36, Pergamon Press.

[192] Pareto, V. (1964). Cours d'Économie Politique: Nouvelle édition par G. -H. Bousquet et G. Busino, Librairie Droz, Geneva, pp. 299-345.

[193] Poiraud-Casanova, S. and Thomas-Agnan, C. (2000). About monotone regression quantiles. Statistics and Probability Letters, 48, 101-104.

[194] Prudnikov, A. P., Brychkov, Y. A. and Marichev, O. I. (1986). Integrals and Series, volumes 1, 2 and 3. Gordon and Breach Science Publishers, Amsterdam.

[195] Rice, S. O. (1945). Mathematical analysis of random noise. The Bell System Technical Journal, 24, 46-156.

[196] Schoenecker, S. and Luginbuhl, T. (2016). Characteristic functions of the product of two Gaussian random variables and the product of a Gaussian and a gamma random variable. IEEE Signal Processing Letters, 23, 644-647.

[197] Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. Journal of Econometrics, 102(1), 111-141.

[198] Tsallis, C. (2009). Nonadditive entropy and nonextensive statistical mechanics-An overview after 20 years. Brazilian Journal of Physics, 39, 337-356.

[199] Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanics, 18, 293-297.

[200] Zhu, D. and Galbraith, J. W. (2010). A generalized asymmetric Student-t distribution with application to financial econometrics. Journal of Econometrics, 157, 297-305.

[201] Arnold, B. C. and Strauss, D. (1991). Bivariate distributions with conditionals in prescribed exponential families. Journal of the Royal Statistical Society, B, 53, 365-375.

[202] Balakrishnan, N. and Lai, C. D. (2009). Continuous Bivariate Distributions, second edition. Springer Verlag, New York.

[203] Barges, M., Cossette, H. and Marceau, E. (2009). TVaR-based capital allocation with copulas. Insurance: Mathematics and Economics, 45, 348-361.

[204] Becker, P. J. and Roux, J. J. J. (1981). A bivariate extension of the gamma distribution. South African Statistical Journal, 15, 1-12.

[205] Block, H. W. and Basu, A. P. (1976). A continuous bivariate exponential distribution. Journal of the American Statistical Association, 64, 1031-1037.

[206] Chacko, M. and Thomas, P. Y. (2007). Estimation of a parameter of bivariate Pareto distribution by ranked set sampling. Journal of Applied Statistics, 34, 703-714.

[207] Cheriyan, K. C. (1941). A bivariate correlated gamma-type distribution function. Journal of the Indian Mathematical Society, 5, 133-144.

[208] Chiragiev, A. and Landsman, Z. (2007). Multivariate Pareto portfolios: TCE based capital allocation and divided differences. Scandinavian Actuarial Journal, 261-280.

[209] Cossette, H., Cote, M. -P., Marceau, E. and Moutanabbir, K. (2013). Risk measures and capital allocation using the Farlie-Gumbel-Morgenstern copula. Insurance: Mathematics and Economics, 52, 560-572.

[210] Cossette, H., Mailhot, M. and Marceau, E. (2012). TVaR-based capital allocation for multivariate compound distributions. Insurance: Mathematics and Economics, 50, 247-256.

[211] Cossette, H., Marceau, E. and Perreault, S. (2015). On two families of bivariate distributions with exponential marginals: Aggregation and capital allocation. Insurance: Mathematics and Economics, 64, 214-224.

[212] Crovelli, R. A. (1973). A bivariate precipitation model. Journal of the American Meteorological Society, 1, 130-134.

[213] Dhaene, J., Henrard, L., Landsman, Z., Vandendorpe, A. and Vanduffel, S. (2008). Some results on the CTE-based capital allocation rule. Insurance: Mathematics and Economics, 42, 855-863.

[214] Dussauchoy, A. and Berland, R. (1975). A multivariate gamma-type distribution whose marginal laws are gamma, and which has a property similar to a characteristic property of the normal case. In: A Modern Course on Distributions in Scientific Work, volume I: Models and Structures, editors G. P. Patil, S. Kotz and J. K. Ord, pp. 319-328, Reidel, Dordrecht.

[215] El-Bassiouny, A. H. and Jones, M. C. (2009). A bivariate F distribution with marginals on arbitrary numerator and denominator degrees of freedom, and related bivariate beta and t distributions. Statistical Methods and Applications, 18, 465-481.

[216] Furman, E. and Landsman, Z. (2005). Risk capital decomposition for a multivariate dependent gamma portfolio. Insurance: Mathematics and Economics, 37, 635-649.

[217] Furman, E. and Landsman, Z. (2008). Economic capital allocations for non-negative portfolios of dependent risks. ASTIN Bulletin, 38, 601-619.

[218] Hesselager, O. (1994). Recursions for certain bivariate counting distributions and their compound distributions. University of Copenhagen, Laboratory of Actuarial Mathematics.

[219] Jones, M. C. (2002). A dependent bivariate t distribution with marginals on different degrees of freedom. Statistics and Probability Letters, 56, 163-170.

[220] Landsman, Z. M. and Valdez, E. A. (2003). Tail conditional expectations for elliptical distributions. North American Actuarial Journal, 7, 55-71.

[221] Lawrance, A. J. and Lewis, P. A. W. (1980). The exponential autoregressive-moving average EARMA(p, q) process. Journal of the Royal Statistical Society, B, 42, 150-161.

[222] Lee, H. and Cha, J. H. (2014). On construction of general classes of bivariate distributions. Journal of Multivariate Analysis, 127, 151-159.

[223] Lee, P. A. (1981). The correlated bivariate inverted beta distribution. Biometrical Journal, 23, 693-703.

[224] Mardia, K. V. (1970). Families of Bivariate Distributions. Griffin, London.

[225] Mirhosseini, S. M., Amini, M., Kundu, D. and Dolati, A. (2015). On a new absolutely continuous bivariate generalized exponential distribution. Statistical Methods and Applications, 24, 61-83.

[226] Mohsin, M., Gebhardt, A., Pilz, J. and Spock, G. (2013). A new bivariate gamma distribution generated from functional scale parameter with application to drought data. Stochastic Environmental Research and Risk Assessment, 27, 1039-1054.

[227] Nadarajah, S. and Gupta, A. K. (2006). Some bivariate gamma distributions. Applied Mathematics Letters, 19, 767-774.

[228] Nagar, D. K., Bran-Cardona, P. A. and Gupta, A. K. (2009). Multivariate generalization of the hypergeometric function type I distribution. Acta Applicande Mathematicae, 105, 111-122.

[229] Nagar, D. K. and Sepulveda-Murillo, F. H. (2011). Properties of the bivariate confluent hypergeometric function kind 1 distribution. Revista de la Union Matematica Argentina, 52, 11-21.

[230] Panjer, H. H. (2002). Measurement of risk, solvency requirements and allocation of capital within financial conglomerates. Research Report, Institute of Insurance and Pension Research, University of Waterloo.

[231] Tasche, D. (1999). Risk contributions and performance measurement. Working Paper, Technische Universitatt Mnchen.

[232] Unnikrishnan Nair, N. and Sankaran, P. G. (2014). Modelling lifetimes with bivariate Schur-constant equilibrium distributions from renewal theory. Metron, 72, 331-349.

[233] Albrecher, H. and Teugels, J. (2006). Exponential behavior in the presence of dependence in risk theory. Journal of Applied Probability, 43, 265-285.

[234] Ambagaspitiya, R. S. (2009). Ultimate ruin probability in the Sparre Andersen model with dependent claim sizes and claim occurrence times. Insurance: Mathematics and Economics, 44, 464-472.

[235] Asimit, V. and Badescu, A. (2010). Extremes on the discounted aggregate claims in a time dependent risk model. Scandinavian Actuarial Journal, 93-104.

[236] Asimit, V. and Chen, Y. (2015). Asymptotic results for conditional measures of association of a random sum. Insurance: Mathematics and Economics, 60, 11-18.

[237] Asimit, V., Hashorva, E. and Kortschak, D. (2014a). Tail asymptotics of randomly weighted large risks. arXiv: 1405.0593

[238] Asimit, V., Hashorva, E. and Kortschak, D. (2014b). Asymptotic tail probability of randomly weighted large risks. arXiv: 1405.0593

[239] Asimit, V., Hashorva, E. and Kortschak, D. (2016). Aggregation of randomly weighted large risks. IMA Journal of Management Mathematics, doi: 10.1093/imaman/dpv020

[240] Asimit, V. and Jones, B. (2008a). Dependence and the asymptotic behavior of large claims reinsurance. Insurance: Mathematics and Economics, 43, 407-411.

[241] Asimit, V. and Jones, B. (2008b). Asymptotic tail probabilities for large claims reinsurance of a portfolio of dependent risks. ASTIN Bulletin, 38, 147-159.

[242] Chueng, E. C. K., Landriault, D., Willmot, G. E. and Woo, J. K. (2010). Structural properties of Gerber-Shiu functions in dependent Sparre Andersen models. Insurance: Mathematics and Economics, 46, 117-126.

[243] Cossette, H., Marceau, E. and Fouad, M. F. (2008). On the compound Poisson risk model with dependence based on a generalized Farlie-Gumbel-Morgenstern copula. Insurance: Mathematics and Economics, 43, 444-455.

[244] Harvey, C. R., Liechty, J. C., Liechty, M. W. and Müller, P. (2010). Portfolio selection with higher moments. Quantitative Finance, 10, 469-485.

[245] Mao, Z. -C. and Zhao, X. (2013). The expectation and variance of dependent random sums using copulas. IMA Journal of Management Mathematics, 25, 421-433.

[246] Marceau, E. (2008). On the discrete-time compound renewal risk model with dependence. Insurance: Mathematics and Economics, 44, 245-259.

[247] Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. Mitteilingsblatt für Mathematishe Statistik, 8, 234-235.

[248] Nadarajah, S. (2015). Expansions for bivariate copulas. Statistics and Probability Letters, 100, 77-84.

[249] Nelsen, R. B. (2006). An Introduction to Copulas, second edition. Springer Verlag, New York.

[250] Papush, D. E., Patrik, G. S., Podgaits, F. (2001, March). Approximations of the aggregate loss distribution. In CAS Forum (No. Winter, pp. 175-186).

[251] Polanski, A. and Stoja, E. (2010). Incorporating higher moments into value-at-risk forecasting. Journal of Forecasting, 29, 523-535.

[252] Sihem, M. and Slaheddine, H. (2014). The impact of higher order moments on market risk assessment. Procedia Economics and Finance, 13, 143-153.

Appendix A

Appendix to Chapter 5

Here, we present the football scores data what we use for the application in Chapter 5. Note that X = the number of goals scored by ACF Fiorentina and Y = the number of goals scored by Juventus.

X	Y
1	2
0	0
1	1
1	2
1	1
0	1
1	1
3	2
1	1
1	1
1	2
3	3
0	1
1	2
1	1
1	3
3	3
0	1
1	1
1	2
1	0
3	0
1	2
1	1
0	1
0	1

Table A.1: The football data.

Appendix B

Appendix to Chapter 6

Here, we illustrate the derivations of the given expressions in Section 6.3 for the bivariate skew normal distribution (Balakrishnan and Lai, 2009), the bivariate t distribution (Kotz and Nadarajah, 2004), the bivariate skew t distribution (Azzalimi and Capitanio, 2003), the standard symmetric bivariate Laplace distribution (Kotz et al, 2001, equation (5.1.2)), the asymmetric bivariate Laplace distribution (Kotz et al. 2001), the bivariate hyperbolic distribution (Balakrishnan and Lai, 2009), and the conditionaly specified bivariate Gumbel distribution (Balakrishnan and Lai, 2009).

Derivations for the Bivariate skew normal distribution (Balakrishnan and Lai,2009) It has the joint pdf specified by

$$f(x,y) = C\exp\left(-\frac{x^2+y^2}{2a^2}\right)\Phi\left(\alpha x\right)\Phi\left(\beta y\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $-\infty < \alpha < +\infty$ and $-\infty < \beta < +\infty$, where C denotes the normalizing constant.

Based on the definition, $x = r\cdot\sin\theta$, $y = r\cdot\cos\theta$ and $f(x,y) = r\cdot f(r,\theta)$. We have

$$f(r,\theta) = Cr\exp\left(-\frac{r^2}{2a^2}\right)\Phi\left(\alpha r\sin\theta\right)\Phi\left(\beta r\cos\theta\right)$$

Note that $\Phi(x) = \frac{1}{2} + \frac{1}{2}\text{erf}(\frac{x}{\sqrt{2}}) = \frac{1}{2} + \frac{1}{2}\frac{2}{\sqrt{\pi}}\sum_{k=0}^{\infty}\frac{(-1)^k(x/\sqrt{2})^{2k+1}}{k!(2k+1)}$.

$$\begin{aligned}
f(r,\theta) &= Cr\exp\left(-\frac{r^2}{2a^2}\right)\Big[\frac{1}{2}+\frac{1}{\sqrt{\pi}}\sum_{k=0}^{\infty}\frac{(-1)^k 2^{-k-\frac{1}{2}}(\alpha r\sin\theta)^{2k+1}}{k!(2k+1)}\ \Big]\cdot\Big[\frac{1}{2}+\frac{1}{\sqrt{\pi}}\sum_{l=0}^{\infty}\frac{(-1)^l 2^{-l-\frac{1}{2}}(\beta r\cos\theta)^{2l+1}}{l!(2l+1)}\ \Big] \\
&= Cr\exp\left(-\frac{r^2}{2a^2}\right)\Big[\frac{1}{4}+\frac{1}{2\sqrt{\pi}}\sum_{k=0}^{\infty}\frac{(-1)^k 2^{-k-\frac{1}{2}}(\alpha r)^{2k+1}\sin^{2k+1}(\theta)}{k!(2k+1)}+\frac{1}{2\sqrt{\pi}}\sum_{l=0}^{\infty}\frac{(-1)^l 2^{-l-\frac{1}{2}}(\beta r)^{2l+1}\cos^{2l+1}(\theta)}{l!(2l+1)} \\
&\quad +\frac{1}{\pi}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^{k+l}2^{-k-l-1}r^{2k+2l+2}\alpha^{2k+1}\beta^{2l+1}}{k!l!(2k+1)(2l+1)}\sin^{2k+1}(\theta)\cos^{2l+1}(\theta)\Big].
\end{aligned} \tag{B.2}$$

Because both $2k+1$ and $2l+1$ are even for $k,l \in \mathbb{Z}_+, \mathbb{Z}_+ = \{0,1,2,3,...\}$, we have $\int_0^{2\pi}\sin^{2k+1}(\theta)d\theta = 0$, $\int_0^{2\pi}\cos^{2l+1}(\theta)d\theta = 0$, and $\int_0^{2\pi}\sin^{2k+1}(\theta)\cos^{2l+1}(\theta)d\theta = 0$. Therefore,

$$\begin{aligned}
f_R(r) &= \int_0^{2\pi} f(r,\theta)d\theta \\
&= \frac{1}{4}\cdot C\cdot r\cdot\exp(-\frac{r^2}{2a^2})\theta|_0^{2\pi} \\
&= \frac{\pi Cr}{2}\exp(-\frac{r^2}{2a^2}).
\end{aligned}$$

The marginal density distribution of Θ can be expressed as

$$\begin{aligned}
f_\Theta(\theta) &= \frac{1}{4}C\int_0^{\infty} r\exp(-\frac{r^2}{2a^2})dr + \frac{1}{4}C\int_0^{\infty} r\exp(-\frac{r^2}{2a^2})\text{erf}(\frac{\alpha r\sin\theta}{\sqrt{2}})dr \\
&\quad +\frac{1}{4}C\int_0^{\infty} r\exp(-\frac{r^2}{2a^2})\text{erf}(\frac{\beta r\cos\theta}{\sqrt{2}})dr + \frac{1}{4}C\int_0^{\infty} r\exp(-\frac{r^2}{2a^2})\text{erf}(\frac{\alpha r\sin\theta}{\sqrt{2}})\text{erf}(\frac{\beta r\cos\theta}{\sqrt{2}})dr.
\end{aligned}$$

We can devide it into four parts to derive the explict forms.

$$\begin{aligned}
(1)\ \ \frac{1}{4}C\int_0^{\infty} r\exp(-\frac{r^2}{2a^2})dr &= \frac{1}{4}C\frac{1}{2}(\frac{1}{2a^2})^{-2/2}\Gamma(\frac{2}{2}) \\
&= \frac{a^2C}{4},
\end{aligned}$$

as $\int_0^{\infty} x^{\alpha-1}\exp(-px^{\mu})dx = \frac{1}{\mu}p^{-\alpha/\mu}\Gamma(\frac{\alpha}{\mu})$.

$$
\begin{aligned}
(2)\ \ \frac{1}{4}C\int_0^\infty r\exp(-\frac{r^2}{2a^2})\mathrm{erf}(\frac{\alpha r\sin\theta}{\sqrt{2}})dr &= \frac{1}{4}Ca^2\frac{\alpha\sin\theta/\sqrt{2}}{\sqrt{\frac{\alpha^2\sin^2\theta}{2}+\frac{1}{2a^2}}} \\
&= \frac{a^2C}{4}\frac{\alpha\sin\theta}{\sqrt{2}\frac{1}{\sqrt{2}a}\sqrt{1+a^2\alpha^2\sin^2\theta}} \\
&= \frac{Ca^3\alpha\sin\theta}{4\sqrt{a^2\alpha^2\sin^2\theta+1}},
\end{aligned}
$$

as $\int_0^\infty x^n e^{-px^2}\mathrm{erf}(cx)dx = \frac{1}{2p}\frac{c}{\sqrt{c^2+p}}$, where $n=1, p=\frac{1}{2a^2}, c=\frac{\alpha\sin\theta}{\sqrt{2}}$.

$$
\begin{aligned}
(3)\ \ \frac{1}{4}C\int_0^\infty r\exp(-\frac{r^2}{2a^2})\mathrm{erf}(\frac{\beta r\cos\theta}{\sqrt{2}})dr &= \frac{1}{4}Ca^2\frac{\beta\cos\theta/\sqrt{2}}{\sqrt{\frac{\beta^2\cos^2\theta}{2}+\frac{1}{2a^2}}} \\
&= \frac{Ca^3\beta\cos\theta}{4\sqrt{a^2\beta^2\cos^2\theta+1}},
\end{aligned}
$$

as $\int_0^\infty x^n e^{-px^2}\mathrm{erf}(cx)dx = \frac{1}{2p}\frac{c}{\sqrt{c^2+p}}$, where $n=1, p=\frac{1}{2a^2}, c=\frac{\beta\cos\theta}{\sqrt{2}}$.

$$
\begin{aligned}
(4)\ \ :\ \ & \frac{1}{4}C\int_0^\infty r\exp(-\frac{r^2}{2a^2})\mathrm{erf}(\frac{\alpha\sin\theta}{\sqrt{2}}r)\mathrm{erf}(\frac{\beta\cos\theta}{\sqrt{2}}r)dr \\
= & \frac{C2a^2}{4\pi}\Big(\frac{\alpha\sin\theta/\sqrt{2}}{\sqrt{\frac{\alpha^2\sin^2\theta}{2}+\frac{1}{2a^2}}}\cdot\arctan\frac{\beta\cos\theta/\sqrt{2}}{\sqrt{\frac{\alpha^2\sin^2\theta}{2}+\frac{1}{2a^2}}}+\frac{\beta\cos\theta/\sqrt{2}}{\sqrt{\frac{\beta^2\cos^2\theta}{2}+\frac{1}{2a^2}}}\cdot\arctan\frac{\alpha\sin\theta/\sqrt{2}}{\sqrt{\frac{\beta^2\cos^2\theta}{2}+\frac{1}{2a^2}}}\Big) \\
= & \frac{Ca^2}{2\pi}\Big(\frac{a\alpha\sin\theta}{\sqrt{a^2\alpha^2\sin^2\theta+1}}\cdot\arctan\frac{a\beta\cos\theta}{\sqrt{a^2\alpha^2\sin^2\theta+1}}+\frac{a\beta\cos\theta}{\sqrt{a^2\beta^2\cos^2\theta+1}}\arctan\frac{a\alpha\sin\theta}{\sqrt{a^2\beta^2\cos^2\theta+1}},
\end{aligned}
$$

as $\int_0^\infty x^n e^{-px^2}\mathrm{erf}(bx)\mathrm{erf}(cx)dx = \frac{1}{\pi p}(\frac{b}{\sqrt{b^2+p}}\arctan\frac{c}{\sqrt{b^2+p}}+\frac{c}{\sqrt{c^2+p}}\arctan\frac{b}{\sqrt{c^2+p}})$, where $n=1, p=\frac{1}{2a^2}, b=\frac{\alpha\sin\theta}{\sqrt{2}}, c=\frac{\beta\cos\theta}{\sqrt{2}}$.

The explicit form of $f_\Theta(\theta)$ is,

$$
\begin{aligned}
f_\Theta(\theta) &= \frac{a^2C}{4}+\frac{Ca^3\alpha\sin\theta}{4\sqrt{a^2\alpha^2\sin^2\theta+1}}+\frac{Ca^3\beta\cos\theta}{4\sqrt{a^2\beta^2\cos^2\theta+1}} \\
&\quad+\frac{Ca^3\alpha\sin\theta}{2\pi\sqrt{\alpha^2a^2\sin^2\theta+1}}\arctan\frac{a\beta\cos\theta}{\sqrt{\alpha^2a^2\sin^2\theta+1}} \\
&\quad+\frac{Ca^3\beta\cos\theta}{2\pi\sqrt{a^2\beta^2\cos^2\theta+1}}\arctan\frac{\alpha a\sin\theta}{\sqrt{a^2\beta^2\cos^2\theta+1}}
\end{aligned}
$$

The the $p-th$ moment of R can be derived as

$$\begin{aligned}
E[R^p] &= \int_0^\infty r^p \frac{\pi Cr}{2} \exp\left(-\frac{r^2}{2a^2}\right) dr \\
&= \frac{\pi C}{2} \int_0^\infty r^{p+1} \exp(-\frac{r^2}{2a^2}) dr \\
&= \frac{\pi C}{2} \frac{1}{2} (\frac{1}{2a^2})^{-p/2-1} \Gamma(\frac{p}{2}+1) \\
&= \pi C a^{p+2} 2^{p/2-1} \Gamma(\frac{p}{2}+1)
\end{aligned}$$

Derivations for the bivariate t distribution (Kotz and Nadarajah, 2004) It has the joint pdf specified by

$$f(x,y) = C\left(1 + ax^2 + by^2 + 2cxy\right)^{-\frac{\nu+2}{2}}$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $a > 0$, $b > 0$, $-\infty < c < +\infty$ and $\nu > 0$, where C denotes the normalizing constant. Based on the definition, $x = r \cdot \sin\theta$, $y = r \cdot \cos\theta$ and $f(x,y) = r \cdot f(r,\theta)$. We have

$$\begin{aligned}
f(r,\theta) &= Cr(1 + ar^2 \sin^2\theta + br^2 \cos^2\theta + cr^2 \sin 2\theta)^{-\frac{\nu+2}{2}} \\
&= Cr(1 + ar^2 + (b-a)r^2 \cos^2\theta + cr^2 \sin(2\theta))^{-\frac{\nu+2}{2}} \\
&= Cr(1 + ar^2 + \frac{b-a}{2}(1+\cos 2\theta)r^2 + cr^2 \sin(2\theta))^{-\frac{\nu+2}{2}} \\
&= Cr\left(1 + \frac{a+b}{2}r^2 + \left[\frac{b-a}{2}\cos(2\theta) + c\sin(2\theta)\right]r^2\right)^{-\frac{\nu+2}{2}}.
\end{aligned}$$

The marginal density distribution of R is

$$\begin{aligned}
f_R(r) &= \int_0^{2\pi} f(r,\theta)d\theta \\
&= Cr\int_0^{2\pi} \left[(1 + \frac{a+b}{2}r^2) + (\frac{b-a}{2}\cos 2\theta + c\sin 2\theta)r^2\right]^{-\frac{\nu+2}{2}} d\theta \\
&= Cr\int_0^{2\pi} \sum_{k=0}^{\infty} \binom{-\frac{v+2}{2}}{k} (\frac{b-a}{2}\cos 2\theta + c\sin 2\theta)^k r^{2k} [1 + \frac{a+b}{2}r^2]^{-\frac{\nu+2}{2}-k} d\theta \\
&= C\sum_{k=0}^{\infty}\sum_{l=0}^{k} \binom{-\frac{\nu+2}{2}}{k}\binom{k}{l} c^l (\frac{b-a}{2})^{k-l} \int_0^{2\pi} \sin^l(2\theta)\cos^{k-l}(2\theta)d\theta \; r^{2k+1}[1 + \frac{a+b}{2}r^2]^{-\frac{\nu+2}{2}-k},
\end{aligned}$$

where

$$\begin{aligned}\int_0^{2\pi} \sin^l(2\theta)\cos^{k-l}(2\theta)d\theta &= \frac{1}{2}\int_0^{4\pi} \sin^l(x)\cos^{k-l}(x)dx \\ &= \int_0^{2\pi} \sin^l(x)\cos^{k-l}(x)dx \\ &= \frac{[1+(-1)^l][1+(-1)^{k-l}]}{2} B(\frac{l+1}{2}, \frac{k-l+1}{2}),\end{aligned}$$

as $\int_\alpha^{\alpha+2\pi} \sin^m x \cos^n x dx = \frac{[1+(-1)^m][1+(-1)^n]}{2} B(\frac{m+1}{2}, \frac{n+1}{2})$. Therefore,

$$\begin{aligned}f_R(r) &= \frac{C}{2}\sum_{k=0}^{\infty}\sum_{\ell=0}^{k} \binom{-\frac{\nu+2}{2}}{k}\binom{k}{\ell} c^\ell \left(\frac{b-a}{2}\right)^{k-\ell} \left[1+(-1)^k+(-1)^\ell+(-1)^{k-\ell}\right] \\ &\cdot \quad B\left(\frac{\ell+1}{2}, \frac{k-\ell+1}{2}\right) r^{2k+1}\left[1+\frac{a+b}{2}r^2\right]^{-\frac{\nu+2}{2}-k}.\end{aligned}$$

Let $f(r,\theta) = Cr(1+Dr^2)^{-\frac{\nu+2}{2}}$, where $D = \frac{a+b}{2} + \frac{b-a}{2}\cos(2\theta) + c\sin(2\theta)$.

$$\begin{aligned}f_\Theta(\theta) &= \int_0^\infty f(r,\theta)dr \\ &= C\int_0^\infty r(1+Dr^2)^{-\frac{\nu+2}{2}}dr\end{aligned}$$

Let $Dr^2 = z$, $r = \sqrt{\frac{z}{D}}$ and $dr = \frac{1}{2\sqrt{D}}z^{-\frac{1}{2}}dz$.

$$\begin{aligned}f_\Theta(\theta) &= C\int_0^\infty \sqrt{\frac{z}{D}}(1+z)^{-\frac{\nu+2}{2}}\frac{1}{2\sqrt{D}}z^{-1/2}dz \\ &= \frac{C}{2D}\int_0^\infty (1+z)^{-\frac{\nu+2}{2}}dz \\ &= \frac{C}{2D}\cdot B(1,\frac{\nu}{2}) \\ &= \frac{C}{a+b+(b-a)\cos(2\theta)+2c\sin(2\theta)}\frac{\Gamma(\frac{\nu}{2}}{\Gamma(\frac{\nu}{2}+1)} \\ &= \frac{2C}{\nu[a+b+(b-a)\cos(2\theta)+2c\sin(2\theta)]}\end{aligned}$$

as $\int_0^\infty \frac{x^{\alpha-1}}{(x+\mu)^p}dx = \mu^{\alpha-p}B(\alpha, p-\alpha)$, where $\alpha = 1, \mu = 1, p = \frac{\nu+2}{2}$, and $B(a,b)$ is the Beta function. The pth moment for R is

$$\begin{aligned}E[R^p] &= \frac{C}{2}\sum_{k=0}^{\infty}\sum_{l=0}^{k}\binom{-\frac{\nu+2}{2}}{k}\binom{k}{l}c^l(\frac{b-a}{2})^{k-l}[1+(-1)^k+(-1)^l+(-1)^{k-l}]B(\frac{l+1}{2},\frac{k-l+1}{2}) \\ &\cdot \quad \int_0^\infty r^{2k+p+1}[1+\frac{a+b}{2}r^2]^{-\frac{v+2}{2}-k}dr,\end{aligned}$$

where, $\int_0^\infty r^{2k+p+1}[1+\frac{a+b}{2}r^2]^{-\frac{v+2}{2}-k}dr = \int_0^\infty \left(\frac{2}{a+b}\right)^{k+\frac{p}{2}+\frac{1}{2}} z^{k+\frac{p}{2}+\frac{1}{2}}(1+z)^{-\frac{\nu+2}{2}-k}\frac{1}{\sqrt{2(a+b)}}z^{-\frac{1}{2}}dz$ when $z=\frac{a+b}{2}r^2$, $r=\sqrt{\frac{2z}{a+b}}$ and $dr=\frac{1}{\sqrt{2(a+b)}}z^{-\frac{1}{2}}dz$.

By applying the equation $\int_0^\infty \frac{x^{\alpha-1}}{(x+\mu)^p}dx = \mu^{\alpha-p}B(\alpha,p-\alpha)$, the explicit form of the pth moment of R is

$$
\begin{aligned}
E\left(R^p\right) &= \frac{2^{\frac{p}{2}-1}C}{(b+a)^{\frac{p}{2}+1}}\sum_{k=0}^{\infty}\sum_{\ell=0}^{k}\binom{-\frac{\nu+2}{2}}{k}\binom{k}{\ell}(2c)^{\ell}\frac{(b-a)^{k-\ell}}{(b+a)^{k}}\left[1+(-1)^k+(-1)^{\ell}+(-1)^{k-\ell}\right] \\
&\cdot \quad B\left(\frac{\ell+1}{2},\frac{k-\ell+1}{2}\right)B\left(k+\frac{p}{2}+1,\frac{\nu-p}{2}\right)
\end{aligned}
$$

for $r>0$, $0\le\theta\le 2\pi$ and $p<\nu$.

Derivations for the bivariate skew t distribution (Azzalini and Capitanio, 2003)

It has the joint pdf specified by

$$
f(x,y)=\frac{1}{\pi}\left(1+\frac{x^2+y^2}{\nu}\right)^{-\frac{\nu+2}{2}} T_{\nu+2}\left((ax+by)\sqrt{\frac{\nu+2}{x^2+y^2+\nu}}\right)
$$

for $-\infty<x<+\infty$, $-\infty<y<+\infty$, $-\infty<a<+\infty$, $-\infty<b<+\infty$ and $\nu>0$, where $T_a(\cdot)$ denotes the cdf of a standard Student's t random variable with degree of freedom a. By definition,

$$
T_{\nu+2}=\frac{1}{2}+x\cdot\frac{\Gamma(\frac{1}{2}(\nu+3))}{\sqrt{\pi(\nu+2)}\Gamma(\frac{1}{2}(\nu+2))}{}_2F_1(\frac{1}{2},\frac{1}{2}(\nu+3);\frac{3}{2};-\frac{x^2}{\nu+2}).
$$

Note that

$$
{}_2F_1(\frac{1}{2},\frac{1}{2}(\nu+3);\frac{3}{2};-\frac{x^2}{\nu+2})=\sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{\nu+3}{2})_n(-1)^n}{(\frac{3}{2})_n n!(\nu+2)^n}x^{2n}.
$$

So,

$$
T_{\nu+2}(x)=\frac{1}{2}+\frac{\Gamma(\frac{1}{2}(\nu+3))}{\sqrt{\pi}\Gamma(\frac{1}{2}(\nu+2))}\sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{\nu+3}{2})_n(-1)^n}{(\frac{3}{2})_n n!(\nu+2)^{n+\frac{1}{2}}}x^{2n+1}.
$$

Here in our case, the $x^{2n+1}=r^{2n+1}(a\sin\theta+b\cos\theta)^{2n+1}(\nu+2)^{n+\frac{1}{2}}(r^2+\nu)^{-n-\frac{1}{2}}$.

The marginal density distribution of R can be derived as

$$
\begin{aligned}
f_R(r) &= \frac{r}{\pi}(\frac{\nu+r^2}{\nu})^{\frac{\nu}{2}-1}\Big[\pi+\frac{\Gamma(\frac{1}{2}(\nu+3))}{\sqrt{\pi}\Gamma(\frac{1}{2}(\nu+2))} \\
&\cdot \sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{\nu+3}{2})_n(-1)^n}{(\frac{3}{2})_n n!(r^2+\nu)^{n+\frac{1}{2}}}r^{2n+1}\int_0^{2\pi}(a\sin\theta+b\cos\theta)^{2n+1}\Big]d\theta
\end{aligned}
\tag{B.3}
$$

Note that $\int_0^{2\pi}(a\cos x+b\sin x)^m dx=\frac{1+(-1)^m}{2}\frac{(n-1)!}{(n)!}2\pi(a^2+b^2)^{m/2}$. Here, $m=2n+1$, which is always a odd number when $n\in\mathbb{Z}_+$. So, $\frac{1+(-1)^{2n+1}}{2}=0$. Therefore,

$$
f_R(r) = r(1+\frac{r^2}{\nu})^{-\frac{\nu}{2}-1}.
$$

The marginal density distribution of Θ can be derived as

$$
\begin{aligned}
f_\Theta(\theta) &= \int_0^\infty f(r,\theta)dr \\
&= \int_0^\infty \frac{r}{\pi}(1+\frac{r^2}{\nu})^{-\frac{\nu+2}{2}}\Big[\frac{1}{2} \\
&+ \frac{\Gamma(\frac{1}{2}(\nu+3))}{\sqrt{\pi}\Gamma(\frac{1}{2}(\nu+2))}\sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{\nu+3}{2})_n(-1)^n(a\sin\theta+b\cos\theta)^{2n+1}}{(\frac{3}{2})_n n!}(\frac{r^2}{r^2+\nu})^{n+\frac{1}{2}}\Big]dr \\
&= \frac{1}{2\pi}\int_0^\infty\frac{rdr}{(1+\frac{r^2}{\nu})^{\frac{\nu+2}{2}}} \\
&+ \frac{\Gamma(\frac{1}{2}(\nu+3))}{\pi^{3/2}\Gamma(\frac{1}{2}(\nu+2))}\sum_{n=0}^{\infty}\frac{(\frac{1}{2})_n(\frac{\nu+3}{2})_n(-1)^n(a\sin\theta+b\cos\theta)^{2n+1}}{(\frac{3}{2})_n\cdot n!}\int_0^\infty r(1+\frac{r^2}{\nu})^{-\frac{\nu+2}{2}}(\frac{r^2}{r^2+\nu})^{n+\frac{1}{2}}dr
\end{aligned}
$$

$$
\begin{aligned}
(1) \quad : \quad & \int_0^\infty \frac{rdr}{(1+\frac{r^2}{\nu})^{\frac{\nu+2}{2}}} \\
= \quad & \frac{\nu}{2}\int_0^\infty \frac{1}{(1+\frac{r^2}{\nu})^{\frac{\nu}{2}+1}} d\frac{r^2}{\nu} \\
= \quad & \frac{\nu}{2}\int_0^\infty (1+x)^{-\frac{\nu}{2}-1}dx \\
= \quad & \frac{\nu}{2}\cdot B(1,\frac{\nu}{2}) \\
= \quad & \frac{\nu}{2}\cdot\frac{2}{\nu} \\
= \quad & 1,
\end{aligned}
$$

as $\int_0^\infty (x^\mu + z^\mu)^{-p}dx = \mu^{-1}z^{1-\mu p}B(1/\mu, p-1/\mu)$.

$$
\begin{aligned}
(2) \quad : \quad & \int_0^\infty r(1+\frac{r^2}{\nu})^{-\frac{\nu}{2}-1}(\frac{r^2}{r^2+\nu})^{n+\frac{1}{2}}dr \\
= \quad & \int_0^\infty r(\nu+r^2)^{-\frac{\nu}{2}-1}\nu^{\frac{\nu}{2}+1}r^{2n+1}(r^2+\nu)^{-n-\frac{1}{2}}dr \\
= \quad & \nu^{\nu/2+1}\int_0^\infty r^{2n+2}(r^2+\nu)^{-(\frac{\nu}{2}+n+\frac{3}{2})}dr \\
= \quad & \frac{\nu^{\nu/2+1}}{2}\int_0^\infty x^{n+\frac{1}{2}}(x+\nu)^{-(\frac{\nu}{2}+n+\frac{3}{2})}dx \\
= \quad & \frac{\nu^{\frac{\nu}{2}+1}}{2}\cdot\nu^{-\frac{\nu}{2}}\cdot B(n+\frac{3}{2},\frac{\nu}{2}) \\
= \quad & \frac{\nu}{2}B(n+\frac{3}{2},\frac{\nu}{2}),
\end{aligned}
$$

as $\int_0^\infty \frac{x^{\alpha-1}}{(x+\nu)^p}dx = \nu^{\alpha-p}B(\alpha, p-\alpha)$.

Therefore, the marginal density distribution of Θ is

$$
f_\Theta(\theta) \quad = \quad \frac{1}{2\pi} + \frac{\nu\Gamma\left(\frac{\nu+3}{2}\right)}{2\pi^{3/2}\Gamma\left(\frac{\nu+2}{2}\right)}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{1}{2}\right)_k\left(\frac{\nu+3}{2}\right)_k}{k!\left(\frac{3}{2}\right)_k}\left(a\sin\theta+b\cos\theta\right)^{2k+1}B\left(k+\frac{3}{2},\frac{\nu}{2}\right)
$$

The pth moment of R can be derived as

$$\begin{aligned} E[R^p] &= \int_0^\infty r^{p+1}(1+\frac{r^2}{\nu})^{-\frac{\nu}{2}-1}dr \\ &= \frac{1}{2}\int_0^\infty \frac{x^{p/2}}{(1+\frac{x}{\nu})^{\nu/2+1}}dx \\ &= \frac{\nu^{\frac{\nu}{2}+1}}{2}\int_0^\infty \frac{x^{p/2}}{(\nu+x)^{\frac{\nu}{2}+1}}dx \\ &= \frac{\nu^{\frac{\nu}{2}+1}}{2}\nu^{\frac{p-\nu}{2}}B(\frac{p}{2}+1,\frac{\nu-p}{2}) \\ &= 2^{-1}\nu^{\frac{p}{2}+1}B(\frac{p}{2}+1,\frac{\nu-p}{2}). \end{aligned}$$

Derivations for the standard symmetric bivariate Laplace distribution (Kotz et al., 2001,equation (5.1.2) It has the joint pdf specified by

$$f(x,y) = \frac{1}{\pi}K_0\left(\sqrt{2\left(x^2+y^2\right)}\right)$$

for $-\infty < x < +\infty$ and $-\infty < y < +\infty$.

Based on the definition, $x = r\cdot\sin\theta$, $y = r\cdot\cos\theta$ and $f(x,y) = r\cdot f(r,\theta)$. We have

$$f(r,\theta) = \frac{1}{\pi}rK_0\left(\sqrt{2}r\right),$$

The marginal distribution of R can be derived as

$$\begin{aligned} f_R(r) &= \int_0^{2\pi}\frac{r}{\pi}K_0\left(\sqrt{2}r\right)d\theta \\ &= 2rK_0\left(\sqrt{2}r\right) \end{aligned}$$

The marginal distribution of Θ can be derived as

$$\begin{aligned} f_\Theta(\theta) &= \frac{1}{\pi}\int_0^\infty rK_0\left(\sqrt{2}r\right)dr \\ &= \frac{1}{\pi}(\sqrt{2})^{-2}\Gamma(1)\Gamma(1) \\ &= \frac{1}{2\pi}, \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}K_\nu(cx)dx = 2^{\alpha-2}c^{-\alpha}\Gamma(\frac{\alpha+\nu}{2})\Gamma(\frac{\alpha-\nu}{2})$, and $\alpha = 2, \nu = 0, c = \sqrt{2}$.

The pth moment of R can be derived as

$$\begin{aligned} E[R^p] &= \int_0^\infty r^p f_R(r)dr \\ &= 2\int_0^\infty r^{p+1}K_0\left(\sqrt{2}r\right)dr \\ &= 2\cdot 2^p(\sqrt{2})^{-p-2}\Gamma(\frac{p}{2}+1)\Gamma(\frac{p}{2}+1) \\ &= 2^{p/2}\cdot\Gamma^2\left(\frac{p}{2}+1\right) \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}K_\nu(cx)dx = 2^{\alpha-2}c^{-\alpha}\Gamma(\frac{\alpha+\nu}{2})\Gamma(\frac{\alpha-\nu}{2})$, and $\alpha = p+2, \nu = 0, c = \sqrt{2}$.

Derivations for the asymmetric bivariate Laplace distribution (Kotz et al., 2001,equation (6.5.3) It has the joint pdf specified by

$$f(x,y) = C\exp\left(\alpha x + \beta y\right)\left(x^2+y^2\right)^{\frac{\nu}{2}}K_\nu\left(\gamma\sqrt{x^2+y^2}\right)$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha < 0$, $\beta < 0$, $\gamma > 0$ and $\nu > 0$, where C denotes the normalizing constant.

Based on the definition, $x = r\cdot\sin\theta$, $y = r\cdot\cos\theta$ and $f(x,y) = r\cdot f(r,\theta)$. We have

$$f(r,\theta) = Cr^{1+\nu}\exp\left(\alpha r\sin\theta + \beta r\cos\theta\right)K_\nu\left(\gamma r\right).$$

The marginal distribution of R can be derived as

$$\begin{aligned} f_R(r) &= Cr^{1+\nu}K_\nu(\gamma\cdot r)\int_0^{2\pi}\exp\left(\alpha r\sin\theta + \beta r\cos\theta\right)d\theta \\ &= Cr^{1+\nu}K_\nu\left(\gamma\cdot r\right)2\pi I_0\left(r\sqrt{\alpha^2+\beta^2}\right) \\ &= 2\pi Cr^{1+\nu}K_\nu\left(\gamma\cdot r\right)I_0\left(r\sqrt{\alpha^2+\beta^2}\right). \end{aligned}$$

The marginal distribution of Θ can be derived as

$$\begin{aligned} f_\Theta(\theta) &= \int_0^\infty Cr^{1+\nu}\exp\left(\alpha r\sin\theta+\beta r\cos\theta\right)K_\nu\left(\gamma r\right)dr \\ &= \frac{\sqrt{\pi}C(2\gamma)^\nu\Gamma(2+2\nu)}{\Gamma\left(\frac{5}{2}+\nu\right)\left(\gamma-\alpha\sin\theta-\beta\cos\theta\right)^{2+2\nu}}{}_2F_1\left(2+2\nu,\frac{1}{2}+\nu;\frac{5}{2}+\nu;-\frac{\gamma+\alpha\sin\theta+\beta\cos\theta}{\gamma-\alpha\sin\theta-\beta\cos\theta}\right) \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}e^{-px}K_\nu(cx)dx = \frac{(2c)^\nu\sqrt{\pi}}{(p+c)\alpha+\nu}\Gamma\begin{bmatrix}\alpha-\nu,\alpha+\nu \\ \alpha+\frac{1}{2}\end{bmatrix}{}_2F_1\left(\alpha+\nu,\nu+\frac{1}{2};\alpha+\frac{1}{2};\frac{p-c}{p+c}\right)$, and $\alpha = 2+\nu, p=-(\alpha\sin\theta+\beta\cos\theta), \nu=\nu, c=\gamma$.

The pth moment of R can be derived as

$$\begin{aligned} E[R^p] &= \int_0^\infty r^p\cdot 2\pi Cr^{1+\nu}K_\nu\left(\gamma\cdot r\right)I_0\left(r\sqrt{\alpha^2+\beta^2}\right)dr \\ &= 2\pi C\int_0^\infty r^{1+\nu+p}I_0\left(r\sqrt{\alpha^2+\beta^2}\right)K_\nu\left(\gamma\cdot r\right)dr \\ &= 2^{\nu+p+1}\pi C\gamma^{-2-\nu-p}\Gamma\left(1+\nu+\frac{p}{2}\right)\Gamma\left(1+\frac{p}{2}\right){}_2F_1\left(1+\nu+\frac{p}{2},1+\frac{p}{2};1;\frac{\alpha^2+\beta^2}{\gamma^2}\right), \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}I_\mu(bx)K_\nu(cx)dx = 2^{\alpha-2}b^\mu c^{-\alpha-\mu}\Gamma\begin{bmatrix}(\alpha+\mu+\nu)/2,(\alpha+\mu-\nu)/2 \\ \mu+1\end{bmatrix}\cdot{}_2F_1\left(\frac{\alpha+\mu+\nu}{2},\frac{\alpha+\mu-\nu}{2};\mu+1;\frac{b^2}{c^2}\right)$, and $\alpha=p+\nu+2, \mu=0, b=\sqrt{\alpha^2+\beta^2}, \nu=\nu, c=\gamma$.

Derivations for the bivariate hyperbolic distribution (Balakrishnan and Lai, 2009) It has the joint pdf specified by

$$f(x,y) = C\exp\left[-\alpha\left(x^2+y^2\right)-\beta x-\gamma y\right]$$

for $-\infty < x < +\infty$, $-\infty < y < +\infty$, $\alpha > 0$, $\beta > 0$ and $\gamma > 0$, where C denotes the normalizing constant.

Based on the definition, $x = r\cdot\sin\theta$, $y = r\cdot\cos\theta$ and $f(x,y) = r\cdot f(r,\theta)$. We have

$$f(r,\theta) = Cr\exp\left(-\alpha r^2-\beta r\sin\theta-\gamma r\cos\theta\right).$$

The marginal distribution of R can be derived as

$$\begin{aligned} f_R(r) &= Cr\exp\left(-\alpha r^2\right)\int_0^{2\pi}\exp\Big[(-\beta r)\sin\theta+(-\gamma r)\cos\theta\Big]d\theta \\ &= 2\pi Cr\exp\left(-\alpha r^2\right)I_0\left(r\sqrt{\beta^2+\gamma^2}\right) \end{aligned}$$

The marginal distribution of Θ can be derived as

$$\begin{aligned} f_\Theta(\theta) &= C\int_0^\infty r\exp\left[-\alpha r^2-(\beta\sin\theta+\gamma\cos\theta)\,r\right]dr \\ &= C\Gamma(2)\Gamma^{-1}(2\alpha)\exp\left(\frac{(\beta\sin\theta+\gamma\cos\theta)^2}{8\alpha}\right)D_{-2}\left(\frac{\beta\sin\theta+\gamma\cos\theta}{\sqrt{2\alpha}}\right) \\ &= \frac{C}{2\alpha}\exp\left[\frac{(\beta\sin\theta+\gamma\cos\theta)^2}{8\alpha}\right]D_{-2}\left(\frac{\beta\sin\theta+\gamma\cos\theta}{\sqrt{2\alpha}}\right), \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}e^{-px^2-qx}dx=\Gamma(\alpha)(2p)^{-\alpha/2}\exp\left[\frac{q^2}{8p}\right]D_{-\alpha}\left(\frac{q}{\sqrt{2p}}\right)$, and $\alpha=2, p=\alpha, q=\beta\sin\theta+\gamma\cos\theta$.

The pth moment of R can be derived as

$$\begin{aligned} E[R^p] &= \int_0^\infty r^p 2\pi Cr\exp\left(-\alpha r^2\right)I_0\left(r\sqrt{\beta^2+\gamma^2}\right)dr \\ &= 2\pi C\int_0^\infty r^{p+1}\exp\left(-\alpha r^2\right)I_0\left(r\sqrt{\beta^2+\gamma^2}\right)dr \\ &= 2\pi C2^{-1}\alpha^{-\frac{p}{2}-1}\Gamma\left(\frac{p}{2}+1\right){}_1F_1\left(\frac{p}{2}+1;1;\frac{\beta^2+\gamma^2}{4\alpha}\right) \\ &= \pi C\alpha^{-\frac{p}{2}-1}\Gamma\left(\frac{p}{2}+1\right){}_1F_1\left(\frac{p}{2}+1;1;\frac{\beta^2+\gamma^2}{4\alpha}\right), \end{aligned}$$

as $\int_0^\infty x^{\alpha-1}e^{-px^2}I_\nu(cx)dx=2^{-\nu-1}c^\nu p^{-(\alpha+\nu)/2}\Gamma\begin{bmatrix}(\alpha+\nu)/2\\ \nu+1\end{bmatrix}{}_1F_1\left(\frac{\alpha+\nu}{2};\nu+1;\frac{c^2}{4p}\right)$, and $\alpha=p+2, p=\alpha, \nu=0, c=\sqrt{\beta^2+\gamma^2}$.

Derivations for the conditionally specified bivariate Gumbel distribution (Balakrishnan and Lai, 2009, Section 12.13.1) It has the joint pdf specified by

$$f(x,y) = C\exp\left[-x-y-\exp(-x)-\exp(-y)-\alpha\exp(-x-y)\right]$$

for $-\infty<x<+\infty$, $-\infty<y<+\infty$ and $0<\alpha<1$, where C denotes the normalizing constant.

Based on the definition, $x = r \cdot \sin\theta$, $y = r \cdot \cos\theta$ and $f(x, y) = r \cdot f(r, \theta)$. We have

$$\begin{aligned} f(r,\theta) &= Cr\exp\left[-r\sin\theta - r\cos\theta - \exp(-r\sin\theta) - \exp(-r\cos\theta) - \alpha\exp(-r\sin\theta - r\cos\theta)\right] \\ &= Cr\exp\left[-r\left(\sin\theta+\cos\theta\right)\right]\exp\left[-\exp\left(-r\sin\theta\right)\right]\exp\left[-\exp\left(-r\cos\theta\right)\right] \\ &\quad \cdot\exp\left[-\alpha\exp\left(-r\sin\theta - r\cos\theta\right)\right] \\ &= Cr\exp\left[-r\left(\sin\theta+\cos\theta\right)\right]\sum_{i=0}^{\infty}\frac{(-1)^i\exp\left(-ir\sin\theta\right)}{i!}\sum_{j=0}^{\infty}\frac{(-1)^j\exp\left(-jr\cos\theta\right)}{j!} \\ &\quad \cdot\sum_{k=0}^{\infty}\frac{(-1)^k\alpha^k\exp\left(-kr(\sin\theta+\cos\theta)\right)}{k!} \\ &= Cr\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\alpha^k}{i!j!k!}\exp\Big[(-r)\cdot(1+i+k)\sin\theta + (-r)\cdot(1+j+k)\cos\theta\Big] \end{aligned}$$

The marginal distribution of R can be derived as

$$f_R(r) = 2\pi Cr\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\alpha^k}{i!j!k!}I_0\left(r\sqrt{(1+i+k)^2+(1+j+k)^2}\right)$$

The marginal distribution of Θ can be derived as

$$\begin{aligned} f_\Theta(\theta) &= C\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\alpha^k}{i!j!k!}\int_0^\infty r\exp\Big[-((1+i+k)\sin\theta+(1+j+k)\cos\theta)r\Big]dr \\ &= C\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^{i+j+k}\alpha^k}{i!j!k!}\left[(1+i+k)\sin\theta+(1+j+k)\cos\theta\right]^{-2}, \end{aligned}$$

as $\int_0^\infty r^{\alpha-1}e^{-pr}dr = p^{-\alpha}\Gamma(\alpha)$, and $\alpha = 2, p = (1+i+k)\sin\theta + (1+j+k)\cos\theta$.

Appendix C

Appendix to Chapter 7

Here, we illstrate the derivation for the characteristic function of the product of two random variables. One follows the standard normal distribution, and the other follows one of the following twelve distributions. They are the Pareto distribution, the Triangular distribution, the Argus distribution, the Student's t distribution, the Rice distribution, the Laplace distribution, the Exponential distribution, the Normal Inverse Gamma distribution, the Uniform distribution, the Rayleigh distribution, the q-Gaussian distribution, and the Normal Exponential Gamma distribution.

Derivation for the Pareto distribution (Pareto, 1964)

If y is a Pareto random variable with $y \geq k > 0$ and $\alpha > 0$, the density function of y is

$$f_Y(y) = \alpha K^{\alpha} y^{-\alpha-1}.$$

The characteristic function of the product can be derived as

$$\begin{aligned}
\phi_w(t) &= \alpha K^{\alpha} \int_K^{\infty} y^{-\alpha-1} \exp\left(-\frac{t^2}{2}y^2\right) dy \\
&= \alpha K^{\alpha} \frac{1}{2(\frac{t^2}{2})^{-\alpha/2}} \Gamma\left(-\frac{\alpha}{2}, \frac{t^2}{2}K^2\right) \\
&= \alpha K^{\alpha} 2^{-\alpha/2-1} t^{\alpha} \Gamma\left(-\frac{\alpha}{2}, \frac{t^2}{2}K^2\right),
\end{aligned}$$

as $\int_{\mu}^{\infty} x^{-m} \exp\left(-\beta x^{n}\right) dx = \frac{1}{n\beta^{z}} \Gamma\left(z, \beta\mu^{n}\right), z = \frac{1-m}{n}$, and $n = 2, \beta = \frac{t^2}{2}, \mu = K, m = \alpha + 1, z = \frac{-\alpha}{2}$.

Derivation for the Triangular distribution

The density function of y is

$$f_Y(y) = \begin{cases} \dfrac{2(y-a)}{(b-a)(c-a)}, & \text{if } a < y < c, \\ \\ \dfrac{2(b-y)}{(b-a)(b-c)}, & \text{if } c \leq y < b, \\ \\ 0, & \text{elsewhere,} \end{cases}$$

The characteristic function of the product can be derived as

$$
\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_a^c \frac{2}{(b-a)(c-a)}(y-a)\exp\left(-\frac{t^2}{2}y^2\right)dy + \int_c^b \frac{2}{(b-a)(b-c)}(b-y)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{2}{(b-a)(c-a)}\left[\int_a^c y\exp\left(-\frac{t^2}{2}y^2\right)dy - \int_a^c a\exp\left(-\frac{t^2}{2}y^2\right)dy\right] \\
&\quad +\frac{2}{(b-a)(b-c)}\left[\int_c^b b\exp\left(-\frac{t^2}{2}y^2\right)dy - \int_c^b y\exp\left(-\frac{t^2}{2}y^2\right)dy\right] \\
&= \frac{2}{(b-a)(c-a)}\left[\frac{1}{t^2}\exp\left(-\frac{a^2t^2}{2}\right) - \frac{1}{t^2}\exp\left(-\frac{c^2t^2}{2}\right)\right] \\
&\quad -\frac{2}{(b-a)(c-a)}\left[a\frac{\Gamma\left(\frac{1}{2},\frac{t^2}{2}a^2\right)}{\sqrt{2}t} - a\frac{\Gamma\left(\frac{1}{2},\frac{t^2}{2}c^2\right)}{\sqrt{2}t}\right] \\
&\quad +\frac{2}{(b-a)(b-c)}\left[b\frac{\Gamma\left(\frac{1}{2},\frac{t^2}{2}c^2\right)}{\sqrt{2}t} - b\frac{\Gamma\left(\frac{1}{2},\frac{t^2}{2}b^2\right)}{\sqrt{2}t}\right] \\
&\quad -\frac{2}{(b-a)(b-c)}\left[\frac{1}{t^2}\exp\left(-\frac{c^2t^2}{2}\right) - \frac{1}{t^2}\exp\left(-\frac{b^2t^2}{2}\right)\right] \\
&= \frac{2}{t^2(b-a)(c-a)}\left[\exp\left(-\frac{t^2a^2}{2}\right) - \exp\left(-\frac{t^2c^2}{2}\right)\right] \\
&\quad -\frac{2}{t^2(b-a)(b-c)}\left[\exp\left(-\frac{t^2c^2}{2}\right) - \exp\left(-\frac{t^2b^2}{2}\right)\right] \\
&\quad -\frac{\sqrt{2}a}{t(b-a)(c-a)}\left[\Gamma\left(\frac{1}{2},\frac{t^2a^2}{2}\right) - \Gamma\left(\frac{1}{2},\frac{t^2c^2}{2}\right)\right] \\
&\quad +\frac{\sqrt{2}b}{t(b-a)(b-c)}\left[\Gamma\left(\frac{1}{2},\frac{t^2c^2}{2}\right) - \Gamma\left(\frac{1}{2},\frac{t^2b^2}{2}\right)\right]
\end{aligned}
$$

Derivation for the Argus distribution (Albrecht, 1990)

The density function of y is

$$
f_Y(y) = \frac{a^3}{\sqrt{2\pi}\Psi(a)c^2}\exp\left[-\frac{a^2}{2}\left(1-\frac{y^2}{c^2}\right)\right]y\sqrt{1-\frac{y^2}{c^2}},
$$

for $0 < y < c$ and $a > 0$, where $\Psi(x) = \Phi(x) - x\phi(x) - \frac{1}{2}$.

The characteristic function of the product can be derived as

$$
\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \frac{a^3}{\sqrt{2\pi}\Psi(a)c^2}\exp\left[-\frac{a^2}{2}\right]\int_0^c y\left(1-\left(\frac{y}{c}\right)^2\right)^{\frac{1}{2}}\exp\left[\left(\frac{a^2}{2}-\frac{c^2t^2}{2}\right)\left(\frac{y}{c}\right)^2\right]dy \\
&= \frac{a^3}{\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]\int_0^1 \frac{y}{c}\left(1-\left(\frac{y}{c}\right)^2\right)^{\frac{1}{2}}\exp\left[\left(\frac{a^2}{2}-\frac{c^2t^2}{2}\right)\left(\frac{y}{c}\right)^2\right]d\frac{y}{c}
\end{aligned}
$$

Let $\left(\frac{y}{c}\right)^2 = x$, $x \in (0,1)$, then $\frac{y}{c} = \sqrt{x}$ and $d\frac{y}{c} = \frac{1}{2}x^{-\frac{1}{2}}dx$.

$$
\begin{aligned}
\phi_w(t) &= \frac{a^3}{\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]\int_0^1 \sqrt{x}\,(1-x)^{\frac{1}{2}}\exp\left[\left(\frac{a^2}{2}-\frac{c^2t^2}{2}\right)x\right]\frac{1}{2}x^{-\frac{1}{2}}dx \\
&= \frac{a^3}{2\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]\int_0^1 (1-x)^{\frac{1}{2}}\exp\left[\left(\frac{a^2}{2}-\frac{c^2t^2}{2}\right)x\right]dx
\end{aligned}
$$

Considering Confluent Hypergeometric function of the first kind,

$$
{}_1F_1(a;b;z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 e^{zt}t^{a-1}(1-t)^{b-a-1}dt
$$

$$
\begin{aligned}
\phi_w(t) &= \frac{a^3}{2\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]\frac{\Gamma(\frac{3}{2})\Gamma(1)}{\Gamma(\frac{5}{2})}{}_1F_1(1;\frac{5}{2};\frac{a^2-c^2t^2}{2}) \\
&= \frac{a^3}{2\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]\frac{\frac{3}{2}\sqrt{\pi}\cdot 1}{\frac{3}{4}\sqrt{\pi}}{}_1F_1(1;\frac{5}{2};\frac{a^2-c^2t^2}{2}) \\
&= \frac{a^3}{3\sqrt{2\pi}\Psi(a)}\exp\left[-\frac{a^2}{2}\right]{}_1F_1(1;\frac{5}{2};\frac{a^2-c^2t^2}{2})
\end{aligned}
$$

Derivation for the Student's t distribution

The density function of y is

$$
f_Y(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{y^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}},
$$

for $-\infty < y < +\infty$, $\nu > 0$ and $\sigma > 0$.

The characteristic function of product can be derived as,

$$\begin{aligned}
\phi_W(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\int_{-\infty}^{\infty}\left(1+\frac{y^2}{\nu\sigma}\right)^{-\frac{\nu+1}{2}}\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{2\Gamma\left(\frac{\nu+1}{2}\right)}{\sigma\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}\int_{0}^{\infty}\left(1+\frac{y^2}{\nu\sigma}\right)^{-\frac{\nu+1}{2}}\exp\left(-\frac{\nu\sigma^2t^2}{2}\cdot\frac{y^2}{\nu\sigma^2}\right)dy
\end{aligned}$$

Let $x=\frac{y^2}{\nu\sigma^2}$, $x\in(0,\infty)$, then $y=\sqrt{\nu}\sigma x^{\frac{1}{2}}$ and $dy=\frac{\sqrt{\nu}\sigma}{2}x^{-\frac{1}{2}}dx$.

$$\begin{aligned}
\phi_W(t) &= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\int_{0}^{\infty}x^{-\frac{1}{2}}(1+x)^{-\frac{\nu+1}{2}}\exp\left(-\frac{\nu\sigma^2t^2}{2}x\right)dx \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\Gamma\left(\frac{1}{2}\right)\Psi\left(\frac{1}{2},\frac{3}{2}-\frac{\nu+1}{2};\frac{\nu\sigma^2t^2}{2}\right) \\
&= \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\Psi\left(\frac{1}{2},\frac{3}{2}-\frac{\nu+1}{2};\frac{\nu\sigma^2t^2}{2}\right),
\end{aligned}$$

as $\int_0^\infty \frac{x^{\alpha-1}}{(x+z)^q}e^{-px}dx=\Gamma(\alpha)z^{\alpha-q}\Psi\left(\alpha,\alpha+1-q;pz\right)$, and $\alpha=\frac{1}{2}, z=1, q=\frac{\nu+1}{2}, p=\frac{\nu\sigma^2t^2}{2}$.

Derivation for the Rice distribution (Rice, 1945)

The density function of y is

$$f_Y(y)=\frac{y}{\sigma^2}\exp\left(-\frac{y^2+\nu^2}{2\sigma^2}\right)I_0\left(\frac{\nu y}{\sigma^2}\right),$$

for $y>0$, $\sigma>0$ and $\nu\geq 0$.

The characteristic function of product can be derived as

$$\begin{aligned}
\phi_W(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_0^\infty \frac{y}{\sigma^2}\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}+t^2\right)y^2\right)\exp\left(-\frac{\nu^2}{2\sigma^2}\right)I_0\left(\frac{\nu y}{\sigma^2}\right)dy \\
&= \sigma^{-2}\exp\left(-\frac{\nu^2}{2\sigma^2}\right)\int_0^\infty y\exp\left(-\frac{1}{2}\left(\frac{1}{\sigma^2}+t^2\right)y^2\right)I_0\left(\frac{\nu y}{\sigma^2}\right)dy \\
&= \sigma^{-2}\exp\left(-\frac{\nu^2}{2\sigma^2}\right)\cdot 2^{-1}\cdot 1\cdot\left[\frac{1}{2}\left(\frac{1}{\sigma^2}+t^2\right)\right]^{-1}\cdot\Gamma\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]{}_1F_1\left(1;1;\frac{\nu^2/\sigma^4}{2\left(\frac{1}{\sigma^2}+t^2\right)}\right) \\
&= \frac{1}{1+\sigma^2t^2}\exp\left[-\frac{\nu^2}{2\sigma^2}\right]{}_1F_1\left(1;1;\frac{\nu^2}{2\sigma^2+2\sigma^4t^2}\right),
\end{aligned}$$

as $\int_0^{\infty} x^{\alpha-1} e^{-px^2} I_\nu(cx)\, dx = A_\nu^\alpha,\ A_\nu^\alpha = 2^{-\nu-1} c^\nu p^{-(\alpha+\nu)/2} \Gamma\left[\begin{smallmatrix}(\alpha+\nu)/2\\ \nu+1\end{smallmatrix}\right] {}_1F_1\left(\frac{\alpha+\nu}{2}; \nu+1; \frac{c^2}{4p}\right)$,
and $\alpha = 2, \nu = 0, p = \frac{1}{2}\left(\frac{1}{\sigma^2} + t^2\right), c = \frac{\nu}{\sigma^2}$.

Derivation for the Laplace distribution (Laplace, 1774)

The density function of y is

$$f_Y(y) = \frac{1}{2b} \exp\left(-\frac{\mid y-\mu \mid}{b}\right),$$

for $-\infty < y < +\infty$, $-\infty < \mu < +\infty$ and $b > 0$.

The characteristic function of product can be derived as

$$\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_{-\infty}^{\mu} \frac{1}{2b}\exp\left(-\frac{\mid y-\mu \mid}{b}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy + \int_{\mu}^{\infty} \frac{1}{2b}\exp\left(-\frac{\mid y-\mu \mid}{b}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy.
\end{aligned}$$

$$\begin{aligned}
(1) \quad & \int_{-\infty}^{\mu} \frac{1}{2b}\exp\left(-\frac{\mid y-\mu \mid}{b}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
= & \frac{1}{2b}\int_{-\infty}^{\mu} \exp\left(-\frac{\mu-y}{b}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
= & \frac{1}{2b}\int_{0}^{\infty} \exp\left(-\frac{\mu-y}{b}\right)\exp\left(-\frac{t^2}{2}(\mu-y-\mu)^2\right)d(\mu-y) \\
= & \frac{1}{2b}\int_{0}^{\infty} \exp\left(-\frac{x}{b}\right)\exp\left(-\frac{t^2}{2}(x-\mu)^2\right)dx \\
= & \frac{1}{2b}\int_{0}^{\infty} \exp\left(-\frac{t^2}{2}x^2 + \mu t^2 x - \frac{x}{b} - \frac{t^2\mu^2}{2}\right)dx \\
= & \exp\left(-\frac{t^2\mu^2}{2}\right)\cdot\frac{1}{2b}\cdot\int_{0}^{\infty} x^0 \exp\left(-\frac{t^2}{2}x^2 - \left(\frac{1}{b}-\mu t^2\right)x\right)dx \\
= & \exp\left(-\frac{t^2\mu^2}{2}\right)\frac{1}{2bt}\exp\left(\frac{\frac{(1-\mu b t^2)^2}{b^2}}{4t^2}\right)D_{-1}\left(\frac{1-\mu b t^2}{bt}\right) \\
= & \frac{1}{2bt}\exp\left(\frac{1-2\mu b t^2+\mu^2 b^2 t^4 - 2\mu^2 b^2 t^4}{4b^2t^2}\right)D_{-1}\left(\frac{1-\mu b t^2}{bt}\right) \\
= & \frac{1}{2bt}\exp\left(\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(\frac{1}{b}+\frac{ut^2}{2}\right)\right)D_{-1}\left(\frac{1-\mu b t^2}{bt}\right),
\end{aligned}$$

as $\int_0^\infty x^{\nu-1}e^{-\beta x^2-\gamma x}dx = (2\beta)^{-\nu/2}\Gamma(\nu)\exp\left(\frac{\gamma^2}{8\beta}\right)D_{-\nu}\left(\frac{\gamma}{\sqrt{2\beta}}\right)$ and $\nu = 1, \beta = \frac{t^2}{2}, \gamma = \frac{1}{b} - \mu t^2$.

$$
\begin{aligned}
(2) \qquad & \int_\mu^\infty \frac{1}{2b}\exp\left(-\frac{\mid y-\mu \mid}{b}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
= & \frac{1}{2b}\int_0^\infty \exp\left(-\frac{y-\mu}{b}\right)\exp\left(-\frac{t^2}{2}\left[(y-\mu)+\mu\right]^2\right)d(y-\mu) \\
= & \frac{1}{2b}\int_0^\infty \exp\left(-\frac{t^2}{2}x^2 - \frac{t^2}{2}\mu^2 - t^2\mu x - \frac{1}{b}x\right)dx \\
= & \frac{1}{2b}\exp\left(-\frac{t^2}{2}\mu^2\right)\int_0^\infty \exp\left(-\frac{t^2}{2}x^2 - \left(t^2\mu + \frac{1}{b}\right)x\right)dx \\
= & \frac{1}{2b}\exp\left(-\frac{t^2}{2}\mu^2\right)\frac{1}{t}\exp\left(\frac{\frac{(1+\mu bt^2)^2}{b^2}}{4t^2}\right)D_{-1}\left(\frac{1+\mu bt^2}{bt}\right) \\
= & \frac{1}{2bt}\exp\left[\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(\frac{\mu t^2}{2} - \frac{1}{b}\right)\right]D_{-1}\left(\frac{1+\mu bt^2}{bt}\right),
\end{aligned}
$$

as $\int_0^\infty x^{\nu-1}e^{-\beta x^2-\gamma x}dx = (2\beta)^{-\nu/2}\Gamma(\nu)\exp\left(\frac{\gamma^2}{8\beta}\right)D_{-\nu}\left(\frac{\gamma}{\sqrt{2\beta}}\right)$ and $\nu = 1, \beta = \frac{t^2}{2}, \gamma = \frac{1}{b} + \mu t^2$. Hence, the characteristic function of the product of two random variables, which one follows the standard normal distribution and the other follows the Laplace distribution has the explicit form as

$$
\begin{aligned}
\phi_W(t) = & \frac{1}{2bt}\exp\left[\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(\frac{1}{b} + \frac{\mu t^2}{2}\right)\right]D_{-1}\left(\frac{1-\mu bt^2}{bt}\right) \\
& + \frac{1}{2bt}\exp\left[\frac{1}{4b^2t^2} - \frac{\mu}{2}\left(-\frac{1}{b} + \frac{\mu t^2}{2}\right)\right]D_{-1}\left(\frac{\mu bt^2+1}{bt}\right)
\end{aligned}
$$

Derivations for the Exponential distribution

The density function of y is

$$f_Y(y) = \lambda\exp(-\lambda y),$$

for $y > 0$ and $\lambda > 0$.

The characteristic function of the product can be derived as

$$
\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_0^\infty \lambda \exp(-\lambda y)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \lambda\int_0^\infty \exp\left(-\frac{t^2}{2}\left(y^2+\frac{2\lambda}{t^2}y\right)\right)dy \\
&= \lambda\int_0^\infty \exp\left(-\frac{t^2}{2}\left(y^2+\frac{2\lambda}{t^2}y+\frac{\lambda^2}{t^4}\right)+\frac{\lambda^2}{2t^2}\right)dy \\
&= \lambda\cdot\exp\left(\frac{\lambda^2}{2t^2}\right)\int_0^\infty \exp\left(-\left(\frac{ty}{\sqrt{2}}+\frac{\lambda}{\sqrt{2}t}\right)^2\right)dy \\
&= \frac{\lambda\sqrt{2}}{t}\exp\left(\frac{\lambda^2}{2t^2}\right)\int_{\frac{\lambda}{\sqrt{2}t}}^\infty \exp\left(-\left(\frac{ty}{\sqrt{2}}+\frac{\lambda}{\sqrt{2}t}\right)^2\right)d\left(\frac{t}{\sqrt{2}}y+\frac{\lambda}{\sqrt{2}t}\right) \\
&= \frac{\lambda\sqrt{2}}{t}\exp\left(\frac{\lambda^2}{2t^2}\right)\cdot\frac{\sqrt{\pi}}{2}\mathrm{erfc}\left(\frac{\lambda}{\sqrt{2}t}\right) \\
&= \frac{\sqrt{\pi}\lambda}{\sqrt{2}t}\exp\left(\frac{\lambda^2}{2t^2}\right)\cdot\mathrm{erfc}\left(\frac{\lambda}{\sqrt{2}t}\right).
\end{aligned}
$$

Derivation for the Normal Inverse Gamma distribution The density function of y is

$$
f_Y(y) = \frac{\sqrt{\lambda}\beta^\alpha}{\sqrt{2\pi}\sigma^{2\alpha+3}\Gamma(\alpha)}\exp\left[-\frac{2\beta+\lambda(y-\mu)^2}{2\sigma^2}\right],
$$

for $-\infty < y < +\infty$, $\lambda > 0$, $\alpha > 0$ and $\beta > 0$.

The characteristic function of the product can be derived as

$$\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \frac{\sqrt{\lambda}\beta^\alpha}{\sqrt{2\pi}\sigma^{2\alpha+3}\Gamma(\alpha)}\exp\left(-\frac{2\beta}{2\sigma^2}\right)\int_{-\infty}^{\infty}\exp\left[-\frac{\lambda(y-\mu)^2}{2\sigma^2}\right]\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{\sqrt{\lambda}\beta^\alpha}{\sqrt{2\pi}\sigma^{2\alpha+3}\Gamma(\alpha)}\exp\left(-\frac{2\beta}{2\sigma^2}\right)\int_{-\infty}^{\infty}\exp\left[-\frac{(\lambda+t^2\sigma^2)y^2-2\mu\lambda y+\lambda\mu^2}{2\sigma^2}\right]dy \\
&= \frac{\sqrt{\lambda}\beta^\alpha}{\sqrt{2\pi}\sigma^{2\alpha+3}\Gamma(\alpha)}\exp\left(-\frac{2\beta}{2\sigma^2}\right)\exp\left(-\frac{\lambda\mu^2}{2\sigma^2}\right) \\
&\quad \int_{-\infty}^{\infty}\exp\left[-\frac{1}{2\sigma^2/(\lambda+t^2\sigma^2)}\left(y^2-\frac{2\mu\lambda}{\lambda+t^2\sigma^2}y+\frac{\mu^2\lambda^2}{(\lambda+t^2\sigma^2)^2}\right)+\frac{\frac{\mu^2\lambda^2}{(\lambda+t^2\sigma^2)^2}}{2\sigma^2/(\lambda+t^2\sigma^2)}\right]dy \\
&= \frac{\sqrt{\lambda}\beta^\alpha}{\sigma^{2\alpha+2}\Gamma(\alpha)}\exp\left(-\frac{2\beta+\lambda\mu^2}{2\sigma^2}+\frac{\mu^2\lambda^2}{2\sigma^2(\lambda+t^2\sigma^2)}\right)\cdot\frac{1}{\sqrt{\sigma^2t^2+\lambda}} \\
&\quad \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma/\sqrt{\sigma^2t^2+\lambda}}\exp\left[-\frac{1}{2\sigma^2/(\lambda+t^2\sigma^2)}\left(y-\frac{\mu\lambda}{\lambda+t^2\sigma^2}\right)^2\right]dy \\
&= \frac{\sqrt{\lambda}\beta^\alpha}{\sigma^{2\alpha+2}\Gamma(\alpha)}\frac{1}{\sqrt{\sigma^2t^2+\lambda}}\exp\left[-\frac{2\beta+\lambda\mu^2}{2\sigma^2}+\frac{\lambda^2\mu^2}{2\sigma^2\left(\lambda+\sigma^2t^2\right)}\right].
\end{aligned}$$

Derivation for the Uniform distribution

The density function of y is

$$f_Y(y) = \frac{1}{b-a},$$

for $-\infty < a < y < b < +\infty$.

The characteristic function of the product can be derived as

$$\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_a^b \frac{1}{b-a}\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{1}{b-a}\int_a^b \exp\left(-\frac{t^2}{2}y^2\right)dy.
\end{aligned}$$

Let $x = \frac{t^2}{2}y^2$, $x \in \left[\frac{t^2a^2}{2}, \frac{b^2t^2}{2}\right]$, then $y = \frac{1}{t}\sqrt{2x}$ and $dy = \frac{1}{\sqrt{2t}}x^{-\frac{1}{2}}dx$.

$$\begin{aligned}\phi_w(t) &= \frac{1}{(b-a)\sqrt{2t}}\int_{\frac{t^2a^2}{2}}^{\frac{t^2b^2}{2}} x^{-\frac{1}{2}}\exp(-x)dx \\ &= \frac{1}{(b-a)\sqrt{2t}}\left[\Gamma\left(\frac{1}{2}, \frac{t^2a^2}{2}\right) - \Gamma\left(\frac{1}{2}, \frac{t^2b^2}{2}\right)\right].\end{aligned}$$

Derivation for the Rayleigh distribution (Weibull, 1951)

The density function of y is

$$f_Y(y) = 2\lambda^2 y \exp\left(-\lambda^2 y^2\right),$$

for $y > 0$ and $\lambda > 0$.

The characteristic function of the product can be derived as

$$\begin{aligned}\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\ &= 2\lambda^2\int_0^\infty y\exp\left(-\left(\lambda^2 + \frac{t^2}{2}\right)y^2\right)dy.\end{aligned}$$

Let $(\lambda^2 + \frac{t^2}{2})y^2 = x$, $x \in (0, \infty)$, then $y = (\lambda^2 + \frac{t^2}{2})^{-\frac{1}{2}}x^{\frac{1}{2}}$ and $dy = (\lambda^2 + \frac{t^2}{2})^{-\frac{1}{2}} \cdot \frac{1}{2} \cdot x^{-\frac{1}{2}}dx$.

$$\begin{aligned}\phi_w(t) &= 2\lambda^2\int_0^\infty \left(\lambda^2 + \frac{t^2}{2}\right)^{-\frac{1}{2}} x^{\frac{1}{2}}\exp(-x)\left(\lambda^2 + \frac{t^2}{2}\right)^{-\frac{1}{2}} \cdot \frac{1}{2} \cdot x^{-\frac{1}{2}}dx \\ &= \frac{\lambda^2}{\lambda^2 + \frac{t^2}{2}}\int_0^\infty \exp(-x)dx \\ &= \frac{2\lambda^2}{2\lambda^2 + t^2}.\end{aligned}$$

Derivation for the q-Gaussian distribution (Tsallis, 2009)

The density function of y is

$$f_Y(y) = \frac{\sqrt{\beta}}{C}\left[1 - (1-q)\beta y^2\right]^{\frac{1}{1-q}},$$

for $-\infty < y < +\infty$ if $1 \leq q < 3$, $-\frac{1}{\sqrt{\beta(1-q)}} < y < +\frac{1}{\sqrt{\beta(1-q)}}$ if $q < 1$ and $\beta > 0$, where C denotes the normalizing constant.

When $1 \leq q < 3$, that is, $-\infty < y < +\infty$. The characteristic function of the product can be derived as,

$$\begin{aligned}\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\ &= \frac{\sqrt{\beta}}{C}\int_{-\infty}^{\infty}\left[1-(1-q)\beta y^2\right]^{\frac{1}{1-q}}\exp\left(-\frac{t^2}{2}y^2\right)dy \\ &= \frac{2\sqrt{\beta}}{C}\int_{0}^{\infty}\left[1+(q-1)\beta y^2\right]^{-\frac{1}{q-1}}\exp\left(-\frac{t^2}{2}y^2\right)dy.\end{aligned}$$

Let $x=(q-1)\beta y^2$, $x \in (0,\infty)$, then $y = [(q-1)\beta]^{-\frac{1}{2}}x^{\frac{1}{2}}$ and $dy = \frac{1}{2}\left[(q-1)\beta\right]^{-\frac{1}{2}}x^{-\frac{1}{2}}dx$.

$$\begin{aligned}\phi_w(t) &= \frac{1}{C[q-1]^{\frac{1}{2}}}\int_{0}^{\infty}x^{-\frac{1}{2}}(1+x)^{-\frac{1}{q-1}}\exp\left(-\frac{t^2}{2\beta(q-1)}x\right)dx \\ &= \frac{1}{C(q-1)^{\frac{1}{2}}}\Gamma\left(\frac{1}{2}\right)\Psi\left(\frac{1}{2},\frac{3}{2}-\frac{1}{q-1};\frac{t^2}{2\beta(q-1)}\right) \\ &= \frac{\sqrt{2}}{C\sqrt{q-1}}\Psi\left(\frac{1}{2},\frac{3}{2}-\frac{1}{q-1};\frac{t^2}{2\beta(q-1)}\right),\end{aligned}$$

as $\int_0^\infty x^{\alpha-1}(x+z)^{-q}e^{-px}dx = \Gamma(\alpha)z^{\alpha-q}\Psi(\alpha,\alpha+1-q;zp)$, and $\alpha=\frac{1}{2}, z=1, q=\frac{1}{q-1}, p=\frac{t^2}{2\beta(q-1)}$.

When $q<1$, that is $-\frac{1}{\sqrt{\beta(1-q)}} < y < +\frac{1}{\sqrt{\beta(1-q)}}$. The characteristic function of the product can be derived as,

$$
\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \frac{\sqrt{\beta}}{C}\int_{-\infty}^{\infty}\left[1-(1-q)\beta y^2\right]^{\frac{1}{1-q}}\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{\sqrt{\beta}}{C}\int_{-\frac{1}{\sqrt{\beta(1-q)}}}^{\frac{1}{\sqrt{\beta(1-q)}}}\left[1-(1-q)\beta y^2\right]^{\frac{1}{1-q}}\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{2\sqrt{\beta}}{C}\int_{0}^{\frac{1}{\sqrt{\beta(1-q)}}}\left[1-(1-q)\beta y^2\right]^{\frac{1}{1-q}}\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \frac{2\sqrt{\beta}}{C\sqrt{\beta(1-q)}}\int_0^1\left[1-\left(\sqrt{\beta(1-q)}y\right)^2\right]^{\frac{1}{1-q}} \\
&\quad \exp\left(-\frac{t^2}{2(\sqrt{(1-q)\beta})^2}\left(\sqrt{(1-q)\beta}y\right)^2\right)d\sqrt{(1-q)\beta}y \\
&= \frac{\sqrt{\beta}}{C\sqrt{\beta(1-q)}}\int_0^1 x^{-\frac{1}{2}}(1-x)^{\frac{1}{1-q}}\exp\left(-\frac{t^2}{2\beta(1-q)}x\right)dx \\
&= \frac{1}{C\sqrt{1-q}}\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{2-q}{1-q}\right)}{\Gamma\left(\frac{1}{2}+\frac{2-q}{1-q}\right)}{}_1F_1\left(\frac{1}{2};\frac{1}{2}+\frac{2-q}{1-q};\frac{-t^2}{2\beta(1-q)}\right) \\
&= \frac{B(\frac{1}{2},\frac{2-q}{1-q})}{C\sqrt{1-q}}{}_1F_1\left(\frac{1}{2};\frac{1}{2}+\frac{2-q}{1-q};\frac{-t^2}{2\beta(1-q)}\right),
\end{aligned}
$$

as ${}_1F_1(a;b;z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)}\int_0^1 t^{a-1}(1-t)^{b-a-1}e^{zt}dt$, and $a = \frac{1}{2}, b = \frac{1}{2}+\frac{2-q}{1-q}, z = \frac{-t^2}{2\beta(1-q)}$.

Derivation for the Normal Exponential Gamma distribution

The density function of y is

$$
f_Y(y) = C\exp\left(-\frac{y^2}{4\theta^2}\right)D_{-2k-1}\left(\frac{\mid y\mid}{\theta}\right),
$$

for $-\infty < y < +\infty$, $k > 0$ and $\theta > 0$, where C denotes the normalizing constant.

The characteristic function of the product can be derived as,

$$
\begin{aligned}
\phi_w(t) &= E\left[\exp\left(-\frac{t^2}{2}y^2\right)\right] \\
&= \int_{-\infty}^{\infty} C\exp\left(-\frac{y^2}{4\theta^2}\right) D_{-2k-1}\left(\frac{|\,y\,|}{\theta}\right)\exp\left(-\frac{t^2}{2}y^2\right)dy \\
&= \int_{-\infty}^{0} C\exp\left(-\frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right)y^2\right) D_{-2k-1}\left(\frac{-y}{\theta}\right)dy \\
&\quad + \int_{0}^{\infty} C\exp\left(-\frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right)y^2\right) D_{-2k-1}\left(\frac{y}{\theta}\right)dy
\end{aligned}
$$

Let $\zeta = -y$ when $y < 0$, thus, $\zeta \in (0, \infty)$.

$$
\begin{aligned}
\phi_w(t) &= \int_{0}^{\infty} C\exp\left(-\frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right)\zeta^2\right) D_{-2k-1}\left(\frac{\zeta}{\theta}\right)d\zeta \\
&\quad + \int_{0}^{\infty} C\exp\left(-\frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right)y^2\right) D_{-2k-1}\left(\frac{y}{\theta}\right)dy \\
&= 2C\int_{0}^{\infty} \exp\left(-\frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right)y^2\right) D_{-2k-1}\left(\frac{y}{\theta}\right)dy \\
&= 2C\cdot 2^{-k-1}\sqrt{\pi}\cdot\theta\cdot\Gamma\begin{bmatrix}1\\ k+3/2\end{bmatrix} {}_2F_1\left(\frac{1}{2},1;k+\frac{3}{2};\frac{\frac{1}{\theta^2}-2\left(\frac{1}{2\theta^2}+t^2\right)}{2/\theta^2}\right) \\
&= C\cdot 2^{-k}\cdot\sqrt{\pi}\cdot\theta\cdot\Gamma^{-1}\left(k+\frac{3}{2}\right){}_2F_1\left(\frac{1}{2},1;k+\frac{3}{2};-t^2\theta^2\right),
\end{aligned}
$$

as $\int_0^\infty x^{\alpha-1}e^{-px^2}D_\nu(cx)dx = 2^{(\nu-\alpha)/2}\sqrt{\pi}c^{-\alpha}\Gamma\begin{bmatrix}\alpha\\ (1+\alpha-\nu)/2\end{bmatrix} {}_2F_1\left(\frac{\alpha}{2},\frac{\alpha+1}{2};\frac{1+\alpha-\nu}{2};\frac{c^2-4p}{2c^2}\right)$, and $\alpha = 1, p = \frac{1}{2}\left(\frac{1}{2\theta^2}+t^2\right), \nu = -2k-1, c = \frac{1}{\theta}$.

Appendix D

Appendix to Chapter 8

Here, we present the derivations of the density function, the cumulative distribution function, and the truncated expectation of the aggregated risk, as well as the expected contribution of each risk to the aggregated risk, $E\left[X1_{S>b}\right]$. The selected models are the representatives of the different classes of bivariate distribution discussed in Chapter 8. They are the mixtures of independent exponential distribution, the Gumbel's bivariate exponential distribution, the Block and Basu (1976's) bivariate exponential distribution, the mixtures of independent gamma distribution with real shape parameters, the Negar and Sepulveda-Murillo (2011)'s bivariate confluent hypergeometric distribution, the Mardia (1970's) bivariate distribution in Eq (5.77) of Balakrishnan and Lai(2009), the bivariate Pareto distribution with equal scale parameter, the generalized Pareto bivariate distribution.

Mixtures of independent exponential distribution

$$f(x,y) = \sum_{k=1}^{m} C_k \exp\left(-\alpha_k x - \beta_k y\right)$$

$$\begin{aligned} f(x,s-x) &= \sum_{k=1}^{m} C_k \exp\left(-\alpha_k x - \beta_k s + \beta_k x\right) \\ &= \sum_{k=1}^{m} C_k \exp\left(-\beta_k s\right) \exp\left(-\left(\alpha_k - \beta_k\right) x\right) \end{aligned}$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x, s-x)dx \\
&= \sum_{k=1}^m C_k \exp\left(-\beta_k s\right) \int_0^s \exp\left(-\left(\alpha_k - \beta_k\right) x\right) dx \\
&= \sum_{k=0}^m C_k \exp\left(-\beta_k s\right) \cdot \frac{1}{\alpha_k - \beta_k} \left[1 - \exp\left(-\left(\alpha_k - \beta_k\right) s\right)\right] \\
&= \sum_{k=1}^m C_k \frac{1}{\alpha_k - \beta_k} \left[\exp\left(-\beta_k s\right) - \exp\left(-\alpha_k s\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
F_S(s) &= \int_0^s f_S(t)dt \\
&= \sum_{k=1}^m \frac{C_k}{\alpha_k - \beta_k} \left[\int_0^s \exp\left(-\beta_k t\right) dt - \int_0^s \exp\left(-\alpha_k t\right) dt\right] \\
&= \sum_{k=1}^m \frac{C_k}{\alpha_k - \beta_k} \left[\frac{1 - \exp\left(-\beta_k s\right)}{\beta_k} - \frac{1 - \exp\left(-\alpha_k s\right)}{\alpha_k}\right]
\end{aligned}
$$

$$
\begin{aligned}
g_{x,s}(s) &= \int_0^s x \cdot f(x, s-x)dx \\
&= \sum_{k=1}^m C_k \exp\left(-\beta_k s\right) \cdot \int_0^s x \exp\left(-\left(\alpha_k - \beta_k\right) x\right) dx \\
&= -\frac{\sum_{k=1}^m C_k \exp\left(-\beta_k s\right)}{\left(\alpha_k - \beta_k\right)^2} \int_0^{s(\alpha_k - \beta_k)} t d \exp\left(-t\right) \\
&= -\frac{\sum_{k=1}^m C_k \exp\left(-\beta_k s\right)}{\left(\alpha_k - \beta_k\right)^2} \left[t \exp\left(-t\right) |_0^{s(\alpha_k - \beta_k)} - \int_0^{s(\alpha_k - \beta_k)} \exp\left(-t\right) dt\right] \\
&= -\frac{\sum_{k=1}^m C_k \exp\left(-\beta_k s\right)}{\left(\alpha_k - \beta_k\right)^2} \left[s\left(\alpha_k - \beta_k\right) \exp\left(-\left(\alpha_k - \beta_k\right) s\right) - \left(1 - \exp\left(-\left(\alpha_k - \beta_k\right) s\right)\right)\right] \\
&= \sum_{k=1}^m C_k \exp\left(-\beta_k s\right) \left[\frac{1 - \exp\left(-\left(\alpha_k - \beta_k\right) s\right)}{(\alpha_k - \beta_k)^2} - \frac{s \cdot \exp\left(-\left(\alpha_k - \beta_k\right) s\right)}{\left(\alpha_k - \beta_k\right)}\right]
\end{aligned}
$$

$$\begin{aligned}
E\left[S_{1_{(s>b)}}\right] &= \int_b^{\infty} s f_S(s) ds \\
&= \sum_{k=1}^{m} \frac{C_k}{(\alpha_k-\beta_k)}\left[\int_b^{\infty} s\exp\left(-\beta_k s\right) ds - \int_b^{\infty} s\exp\left(-\alpha_k s\right) ds\right] \\
&= \sum_{k=1}^{m} \frac{C_k}{(\alpha_k-\beta_k)}\left[\frac{\int_{b\beta_k}^{\infty} t\exp\left(-t\right) dt}{\beta_k^2} - \frac{\int_{b\alpha_k}^{\infty} t\exp\left(-t\right) dt}{\alpha_k^2}\right] \\
&= \sum_{k=1}^{m} \frac{C_k}{(\alpha_k-\beta_k)}\left[\frac{(1+b\beta_k)\exp(-b\beta_k)}{\beta_k^2} - \frac{(1+b\alpha_k)\exp(-b\alpha_k)}{\alpha_k^2}\right]
\end{aligned}$$

$$\begin{aligned}
E\left[X_{1_{(s>b)}}\right] &= \int_b^{\infty} g_{x,s}(s) ds \\
&= \sum_{k=1}^{m} C_k\left[\int_b^{\infty} \frac{\exp(-\beta_k s)-\exp(-\alpha_k s)}{(\alpha_k-\beta_k)^2} ds - \int_b^{\infty} \frac{s\exp(-\alpha_k s)}{(\alpha_k-\beta_k)} ds\right] \\
&= \sum_{k=1}^{m} C_k\left[\frac{1}{(\alpha_k-\beta_k)^2}\left(\int_b^{\infty} \exp\left(\beta_k s\right) ds - \int_b^{\infty} \exp\left(-\alpha_k s\right) ds\right) - \frac{1}{(\alpha_k-\beta_k)}\left(\frac{1}{\alpha_k^2}\int_{b\alpha_k}^{\infty} te^{-t} dt\right)\right] \\
&= \sum_{k=1}^{m} C_k\left[\frac{\exp\left(-b\beta_k\right)}{\beta_k(\alpha_k-\beta_k)^2} - \frac{\exp\left(-b\alpha_k\right)}{\alpha_k\left(\alpha_k-\beta_k\right)^2} - \frac{(1+b\alpha_k)\exp\left(-b\alpha_k\right)}{\alpha_k^2(\alpha_k-\beta_k)}\right]
\end{aligned}$$

Gumbel's bivariate exponential distribution

$$f(x,Y) = \left[(1+\theta x)(1+\theta y) - \theta\right]\exp\left(-x-y-\theta xy\right)$$

$$\begin{aligned}
f(x,s-x) &= \left[1+\theta x+\theta\left(s-x\right)+\theta^2 sx-\theta^2 x^2-\theta\right]\exp\left(-s-\theta sx+\theta x^2\right) \\
&= \left[(1+s\theta-\theta)+\theta^2 sx-\theta^2 x^2\right]\exp\left(-s\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)\exp\left(-\frac{\theta}{4}s^2\right)
\end{aligned}$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x, s-x)dx \\
&= (1+s\theta-\theta)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\theta^2\cdot s\cdot\exp\left(-s-\frac{\theta^2}{4}s^2\right)\int_0^s x\cdot\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s x^2\cdot\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= (1+s\theta-\theta)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\theta^2 s\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\frac{\theta^2 s^2}{2}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -s\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\frac{s^2\theta^2}{4}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= \left(1+s\theta-\theta+\frac{s^2\theta^2}{2}-\frac{s^2\theta^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\left(\theta^2 s-s\theta^2\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s \left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= 2\left(1+\theta(s-1)+\frac{s^2\theta^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\exp\left(-\left(\theta i^2\right)\zeta^2\right)d\zeta \\
&\quad -2\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\zeta^2\exp\left(-\left(\theta i^2\right)\zeta^2\right)d\zeta
\end{aligned}
$$

Note that

$$\int_0^{b_1}\exp\left(-a_1t^2\right)dt = \frac{\sqrt{\pi}\mathrm{erf}\left(\sqrt{a_1}b_1\right)}{2\sqrt{a_1}}$$

and

$$\int_0^{b_2} t^2\exp\left(-a_2t^2\right)dt = -\frac{b_2\exp\left(-a_2b_2^2\right)}{2a_2}+\frac{\sqrt{\pi}\mathrm{erf}\left(\sqrt{a_2}b_2\right)}{4\cdot a_2^{3/2}},$$

so,

$$\int_0^{\frac{s}{2}}\exp\left(-\left(\theta i^2\right)\zeta^2\right)d\zeta = \frac{\sqrt{\pi}\mathrm{erf}\left(i\cdot\sqrt{\theta}\cdot\frac{s}{2}\right)}{2i\cdot\sqrt{\theta}}$$

and

$$\int_0^{\frac{s}{2}} \zeta^2 \exp\left(-\left(\theta i^2\right)\zeta^2\right) d\zeta = -\frac{\frac{s}{2}\exp\left(\theta\frac{s^2}{4}\right)}{-2\theta} + \frac{\sqrt{\pi}\mathrm{erf}\left(i\cdot\sqrt{\theta}\cdot\frac{s}{2}\right)}{4\cdot(\theta\cdot i^2)^{3/2}}$$

$$\begin{aligned}
f_S(s) &= 2\left(1+\theta(s-1)+\frac{s^2\theta^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\frac{\sqrt{\pi}\mathrm{erf}\left(\frac{i\cdot\sqrt{\theta}}{2}s\right)}{2i\cdot\sqrt{\theta}} \\
&\quad -2\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\left[\frac{\frac{s}{2}\exp\left(\theta\frac{s^2}{4}\right)}{2\theta} + \frac{\sqrt{\pi}\mathrm{erf}\left(i\cdot\sqrt{\theta}\cdot\frac{s}{2}\right)}{-4\cdot i\cdot\theta^{3/2}}\right] \\
&= \left(-i\sqrt{\frac{\pi}{\theta}} - i\sqrt{\theta\pi}(s-1) - i\sqrt{\pi}\theta^{3/2}\frac{s^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\mathrm{erf}\left(\frac{i\cdot\sqrt{\theta}}{2}s\right) \\
&\quad -\frac{i\sqrt{\theta\pi}}{2}\exp\left(-s-\frac{\theta}{4}s^2\right)\mathrm{erf}\left(\frac{i\cdot\sqrt{\theta}}{2}s\right) - \frac{\theta}{2}s\exp(-s) \\
&= \left(-i\sqrt{\frac{\pi}{\theta}} - i\sqrt{\theta\pi}(s-\frac{1}{2}) - i\sqrt{\pi}\theta^{3/2}\frac{s^2}{4}\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\mathrm{erf}\left(\frac{i\cdot\sqrt{\theta}}{2}s\right) - \frac{\theta}{2}s\exp(-s)
\end{aligned}$$

$$
\begin{aligned}
g_{x,s}(s) &= \int_0^s x\cdot f(x,s-x)dx \\
&= \int_0^s \left[(1+s\theta-\theta)x+\theta^2 sx^2-\theta^2x^3\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= (1+s\theta-\theta)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\frac{s}{2}(1+s\theta-\theta)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\theta^2 s\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\theta^2 s\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\frac{\theta^2 s^3}{4}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)^3\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\frac{3\theta^2 s^2}{2}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\frac{3\theta^2 s^2}{4}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\frac{\theta^2 s^3}{8}\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= \left[\frac{s}{2}+\frac{s^2\theta}{2}-\frac{s\theta}{2}+\frac{\theta^2s^3}{4}-\frac{\theta^2s^3}{8}\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad \left[1+s\theta-\theta+\theta^2s^2-\frac{3\theta^2}{4}s^2\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad +\left(\theta^2s-\frac{3}{2}\theta^2s^2\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&\quad -\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^s\left(x-\frac{s}{2}\right)^3\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)dx \\
&= 2\left[\frac{s}{2}+\frac{s\theta}{2}(s-1)+\frac{\theta^2}{8}s^3\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)d(x-\frac{s}{2}) \\
&\quad +2\left[1+\theta(s-1)+\frac{\theta^2}{4}s^2\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\left(x-\frac{s}{2}\right)\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)d(x-\frac{s}{2}) \\
&\quad +2\left(-\frac{\theta^2}{2}s^2\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\left(x-\frac{s}{2}\right)^2\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)d(x-\frac{s}{2}) \\
&\quad -2\theta^2\exp\left(-s-\frac{\theta}{4}s^2\right)\int_0^{\frac{s}{2}}\left(x-\frac{s}{2}\right)^3\exp\left(\theta\left(x-\frac{s}{2}\right)^2\right)d(x-\frac{s}{2})
\end{aligned}
$$

Note that,

$$\int_0^{b_1}\exp\left(-a_1t^2\right)dt=\frac{\sqrt{\pi}\mathrm{erf}\left(\sqrt{a_1}b_1\right)}{2\sqrt{a_1}},$$

$$\int_0^{b_2} t \exp\left(-a_2 t^2\right) dt = \frac{1 - e^{-a_2 b_2^2}}{2a_2},$$

$$\int_0^{b_3} t^2 \exp\left(-a_3 t^2\right) dt = \frac{-b_3 e^{-a_3 b_3^2}}{2a_3} + \frac{\sqrt{\pi}\mathrm{erf}\left(\sqrt{a_3} b_3\right)}{4a_3^{3/2}},$$

$$\int_0^{b_4} t^3 \exp\left(-a_4 t^2\right) dt = \frac{1 - \left(1 + a_4 b_4^2\right) e^{-a_4 b_4^2}}{2a_4^2}.$$

Thus,

$$\int_0^{\frac{s}{2}} \exp\left(\theta\left(x - \frac{s}{2}\right)^2\right) d(x - \frac{s}{2}) = \frac{\sqrt{\pi}\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2} s\right)}{2i\sqrt{\theta}}$$

$$\int_0^{\frac{s}{2}} \left(x - \frac{s}{2}\right) \exp\left(\theta\left(x - \frac{s}{2}\right)^2\right) d(x - \frac{s}{2}) = \frac{\exp\left(\frac{\theta}{4} s^2\right) - 1}{2\theta}$$

$$\int_0^{\frac{s}{2}} \left(x - \frac{s}{2}\right)^2 \exp\left(\theta\left(x - \frac{s}{2}\right)^2\right) d(x - \frac{s}{2}) = \frac{\frac{s}{2}\exp\left(\frac{\theta}{4} s^2\right)}{2\theta} + \frac{\sqrt{\pi}\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2} s\right)}{-4i\theta^{3/2}}$$

$$\int_0^{\frac{s}{2}} \left(x - \frac{s}{2}\right)^3 \exp\left(\theta\left(x - \frac{s}{2}\right)^2\right) d(x - \frac{s}{2}) = \frac{1 - \left(1 - \frac{\theta}{4} s^2\right) \exp\left(\frac{\theta}{4} s^2\right)}{2\theta^2}$$

$$
\begin{aligned}
g_{x,s}(s) &= \left[\frac{s}{2}+\frac{s\theta}{2}(s-1)+\frac{\theta^2}{8}s^3\right]\frac{\sqrt{\pi}\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2}s\right)}{i\sqrt{\theta}}\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad+\left[1+\theta(s-1)+\frac{\theta^2}{4}s^2\right]\frac{\exp\left(\frac{\theta}{4}s^2\right)-1}{\theta}\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad+\left(-\frac{\theta^2}{2}s^2\right)\left[\frac{\frac{s}{2}\exp\left(\frac{\theta}{4}s^2\right)}{\theta}-\frac{\sqrt{\pi}\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2}s\right)}{2i\theta^{3/2}}\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad-\left[1-\left(1-\frac{\theta}{4}s^2\right)\exp\left(\frac{\theta}{4}s^2\right)\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&= \left[\frac{s\sqrt{\pi}}{2i\sqrt{\theta}}+\frac{s\sqrt{\theta\pi}}{2\cdot i}(s-1)+\frac{\sqrt{\pi}\theta^{3/2}s^3}{8i}+\frac{\sqrt{\pi}\theta^2s^2}{4i\theta^{3/2}}\right]\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2}s\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad+\left[\frac{\exp\left(\frac{\theta}{4}s^2\right)-1}{\theta}+(s-1)\left(\exp\left(\frac{\theta}{4}s^2\right)-1\right)+\frac{\theta s^2}{4}\left(\exp\left(\frac{\theta}{4}s^2\right)-1\right)\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad-\left[\frac{\theta s^2\cdot\frac{s}{2}\exp\left(\frac{\theta}{4}s^2\right)}{2}+1-\left(1-\frac{\theta}{4}s^2\right)\exp\left(\frac{\theta}{4}s^2\right)\right]\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&= \left[-i\cdot\frac{s}{2}\sqrt{\frac{\pi}{\theta}}-i\cdot\frac{3s^2\sqrt{\pi\theta}}{4}-i\frac{\sqrt{\pi}\theta^{3/2}s^3}{8}+\frac{i\cdot s\sqrt{\pi\theta}}{2}\right]\mathrm{erf}\left(\frac{i\sqrt{\theta}}{2}s\right)\exp\left(-s-\frac{\theta}{4}s^2\right)\\
&\quad+\left(\frac{1}{\theta}+s-\frac{\theta}{4}s^3\right)\exp(-s)-\left(\frac{1}{\theta}+s+\frac{\theta}{4}s^2\right)\exp\left(-s-\frac{\theta}{4}s^2\right)
\end{aligned}
$$

Block and Basu (1976)'s bivariate exponential distribution

$$
f(x,y) = \begin{cases} C\exp\left(-\alpha x-\beta y\right), & \text{if } x>y,\\ D\exp\left(-\gamma x-\delta y\right), & \text{if } x\le y,\end{cases}
$$

for $\alpha>0$, $\beta>0$, $\gamma>0$, $\delta>0$, $x>0$ and $y>0$.

$$
f(x,s-x) = \begin{cases} C\exp\left(-\left(\alpha-\beta\right)x-\beta s\right), & \text{if } x>y,\\ D\exp\left(-\left(\gamma-\delta\right)x-\delta s\right), & \text{if } x\le y,\end{cases}
$$

Note that, because $x+y=s$, $x>y$ indicates that $x>s-x$, which is equal to $x>\frac{s}{2}$. That is, when $x \in \left(\frac{s}{2}, s\right]$, $x > y$. Therefore, when $x \in \left[0, \frac{s}{2}\right]$,x $\leq y$.

$$\begin{aligned}
f(s) &= \int_0^s f(x, s-x)dx \\
&= \int_0^{\frac{s}{2}} f(x, s-x)dx + \int_{\frac{s}{2}}^s f(x, s-x)dx \\
&= \int_0^{\frac{s}{2}} D\exp\left(-\left(\gamma-\delta\right)x - \delta s\right)dx + \int_{\frac{s}{2}}^s C\exp\left(-\left(\alpha-\beta\right)x - \beta s\right)dx \\
&= \frac{D\exp\left(-\delta s\right)}{\gamma-\delta}\left[1-\exp\left(-\left(\gamma-\delta\right)\frac{s}{2}\right)\right] + \frac{C\exp\left(-\beta s\right)}{\alpha-\beta}\left[\exp\left(-\frac{\alpha-\beta}{2}s\right) - \exp\left(-\left(\alpha-\beta\right)s\right)\right] \\
&= \frac{C}{\beta-\alpha}\left[\exp\left(-\alpha s\right) - \exp\left(-\frac{\alpha+\beta}{2}s\right)\right] - \frac{D}{\delta-\gamma}\left[\exp\left(-\delta s\right) - \exp\left(-\left(\gamma+\delta\right)\frac{s}{2}\right)\right]
\end{aligned}$$

$$\begin{aligned}
F_S(s) &= \int_0^s f_s(t)dt \\
&= \frac{C}{\beta-\alpha}\left[\int_0^s \exp\left(-\alpha t\right)dt - \int_0^s \exp\left(-\frac{\alpha+\beta}{2}t\right)dt\right] \\
&\quad -\frac{D}{\delta-\gamma}\left[\int_0^s \exp\left(-\delta t\right)dt - \int_0^s \exp\left(-\frac{\gamma+\delta}{2}t\right)dt\right] \\
&= \frac{C}{\beta-\alpha}\left[\frac{1-\exp\left(-\alpha s\right)}{\alpha} - \frac{2}{\alpha+\beta}\left(1-\exp\left(-\frac{\alpha+\beta}{2}s\right)\right)\right] \\
&\quad -\frac{D}{\gamma-\delta}\left[\frac{2}{\gamma+\delta}\left(1-\exp\left(-\frac{\delta+\gamma}{2}s\right)\right) - \frac{1}{\delta}\left(1-\exp\left(-\delta s\right)\right)\right]
\end{aligned}$$

$$\begin{aligned}
E\left[S_{1_{S>b}}\right] &= \int_b^\infty s f_S(s)ds \\
&= \frac{C}{\beta-\alpha}\left[\int_b^\infty s\exp\left(-\alpha s\right)ds - \int_b^\infty s\exp\left(-\frac{\alpha+\beta}{2}s\right)\right] \\
&\quad -\frac{D}{\delta-\gamma}\left[\int_b^\infty s\exp\left(-\delta s\right)ds - \int_b^\infty s\exp\left(-\frac{\gamma+\delta}{2}s\right)ds\right] \\
&= \frac{C}{\beta-\alpha}\left[\frac{1}{\alpha^2}\int_{\alpha b}^\infty \alpha s\exp\left(-\alpha s\right)d\alpha s - \frac{4}{(\alpha+\beta)^2}\int_{\frac{(\alpha+\beta)b}{2}}^\infty \frac{\alpha+\beta}{2}s\exp\left(-\frac{\alpha+\beta}{2}s\right)d\left(\frac{\alpha+\beta}{2}\right)s\right] \\
&\quad -\frac{D}{\delta-\gamma}\left[\frac{1}{\delta^2}\int_{\delta b}^\infty \delta s\exp\left(-\delta s\right)d\delta s - \frac{4}{(\gamma+\delta)^2}\int_{\frac{(\gamma+\delta)b}{2}}^\infty \frac{(\gamma+\delta)}{2}s\exp\left(-\frac{(\gamma+\delta)}{2}s\right)d\frac{(\gamma+\delta)}{2}s\right. \\
&= \frac{C}{\beta-\alpha}\left[\frac{(1+\alpha b)}{\alpha^2}\exp\left(-\alpha b\right) - \frac{4}{(\alpha+\beta)^2}\left(1+\frac{\alpha+\beta}{2}b\right)\exp\left(-\frac{\alpha+\beta}{2}b\right)\right] \\
&\quad -\frac{D}{\delta-\gamma}\left[\frac{1+\delta b}{\delta^2}\exp\left(-\delta b\right) - \frac{4}{(\gamma+\delta)^2}\left(1+\frac{\gamma+\delta}{2}b\right)\exp\left(-\frac{\gamma+\delta}{2}b\right)\right]
\end{aligned}$$

$$
\begin{aligned}
g_{X,S}(s) &= \int_0^s x f(x, s-x) dx \\
&= D \exp\left(-\delta s\right) \int_0^{\frac{s}{2}} x \exp\left(-\left(\gamma - \delta\right) x\right) dx \\
&\quad + C \exp\left(-\beta s\right) \int_{\frac{s}{2}}^{s} x \exp\left(-\left(\alpha - \beta\right) x\right) dx \\
&= D \exp\left(-\delta s\right) \int_0^{\frac{s}{2}} x \exp\left(-\left(\gamma - \delta\right) s\right) dx \\
&\quad + C \exp\left(-\beta s\right) \int_0^{s} x \exp\left(-\left(\alpha - \beta\right) x\right) dx \\
&\quad - C \exp\left(-\beta s\right) \int_0^{\frac{s}{2}} x \exp\left(-\left(\gamma - \delta\right) x\right) dx \\
&= \frac{D \exp(-\delta s)}{(\gamma - \delta)^2} \int_0^{\frac{s}{2}(\gamma-\delta)} x(\gamma - \delta) \exp\left(-\left(\gamma - \delta\right) x\right) dx(\gamma - \delta) \\
&\quad + \frac{C \exp(-\beta s)}{(\alpha - \beta)^2} \int_0^{s(\alpha-\beta)} x(\alpha - \beta) \exp\left(-\left(\alpha - \beta\right) x\right) dx \left(\alpha - \beta\right) \\
&\quad - \frac{C \exp(-\beta s)}{(\gamma - \delta)^2} \int_0^{\frac{s}{2}(\gamma-\delta)} x(\gamma - \delta) \exp\left(-\left(\gamma - \delta\right) x\right) dx \left(\gamma - \delta\right) \\
&= D \left[\frac{\exp\left(-\delta s\right) - \exp\left(-\left(\delta + \gamma\right) s/2\right)}{(\delta - \gamma)^2} - \frac{s \exp\left(-\left(\delta + \gamma\right) s/2\right)}{2(\delta - \gamma)} \right] \\
&\quad + C \left[\frac{\exp\left(-\beta s\right) - \exp\left(-\alpha s\right)}{(\alpha - \beta)^2} - \frac{s \exp\left(-\alpha s\right)}{\alpha - \beta} \right] \\
&\quad - C \left[\frac{\exp\left(-\beta s\right) - \exp\left(-\beta s - \left(\gamma - \delta\right) s/2\right)}{(\delta - \gamma)^2} - \frac{s \exp\left(-\beta s - \left(\gamma - \delta\right) s/2\right)}{2(\delta - \gamma)} \right],
\end{aligned}
$$

as $\int_0^a x \exp(-x)\, dx = 1 - \exp(-a) - a\exp(-a)$.

$$
\begin{aligned}
E\left[X_{1_{S>b}}\right] &= \int_b^\infty g_{X,S}(s) ds \\
&= D\int_b^\infty \left[\frac{\exp(-\delta s) - \exp(-(\delta+\gamma)s/2)}{(\delta-\gamma)^2} - \frac{s\exp(-(\delta+\gamma)s/2)}{2(\delta-\gamma)}\right] ds \\
&\quad +C\int_b^\infty \left[\frac{\exp(-\beta s) - \exp(-\alpha s)}{(\alpha-\beta)^2} - \frac{s\exp(-\alpha s)}{\alpha-\beta}\right] ds \\
&\quad -C\int_b^\infty \left[\frac{\exp(-\beta s) - \exp(-\beta s-(\gamma-\delta)s/2)}{(\delta-\gamma)^2} - \frac{s\exp(-\beta s-(\gamma-\delta)s/2)}{2(\delta-\gamma)}\right] ds \\
&= C\left[-\frac{(b\alpha+1)\exp(-b\alpha)}{\alpha^2(\alpha-\beta)} + \frac{\exp(-b\beta)}{\beta(\alpha-\beta)^2} - \frac{\exp(-b\alpha)}{\alpha(\alpha-\beta)^2}\right] \\
&\quad +D\left[\frac{\exp(-b\delta)}{\delta(\delta-\gamma)^2} - \frac{2\exp(-(\delta+\gamma)b/2)}{(\delta-\gamma)^2(\delta+\gamma)} - \frac{(2+(\delta+\gamma)b)\exp(-(\delta+\gamma)b/2)}{(\delta+\gamma)^2(\delta-\gamma)}\right] \\
&\quad -C\left[\frac{\exp(-b\beta)}{\beta(\delta-\gamma)^2} - \frac{\exp(-b\beta-(\gamma-\delta)b/2)}{(\delta-\gamma)^2(\beta+(\gamma-\delta)/2)}\right] \\
&\quad +C\frac{(2+2b\beta+b(\gamma-\delta))\exp(-b\beta-(\gamma-\delta)b/2)}{(\delta-\gamma)(2\beta+\gamma-\delta)^2},
\end{aligned}
$$

as $\int_b^\infty s\exp(-\alpha s)\, ds = \frac{1+\alpha b}{\alpha^2}\exp(-\alpha b)$ and $\int_b^\infty \exp(-\alpha s)\, ds = \frac{1}{\alpha}\exp(-\alpha b)$.

Mixtures of independent gamma distributions with real shape parameters

$$
f(x,y) = \sum_{k=0}^\infty C_k x^{\alpha_k-1} y^{\beta_k-1}\exp(-\gamma_k x - \delta_k y),
$$

for $\alpha_k > 0$, $\beta_k > 0$, $\gamma_k > 0$, $\delta_k > 0$, $x > 0$ and $y > 0$.

$$
f(x, s-x) = \sum_{k=0}^\infty C_k x^{\alpha_k-1}(s-x)^{\beta_k-1}\exp(-\delta_k s)\exp(-(\gamma_k-\delta_k)x)
$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x,s-x)dx \\
&= \sum_{k=0}^\infty C_k\exp(-\delta_k s)\int_0^s x^{\alpha_k-1}(s-x)^{\beta_k-1}\exp(-(\gamma_k-\delta_k)x)\, dx,
\end{aligned}
$$

as $\int_0^a x^{\alpha-1}(a-x)^{\beta-1}e^{-px}dx = B(\alpha,\beta)a^{\alpha+\beta-1}{}_1F_1(\alpha;\alpha+\beta;-ap)$.

$$\begin{aligned}
f_S(s) &= \sum_{k=0}^{\infty} C_k \exp\left(-\delta_k s\right) B\left(\alpha_k,\beta_k\right) s^{\alpha_k+\beta_k-1}{}_1F_1\left(\alpha_k;\alpha_k+\beta_k;\left(\delta_k-\gamma_k\right)s\right) \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!} s^{\alpha_k+\beta_k+i-1}\exp\left(-\delta_k s\right),
\end{aligned}$$

as ${}_1F_1\left(a,b;z\right) = \sum_{k=0}^{\infty}\frac{(a)_k z^k}{(b)_k k!}$.

$$\begin{aligned}
F(s) &= \int_0^s f_S(t)dt \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!}\int_0^s t^{\alpha_k+\beta_k+i-1}\exp\left(-\delta_k t\right)dt \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!}\delta_k^{-(\alpha_k+\beta_k+i-1)}\delta_k^{-1}\int_0^{\delta_k s}\left(\delta_k t\right)^{\alpha_k+\beta_k+i-1}\exp\left(-\delta_k t\right)dt \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!\cdot\delta_k^{\alpha_k+\beta_k+i}}\gamma\left(\alpha_k+\beta_k+i,\delta_k s\right)
\end{aligned}$$

$$\begin{aligned}
E\left[S_{1_{S>b}}\right] &= \int_b^{\infty} s f_S(s)ds \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!}\int_b^{\infty} s^{\alpha_k+\beta_k+i}\exp\left(-\delta_k s\right)ds \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!}\delta_k^{-(\alpha_k+\beta_k+i)}\delta_k^{-1}\int_{\delta_k b}^{\infty}\left(\delta_k s\right)^{\alpha_k+\beta_k+i}\exp\left(-\delta_k s\right)ds \\
&= \sum_{k=0}^{\infty}\sum_{i=0}^{\infty} \frac{\left(\alpha_k\right)_i\left(\delta_k-\gamma_k\right)^i C_k B\left(\alpha_k,\beta_k\right)}{\left(\alpha_k+\beta_k\right)_i\cdot i!\cdot\delta_k^{\alpha_k+\beta_k+i+1}}\Gamma\left(\alpha_k+\beta_k+i+1,\delta_k b\right)
\end{aligned}$$

$$
\begin{aligned}
g_{X,S}(s) &= \int_b^\infty xf(x,s-x)dx \\
&= \sum_{k=0}^\infty C_k \exp\left(-\delta_k s\right) \int_0^S x^{\alpha_k} \left(s-x\right)^{\beta_k - 1} \exp\left(-\left(\gamma_k - \delta_k\right)x\right)dx \\
&= \sum_{k=0}^\infty C_k \exp\left(-\delta_k s\right) B\left(\alpha_k + 1, \beta_k\right) s^{\alpha_k+\beta_k} {}_1F_1\left(\alpha_k+1; \alpha_k+\beta_k+1; (\delta_k - \gamma_k)s\right) \\
&= \sum_{k=0}^\infty \sum_{i=0}^\infty C_k B(\alpha_k+1, \beta_k) \frac{(\alpha_k+1)_i(\delta_k-\gamma_k)^i}{(1+\alpha_k+\beta_k)_i i!} s^{i+\alpha_k+\beta_k} \exp\left(-\delta_k s\right)
\end{aligned}
$$

$$
\begin{aligned}
E\left[X_{1_{S>b}}\right] &= \int_b^\infty g_{X,S}(s)ds \\
&= \sum_{k=0}^\infty \sum_{i=0}^\infty C_k B(\alpha_k+1, \beta_k) \frac{(\alpha_k+1)_i(\delta_k-\gamma_k)^i}{(1+\alpha_k+\beta_k)_i i!} \int_b^\infty s^{i+\alpha_k+\beta_k} \exp\left(-\delta_k s\right) ds \\
&= \sum_{k=0}^\infty \sum_{i=0}^\infty C_k B(\alpha_k+1, \beta_k) \frac{(\alpha_k+1)_i(\delta_k-\gamma_k)^i}{(1+\alpha_k+\beta_k)_i i!} \left(\delta_k\right)^{-(i+\alpha_k+\beta_k)} \left(\delta_k\right)^{-1} \\
&\qquad \cdot \int_{b\delta_k}^\infty \left(\delta_k s\right)^{i+\alpha_k+\beta_k} \exp\left(-\delta_k s\right) ds \\
&= \sum_{k=0}^\infty \sum_{i=0}^\infty C_k B\left(\alpha_k+1, \beta_k\right) \frac{\left(\alpha_k+1\right)_i \left(\delta_k - \gamma_k\right)^i \Gamma\left(i+\alpha_k+\beta_k+1, \delta_k b\right)}{\left(1+\alpha_k+\beta_k\right)_i i! \delta_k^{\alpha_k+\beta_k+i+1}}
\end{aligned}
$$

Nagar and Sepulveda-Murillo (2011)'s bivariate confluent hypergeometric distribution

$$
f(x,y) = Cx^{p-1}y^{q-1} {}_1F_1\left(\alpha; \beta; -x-y\right),
$$

$$
f(x,s-x) = Cx^{p-1}(s-x)^{q-1} {}_1F_1\left(\alpha; \beta; -s\right)
$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x, s-x)dx \\
&= C \cdot {}_1F_1(\alpha; \beta; -s) \int_0^s x^{p-1} (s-x)^{q-1} dx \\
&= C \cdot B(p,q) \cdot s^{p+q-1} \cdot {}_1F_1(\alpha; \beta; -s),
\end{aligned}
$$

as $\int_0^a x^{\alpha-1}(a-x)^{\beta-1}dx = a^{\alpha+\beta-1}B(\alpha,\beta)$.

$$
\begin{aligned}
F(s) &= \int_0^s f_S(t)dt \\
&= C \cdot B(p,q) \sum_{k=0}^{\infty} \frac{(\alpha)_k(-1)^k}{(\beta)_k k!} \int_0^s t^{p+q+k-1} dt \\
&= C \cdot B(p,q) \sum_{k=0}^{\infty} \frac{(\alpha)_k(-1)^k}{(\beta)_k k!} \frac{s^{p+q+k}}{p+q+k}
\end{aligned}
$$

$$
\begin{aligned}
g_{X,S}(s) &= \int_0^s x f(x, s-x)dx \\
&= C \cdot {}_1F_1(\alpha; \beta; -s) \int_0^s x^p (s-x)^{q-1} dx \\
&= C \cdot B(p+1,q) \cdot s^{p+q} \cdot {}_1F_1(\alpha; \beta; -s)
\end{aligned}
$$

Mardia (1970)'s bivariate distribution in Eq(5.77) of Balakrishnan and Lai (2009)

$$
f(x,y) = C\,(\alpha x+1)^p\,(\beta y+1)^q\,(\gamma x+\delta y+1)^r\,,
$$

for $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $\delta > 0$, $p > 0$, $q > 0$, $r > 0$, $x > 0$ and $y > 0$, provided that p, q and r are integers.

$$
\begin{aligned}
f(x, s-x) &= C \sum_{i=0}^{p} \binom{p}{i} \alpha^i x^i \cdot \sum_{j=0}^{q} \binom{q}{j} \beta^j y^j \cdot \sum_{k=0}^{r} \binom{r}{k} \left(\gamma x + \delta\,(s-x)\right)^k \\
&= C \sum_{i=0}^{p} \binom{p}{i} \alpha^i x^i \cdot \sum_{j=0}^{q} \binom{q}{j} \beta^j y^j \cdot \sum_{k=0}^{r} \sum_{m=0}^{k} \binom{r}{k} \binom{k}{m} \delta^m\,(s-x)^m\,\gamma^{k-m} x^{k-m} \\
&= C \sum_{i=0}^{p} \sum_{j=0}^{q} \sum_{k=0}^{r} \sum_{m=0}^{k} \binom{p}{i} \binom{q}{j} \binom{r}{k} \binom{k}{m} \alpha^i \beta^j \delta^m \gamma^{k-m} x^{i+k+m}\,(s-x)^{m+j}
\end{aligned}
$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x, s-x)dx \\
&= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{k}{m}\alpha^i\beta^j\delta^m\gamma^{k-m}\int_0^s x^{i+k+m}\left(s-x\right)^{m+j}dx \\
&= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{k}{m}\alpha^i\beta^j\delta^m\gamma^{k-m}\cdot B\left(i+k-m+1, j+m+1\right)\cdot s^{i+j+k+1},
\end{aligned}
$$

as $\int_0^a x^{\alpha-1}(a-x)^{\beta-1}dx = a^{\alpha+\beta-1}B(\alpha,\beta)$, and $a = s, \alpha = i+k-m+1, \beta = m+j+1$.

$$
\begin{aligned}
F_S(s) &= \int_0^s f_S(t)dt \\
&= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{k}{m}\alpha^i\beta^j\delta^m\gamma^{k-m}\cdot B\left(i+k-m+1, j+m+1\right)\cdot\frac{s^{i+j+k+2}}{i+j+k+2}
\end{aligned}
$$

$$
\begin{aligned}
g_{X,S}(s) &= \int_0^s x\cdot f(x, s-x)dx \\
&= C\sum_{i=0}^{p}\sum_{j=0}^{q}\sum_{k=0}^{r}\sum_{m=0}^{k}\binom{p}{i}\binom{q}{j}\binom{r}{k}\binom{k}{m}\alpha^i\beta^j\delta^m\gamma^{k-m}\cdot B\left(i+k-m+2, j+m+1\right)\cdot s^{i+j+k+2}
\end{aligned}
$$

Bivariate Pareto distribution with unequal scale parameter

$$
f(x,y) = Cx^{\alpha-1}y^{\beta-1}\left(1+px+qy\right)^{-\gamma},
$$

for $\alpha > 0, \beta > 0, \gamma > 0, p > 0, q > 0, x > 0$ and $y > 0$.

$$
\begin{aligned}
f(x, s-x) &= Cx^{\alpha-1}\left(s-x\right)^{\beta-1}\left(1+px+qs-qx\right)^{-\gamma} \\
&= Cx^{\alpha-1}\left(s-x\right)^{\beta-1}\left(1+qs+(p-q)x\right)^{-\gamma} \\
&= C\left(1+qs\right)^{-\gamma}x^{\alpha-1}\left(s-x\right)^{\beta-1}\left(1+\frac{p-q}{1+qs}x\right)^{-\gamma} \\
&= C(p-q)^{-\gamma}x^{\alpha-1}(s-x)^{\beta-1}\left(\frac{1+qs}{p-q}+x\right)^{-\gamma}
\end{aligned}
$$

$$
\begin{aligned}
f_S(s) &= \int_0^s f(x,s-x)dx \\
&= C\,(p-q)^{-\gamma}\int_0^s x^{\alpha-1}\left((s-x)^{\beta-1}\left(\frac{1+qs}{p-q}+x\right)^{-\gamma}dx\right. \\
&= C\,(p-q)^{-\gamma}\cdot\left(\frac{1+qs}{p-q}\right)^{-\gamma}\cdot s^{\beta+\alpha-1}\cdot B\,(\beta,\alpha)\,{}_2F_1\left(\gamma,\alpha;\alpha+\beta;\frac{(q-p)s}{1+qs}\right) \\
&= C\,(1+qs)^{-\gamma}\,B\,(\alpha,\beta)\,s^{\alpha+\beta-1}{}_2F_1\left(\gamma,\alpha;\alpha+\beta;\frac{(q-p)s}{1+qs}\right),
\end{aligned}
$$

as $\int_0^\mu x^{\nu-1}\,(x+\alpha)^\lambda\,(\mu-x)^{\beta-1}\,dx=\alpha^\lambda\mu^{\beta+\nu-1}B\,(\beta,\nu)\,{}_2F_1\left(-\lambda,\nu;\beta+\nu;-\frac{\mu}{\alpha}\right)$, and $\mu=s,\nu=\alpha,\beta=\beta,\alpha=\frac{1+qs}{p-q},\lambda=-\gamma$.

$$
\begin{aligned}
F(s) &= \int_0^s f_S(t)dt \\
&= CB(\alpha,\beta)\int_0^s t^{\alpha+\beta-1}(1+qt)^{-\gamma}\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k\frac{(q-p)^k}{(1+qt)^k}t^k}{(\alpha+\beta)_k k!}dt \\
&= C\cdot B(\alpha,\beta)\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k}{(\alpha+\beta)_k k!}\int_0^s t^{\alpha+\beta+k-1}(1+qt)^{-(\gamma+k)}dt \\
&= CB(\alpha,\beta)s^{\alpha+\beta}\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k s^k}{(\alpha+\beta)_k k!(k+\alpha+\beta)}\cdot{}_2F_1\left(\gamma+k,\alpha+\beta+k;1+\alpha+\beta+k;-qs\right),
\end{aligned}
$$

as $\int_0^a \frac{x^{\mu-1}}{(1+\beta x)^\nu}dx=\frac{a^\mu}{\mu}{}_2F_1\left(\nu,\mu;1+\mu;-\beta a\right)$, and $a=s,\mu=\alpha+\beta+k,\beta=q,\nu=\gamma+k$.

$$
\begin{aligned}
E\left[S_{1_{S>b}}\right] &= \int_b^\infty s f_S(s)ds \\
&= \int_b^\infty C(1+qs)^{-\gamma}B(\alpha,\beta)s^{\alpha+\beta}{}_2F_1\left(\gamma,\alpha;\alpha+\beta;\frac{(q-p)s}{1+qs}\right)ds \\
&= C\cdot B\,(\alpha,\beta)\int_b^\infty s^{\alpha+\beta}(1+qs)^{-\gamma}\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k s^k(1+qs)^{-k}}{(\alpha+\beta)_K k!}ds \\
&= C\cdot B\,(\alpha,\beta)\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k}{(\alpha+\beta)_K k!}\int_b^\infty s^{\alpha+\beta+k}(1+qs)^{-(\gamma+k)}ds \\
&= C\cdot B\,(\alpha,\beta)\sum_{k=0}^{\infty}\frac{(\gamma)_k(\alpha)_k(q-p)^k}{(\alpha+\beta)_K k!}\cdot\frac{b^{\alpha+\beta-\gamma+1}}{q^{\gamma+k}(\gamma-\alpha-\beta-1)} \\
&\quad\cdot{}_2F_1\left(\gamma+k,\gamma-\alpha-\beta-1;\gamma-\alpha-\beta;-\frac{1}{qb}\right),
\end{aligned}
$$

as $\int_a^\infty \frac{x^{p-1}dx}{(1+\beta x)^\nu}=\frac{a^{p-\nu}}{\beta^\nu(\nu-p)}{}_2F_1(\nu,\nu-p;\nu-p+1;-\frac{1}{\beta a})$ and $a=b,p=\alpha+\beta+k+1,\beta=q,\nu=$

$\gamma + k$.

$$\begin{aligned}
g_{X,S}(s) &= \int_0^s x f(x, s-x) dx \\
&= C(p-q)^{-\gamma} \int_0^s x^{\alpha}(s-x)^{\beta-1} \left(\frac{1+qs}{p-q} + x\right)^{-\gamma} dx \\
&= C(p-q)^{-\gamma} \left(\frac{1+qs}{p-q}\right)^{-\gamma} s^{\alpha+\beta} B\left(\beta, \alpha+1\right) {}_2F_1\left(\gamma, \alpha+1; \beta+\alpha+1; \frac{(q-p)s}{1+qs}\right) \\
&= CB\left(\alpha+1, \beta\right) s^{\alpha+\beta} (1+qs)^{-\gamma} {}_2F_1\left(\gamma, \alpha+1; \alpha+\beta+1; \frac{(q-p)s}{1+qs}\right),
\end{aligned}$$

as $\int_0 \mu x^{\nu-1}(x+\alpha)^{\lambda}(\mu-x)^{\beta-1}dx = \alpha^{\lambda}\mu^{\beta+\nu-1}B(\beta,\nu)_2F_1(-\lambda, \nu; \beta+\nu, -\frac{\mu}{\alpha})$ and $\mu = s, \nu = \alpha+1, \alpha = \frac{1+qs}{p-q}, \lambda = -\gamma, \beta = \beta$.

$$\begin{aligned}
E\left[X_{1_{s>b}}\right] &= \int_b^{\infty} g_{X,S}(s) ds \\
&= C \cdot B\left(\alpha+1, \beta\right) \int_b^{\infty} s^{\alpha+\beta}(1+qs)^{-\gamma} {}_2F_1\left(\gamma, \alpha+1; \alpha+\beta+1; \frac{(q-p)s}{1+qs}\right) ds \\
&= C \cdot B\left(\alpha+1, \beta\right) \sum_{k=0}^{\infty} \frac{(\gamma)_k(\alpha+1)_k(q-p)^k}{(\alpha+\beta+1)_k k!} \int_b^{\infty} s^{\alpha+\beta+k}(1+qs)^{-(\gamma+k)} ds \\
&= C \cdot B\left(\alpha+1, \beta\right) \sum_{k=0}^{\infty} \frac{(\gamma)_k(\alpha+1)_k(q-p)^k}{(\alpha+\beta+1)_k k!} \cdot \frac{b^{\alpha+\beta-\gamma-1}}{q^{\gamma+k}(\gamma-\alpha-\beta-1)} \\
&\quad \cdot {}_2F_1\left(\gamma+k, \gamma-\alpha-\beta-1; \gamma-\alpha-\beta; -\frac{1}{qb}\right)
\end{aligned}$$

Generalized bivariate Pareto distribution

$$f(x,y) = \frac{Cx^{\alpha-1}y^{\beta-1}}{(1+px+qy+rxy)^{\delta}},$$

for $\alpha > 0$, $\beta > 0$, $\gamma > 0$, $\delta > 0$, $p > 0$, $q > 0$, $r > 0$, $x > 0$ and $y > 0$.

$$\begin{aligned}
f(x, s-x) &= \frac{Cx^{\alpha-1}(s-x)^{\beta-1}}{(1+px+q(s-x)+rx(s-x))^{\delta}} \\
&= Cx^{\alpha-1}(s-x)^{\beta-1}\left(1+qs+(p-q+rs)x-rx^2\right)^{-\delta} \\
&= Cx^{\alpha-1}(s-x)^{\beta-1}(x-\mu)^{-\delta}(x-\nu)^{-\delta}(-1)^{-\delta}
\end{aligned}$$

where $\mu = \frac{p-q+rs}{2r} + \sqrt{\frac{1+qs}{r} + \frac{(p-q+rs)^2}{4r^2}}$ and $\nu = \frac{p-q+rs}{2r} - \sqrt{\frac{1+qs}{r} + \frac{(p-q+rs)^2}{4r^2}}$.

$$\begin{aligned} f_S(s) &= \int_0^s f(x, s-x)dx \\ &= C(-1)^{-\delta}\mu^{-\delta}\nu^{-\delta}\int_0^s x^{\alpha-1}(s-x)^{\beta-1}(1-\frac{1}{\mu}x)^{-\delta}(1-\frac{1}{\nu}x)^{-\delta}dx \\ &= C(-1)^{-\delta}\mu^{-\delta}\nu^{-\delta}s^{\alpha+\beta-1}B(\alpha,\beta)F_1\left(\alpha,\delta,\delta,\alpha+\beta;\frac{s}{\mu},\frac{s}{\nu}\right), \end{aligned}$$

as $\int_0^a x^{\alpha-1}(a-x)^{\beta-1}(1-\mu x)^{-p}(1-\nu x)^{-\lambda}dx = a^{\alpha+\beta-1}B(\alpha,\beta)F_1(\alpha,p,\lambda,\alpha+\beta;\mu a,\nu a)$, and $a = s, \alpha = \alpha, \beta = \beta, \mu = \frac{1}{\mu}, \nu = \frac{1}{\nu}, p = \delta, \lambda = \delta$.

$$\begin{aligned} g_{X,S}(s) &= \int_0^s xf(x, s-x)dx \\ &= \int_0^s Cx^{\alpha}(s-x)^{\beta-1}(x-\mu)^{-\delta}(x-\nu)^{-\delta}(-1)^{-\delta}dx \\ &= C(-1)^{-\delta}(\mu\nu)^{-\delta}B(\alpha+1,\beta)s^{\alpha+\beta}F_1\left(\alpha+1,\delta,\delta,\alpha+\beta+1;\frac{s}{\mu},\frac{s}{\nu}\right). \end{aligned}$$

Appendix E

R codes for Chapter 3

```
rm(list=ls())
library(insuranceData)
library(stats4)
data("AutoClaims")
state<-as.character(unique(AutoClaims$STATE))
state<-sort(state)
AC<-AutoClaims

data=NULL
A=NULL
B=NULL
C=NULL
D=NULL
N_male=NULL
N_female=NULL
A.SD <- NULL
B.SD <- NULL
C.SD <- NULL
D.SD <- NULL
```

```
f=function (par1)
{
  tt=1.0e20
  if (par1>0&par2>0) tt=-n*log(par1)-n*par1*log(par2)+(par1+1)*sum(log(x))
  if (is.na(tt)) tt=1.0e20
  if (abs(tt)>1.0e20) tt=1.0e20
  return(tt)}

par2.min=NULL
bootstrap <- function(data,n)
{
  size <- round(length(data)*0.9)
  for(i in 1:n){
  bootres <- sample(data,size,replace = FALSE)
  par2.min <- c(par2.min, min(bootres))
  }
  par2.sd <- sd(par2.min)
  return(par2.sd)
}

for (i in state){
  data<-AC[AC$STATE==i,]
  male <- data[data$GENDER=='M',]
  n <- length(male$PAID)
  x <- male$PAID
  N_male <- c(N_male, n)
  par2<-min(male$PAID)
  est.m <- mle(f,start = list(par1=0.1))
  par1 <- est.m@coef[1][[1]]
```

```
  par1.sd <-summary(est.m)@coef[2]
  par2.sd <- bootstrap(male$PAID,1000)
  B <- c(B, round(par2))
  A <- c(A, par1)
  A.SD <- c(A.SD, par1.sd)
  B.SD <- c(B.SD, par2.sd)
  female <- data[data$GENDER=='F',]
  n <- length(female$PAID)
  x<-female$PAID
  N_female <- c(N_female, n)
  par2<-min(female$PAID)
  est.f <- mle(f,start = list(par1=0.1))
  par1 <- est.f@coef[1][[1]]
  par1.sd <-summary(est.f)@coef[2]
  par2.sd <- bootstrap(female$PAID,1000)
  D <- c(D, round(par2))
  C <- c(C, par1)
  C.SD <- c(C.SD, par1.sd)
  D.SD <- c(D.SD, par2.sd)
}

male_tab <- matrix(c(A,B,N_male,A.SD,B.SD),ncol=5)
rownames(male_tab)<-paste('male',1:13,sep='_')
female_tab <- matrix(c(C,D,N_female, C.SD,D.SD),ncol=5)
rownames(female_tab) <- paste('female',1:13,sep = '_')
result <- rbind(female_tab, male_tab)
colnames(result) <- c(paste('para.',1:2,sep = '_'),'sample_size','para_1.sd','para_2.sd')
output<-result[result[,'sample_size']>10,]
output
write.csv(output,'chapter3_application_table.csv')
```

Appendix F

R codes for Chapter 4

```
rm(list=ls())


install.packages('survival')
install.packages('matlib')


library(survival)
library(rgl)
library(matlib)


data(colon)
colnames(colon)
colcancer<-data.frame(colon[,c('nodes','surg','time')])
colcancer<- colcancer[complete.cases(colcancer),]
colcancer<-colcancer[colcancer$nodes>0,]
colcancer<-colcancer[colcancer$nodes<11,]


nrow(colcancer[colcancer$nodes==0,])
nrow(colcancer[colcancer$nodes>11,])
```

```
node <- as.numeric(unique(colcancer$nodes))
node<-sort(node)

f=function (p)
{
  par1=p[1]
  par2=p[2]
  tt=1.0e20
  if (par1>0&par2>0)
    {tt= -n*log(par1) - par1*n*log(par2) - (par1-1)*sum(log(x))
      + par2^par1*sum(x^par1)}
  if (is.na(tt)) tt=1.0e20
  if (abs(tt)>1.0e20) tt=1.0e20
  return(tt)}


data=NULL
A=NULL
B=NULL
C=NULL
D=NULL
N_short=NULL
N_long=NULL
A.SD <- NULL
B.SD <- NULL
C.SD <- NULL
D.SD <- NULL

for (i in node){
  data<-colcancer[colcancer$nodes==i,]
```

```
  x<-data[data$surg==0, "time"]
  n <- length(x)
  N_short <- c(N_short, n)
  est.s <- optim(fn=f,par=c(1,1),hessian=TRUE)
  par1 <- est.s$par[1]
  par2 <- est.s$par[2]
  mm <- est.s$hessian
  par1.sd <- (Ginv(mm,fractions = TRUE)[1,1])**(1/2)
  par2.sd <- (Ginv(mm,fractions = TRUE)[2,2])**(1/2)
  B <- c(B, par2)
  A <- c(A, par1)
  A.SD <- c(A.SD, par1.sd)
  B.SD <- c(B.SD, par2.sd)
  x<-data[data$surg==1, "time"]
  n <- length(x)
  N_long <- c(N_long, n)
  est.l <- optim(fn=f,par=c(1,1),hessian=TRUE)
  par1 <- est.l$par[1]
  par2 <- est.l$par[2]
  ll<-est.l$hessian
  par1.sd <- (Ginv(ll,fractions = TRUE)[1,1])**(1/2)
  par2.sd <- (Ginv(ll,fractions = TRUE)[2,2])**(1/2)
  D <- c(D, par2)
  C <- c(C, par1)
  C.SD <- c(C.SD, par1.sd)
  D.SD <- c(D.SD, par2.sd)
}

short_tab <- matrix(c(A,B,A.SD,B.SD,N_short),ncol=5)
rownames(short_tab)<-paste('short',1:10,sep='_')
```

```
long_tab <- matrix(c(C,D,C.SD,D.SD,N_long),ncol=5)
rownames(long_tab) <- paste('long',1:10,sep = '_')
result <- rbind(short_tab, long_tab)
colnames(result) <- c(paste('par',1:2,sep = '_'),
                      paste('par.sd.',1:2,sep = '_'),'sample_size')
result

write.csv(result,'chapter4_application_table.csv')
```

Appendix G

R codes for Chapter 5

```
x=c(1,0,1,1,1,0,1,3,1,1,1,3,0,1,1,1,3,0,1,1,1,3,1,1,0,0)

x=3-x

y=c(2,0,1,2,1,1,1,2,1,1,2,3,1,2,1,3,3,1,1,2,0,0,2,1,1,1)

y=3-y

n=26



a=c(1,0,1,0,3,3,1,1,3)

a=3-a

b=c(2,0,1,1,2,3,3,0,0)

b=3-b

obs=c(6,1,8,5,1,2,1,1,1)


ee0=0

eeA=0

eeB=0

eeC=0
```

```
ee1=0
ee2=0
ee3=0
ee4=0
ee5=0
ee6=0
ee7=0
for (j in 1:9)
{xx=a[j]
yy=b[j]



#indep poisson
f0=function (p)
{a1=p[1]
a2=p[2]
tt=1.0e20
if (a1>0&a2>0)
{tt=0
for (i in 1:n)
{tt=tt-log(dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2))}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}

est0=nlm(f0,p=c(1,1))
#cat("Indep Poisson ",est0$minimum,"\n")

a1=est0$estimate[1]
a2=est0$estimate[2]
```

```
e0=n*dpois(xx,lambda=a1)*dpois(yy,lambda=a2)



#model A
fA=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{k=seq(0,min(x[i],y[i]))
tt=tt-log(sum(dpois(x[i]-k,lambda=a1)*dpois(y[i]-k,lambda=a2)*dpois(k,lambda=a3)))}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}

estA=nlm(fA,p=c(1,1,1))
#cat("model A ",estA$minimum,"\n")

a1=estA$estimate[1]
a2=estA$estimate[2]
a3=estA$estimate[3]
k=seq(0,min(xx,yy))
eA=n*sum(dpois(xx-k,lambda=a1)*dpois(yy-k,lambda=a2)*dpois(k,lambda=a3))




#model B
```

```
fB=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{ttt=dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*(1-ppois(max(x[i],y[i]),lambda=a3))
if (y[i]>x[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (y[i]>x[i]) ttt=ttt+dpois(x[i],lambda=a1)*(1-ppois(y[i],lambda=a2))*dpois(y[i],lambda=a3)
if (x[i]>y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (x[i]==y[i]) {ttt=ttt+
                dpois(x[i],lambda=a1)*(1-ppois(y[i],lambda=a2))*dpois(x[i],lambda=a3)}
if (x[i]>y[i]) {ttt=ttt+
                (1-ppois(x[i],lambda=a1))*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)}
if (x[i]==y[i]) {ttt=ttt+
                (1-ppois(x[i],lambda=a1))*dpois(x[i],lambda=a2)*dpois(x[i],lambda=a3)}
if (x[i]==y[i]) {ttt=ttt+
                (1-ppois(x[i],lambda=a1))*(1-ppois(x[i],lambda=a2))*dpois(x[i],lambda=a3)}
tt=tt-log(ttt)}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}

estB=nlm(fB,p=c(1,1,1))
#cat("model B ",estB$minimum,"\n")

a1=estB$estimate[1]
a2=estB$estimate[2]
```

```
a3=estB$estimate[3]
ttt=dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*(1-ppois(max(xx,yy),lambda=a3))
if (yy>xx) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (yy>xx) ttt=ttt+dpois(xx,lambda=a1)*(1-ppois(yy,lambda=a2))*dpois(yy,lambda=a3)
if (xx>yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*(1-ppois(yy,lambda=a2))*dpois(xx,lambda=a3)
if (xx>yy) ttt=ttt+(1-ppois(xx,lambda=a1))*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+(1-ppois(xx,lambda=a1))*dpois(xx,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+(1-ppois(xx,lambda=a1))*(1-ppois(xx,lambda=a2))*dpois(xx,lambda=a3)
eB=n*ttt




#model C
fC=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{ttt=dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*ppois(min(x[i],y[i])-1,lambda=a3)
if (y[i]<x[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (y[i]<x[i]) ttt=ttt+dpois(x[i],lambda=a1)*ppois(y[i]-1,lambda=a2)*dpois(y[i],lambda=a3)
if (x[i]<y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+dpois(x[i],lambda=a1)*ppois(y[i]-1,lambda=a2)*dpois(x[i],lambda=a3)
```

```
if (x[i]<y[i]) ttt=ttt+ppois(x[i]-1,lambda=a1)*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+ppois(x[i]-1,lambda=a1)*dpois(x[i],lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]==y[i]) {ttt=ttt+
                ppois(x[i]-1,lambda=a1)*ppois(x[i]-1,lambda=a2)*dpois(x[i],lambda=a3)}
tt=tt-log(ttt)}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}

estC=nlm(fC,p=c(1,1,1))
#cat("model C ",estC$minimum,"\n")



a1=estC$estimate[1]
a2=estC$estimate[2]
a3=estC$estimate[3]
ttt=dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*ppois(min(xx,yy)-1,lambda=a3)
if (yy<xx) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (yy<xx) ttt=ttt+dpois(xx,lambda=a1)*ppois(yy-1,lambda=a2)*dpois(yy,lambda=a3)
if (xx<yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*ppois(yy-1,lambda=a2)*dpois(xx,lambda=a3)
if (xx<yy) ttt=ttt+ppois(xx-1,lambda=a1)*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+ppois(xx-1,lambda=a1)*dpois(xx,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+ppois(xx-1,lambda=a1)*ppois(xx-1,lambda=a2)*dpois(xx,lambda=a3)
eC=n*ttt

#case 1
f2=function (p)
{a1=p[1]
```

```
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{if (x[i]>0) k=seq(1,x[i],1)
if (x[i]==0&y[i]==0) tt=tt-log(dpois(0,lambda=a1)*dpois(0,lambda=a3))
if (x[i]>0&y[i]==0)
{ttt=dpois(x[i],lambda=a1)*dpois(0,lambda=a3)
ttt=ttt+sum(dpois(x[i]-k,lambda=a1)*dpois(0,lambda=a2)*dpois(k,lambda=a3))
tt=tt-log(ttt)}
if (x[i]>0&y[i]>0)
tt=tt-log(sum(dpois(x[i]-k,lambda=a1)*dpois(y[i]/k,lambda=a2)*dpois(k,lambda=a3)))
}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}


tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f2,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est2=ee

#cat("Case 1 ",est2$minimum,"\n")

a1=est2$estimate[1]
```

```
a2=est2$estimate[2]
a3=est2$estimate[3]
ttt=NA
if (xx>0) k=seq(1,xx,1)
if (xx==0&yy==0) ttt=dpois(0,lambda=a1)*dpois(0,lambda=a3)
if (xx>0&yy==0)
{ttt=dpois(xx,lambda=a1)*dpois(0,lambda=a3)
ttt=ttt+sum(dpois(xx-k,lambda=a1)*dpois(0,lambda=a2)*dpois(k,lambda=a3))}
if (xx>0&yy>0) ttt=sum(dpois(xx-k,lambda=a1)*dpois(yy/k,lambda=a2)*dpois(k,lambda=a3))
e2=n*ttt



#case 4
k=seq(1,100,1)
f11=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{if (x[i]==0&y[i]==0)
{ttt=dpois(0,lambda=a1)*dpois(0,lambda=a2)*(1-dpois(0,lambda=a3))+dpois(0,lambda=a3)
tt=tt-log(ttt)}
if (x[i]>0&y[i]==0)
{ttt=dpois(0,lambda=a2)*sum(dpois(x[i]/k,lambda=a1)*dpois(k,lambda=a3))
tt=tt-log(ttt)}
if (x[i]==0&y[i]>0)
```

```
{ttt=dpois(0,lambda=a1)*sum(dpois(y[i]/k,lambda=a2)*dpois(k,lambda=a3))
tt=tt-log(ttt)}
if (x[i]>0&y[i]>0)
  {tt=tt-log(sum(dpois(x[i]/k,lambda=a1)*dpois(y[i]/k,lambda=a2)*dpois(k,lambda=a3)))}}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}


tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f11,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est11=ee

#cat("Case 4 ",est11$minimum,"\n")


a1=est11$estimate[1]
a2=est11$estimate[2]
a3=est11$estimate[3]
if (xx==0&yy==0)
  {ttt=dpois(0,lambda=a1)*dpois(0,lambda=a2)*(1-dpois(0,lambda=a3))+dpois(0,lambda=a3)}
if (xx>0&yy==0) ttt=dpois(0,lambda=a2)*sum(dpois(xx/k,lambda=a1)*dpois(k,lambda=a3))
if (xx==0&yy>0) ttt=dpois(0,lambda=a1)*sum(dpois(yy/k,lambda=a2)*dpois(k,lambda=a3))
if (xx>0&yy>0) ttt=sum(dpois(xx/k,lambda=a1)*dpois(yy/k,lambda=a2)*dpois(k,lambda=a3))
e11=n*ttt
```

```
#case 5
f5=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{k=seq(1,100,1)
if (x[i]==0&y[i]==0)
{ttt=dpois(0,lambda=a3)+(1-dpois(0,lambda=a3))*dpois(0,lambda=a1)*dpois(0,lambda=a2)
tt=tt-log(ttt)}
if (x[i]>0&y[i]==0)
{ttt=dpois(0,lambda=a2)*sum(dpois(x[i]/k,lambda=a1)*dpois(k,lambda=a3))
tt=tt-log(ttt)}
if (x[i]==0&y[i]>0)
{ttt=dpois(0,lambda=a1)*dpois(y[i],lambda=a2)*(1-ppois(y[i]-1,lambda=a3))
ttt=ttt+dpois(0,lambda=a1)*(1-ppois(y[i],lambda=a2))*dpois(y[i],lambda=a3)
tt=tt-log(ttt)}
if (x[i]>0&y[i]>0)
{k=y[i]+seq(1,100)
ttt=sum(dpois(y[i],lambda=a2)*dpois(k,lambda=a3)*dpois(x[i]/k,lambda=a1))
ttt=ttt+(1-ppois(y[i]-1,lambda=a2))*dpois(y[i],lambda=a3)*dpois(x[i]/y[i],lambda=a1)
tt=tt-log(ttt)}}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}
```

```
tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f5,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est5=ee

#cat("Case 5 ",est5$minimum,"\n")



a1=est5$estimate[1]
a2=est5$estimate[2]
a3=est5$estimate[3]
k=seq(1,100,1)
if (xx==0&yy==0)
  {ttt=dpois(0,lambda=a3)+(1-dpois(0,lambda=a3))*dpois(0,lambda=a1)*dpois(0,lambda=a2)}
if (xx>0&yy==0) ttt=dpois(0,lambda=a2)*sum(dpois(xx/k,lambda=a1)*dpois(k,lambda=a3))
if (xx==0&yy>0)
{ttt=dpois(0,lambda=a1)*dpois(yy,lambda=a2)*(1-ppois(yy-1,lambda=a3))
ttt=ttt+dpois(0,lambda=a1)*(1-ppois(yy,lambda=a2))*dpois(yy,lambda=a3)}
if (xx>0&yy>0)
{k=yy+seq(1,100)
ttt=sum(dpois(yy,lambda=a2)*dpois(k,lambda=a3)*dpois(xx/k,lambda=a1))
ttt=ttt+(1-ppois(yy-1,lambda=a2))*dpois(yy,lambda=a3)*dpois(xx/yy,lambda=a1)}
e5=n*ttt
```

```
#case 6
f6=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{if (x[i]==0&y[i]==0)
{ttt=dpois(0,lambda=a2)*dpois(0,lambda=a3)
tt=tt-log(ttt)}
if (x[i]==0&y[i]>0)
{ttt=dpois(0,lambda=a1)*dpois(y[i],lambda=a2)*(ppois(y[i],lambda=a3)-dpois(0,lambda=a3))
ttt=ttt+dpois(0,lambda=a3)*dpois(y[i],lambda=a2)
ttt=ttt+dpois(0,lambda=a1)*dpois(y[i],lambda=a3)*ppois(y[i]-1,lambda=a2)
tt=tt-log(ttt)}
if (x[i]>0&y[i]>0)
{ttt=0
if (y[i]>1) k=seq(1,(y[i]-1))
if (y[i]>1) ttt=sum(dpois(k,lambda=a3)*dpois(y[i],lambda=a2)*dpois(x[i]/k,lambda=a1))
ttt=ttt+dpois(x[i]/y[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
ttt=ttt+ppois(y[i]-1,lambda=a2)*dpois(y[i],lambda=a3)*dpois(x[i]/y[i],lambda=a1)
tt=tt-log(ttt)}}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}


tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
```

```
{est=nlm(f6,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est6=ee


#cat("Case 6 ",est6$minimum,"\n")



a1=est6$estimate[1]
a2=est6$estimate[2]
a3=est6$estimate[3]
ttt=NA
if (xx==0&yy==0) ttt=dpois(0,lambda=a2)*dpois(0,lambda=a3)
if (xx==0&yy>0)
{ttt=dpois(0,lambda=a1)*dpois(yy,lambda=a2)*(ppois(yy,lambda=a3)-dpois(0,lambda=a3))
ttt=ttt+dpois(0,lambda=a3)*dpois(yy,lambda=a2)
ttt=ttt+dpois(0,lambda=a1)*dpois(yy,lambda=a3)*ppois(yy-1,lambda=a2)}
if (xx>0&yy>0)
{ttt=0
if (yy>1) k=seq(1,(yy-1))
if (yy>1) ttt=sum(dpois(k,lambda=a3)*dpois(yy,lambda=a2)*dpois(xx/k,lambda=a1))
ttt=ttt+dpois(xx/yy,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
ttt=ttt+ppois(yy-1,lambda=a2)*dpois(yy,lambda=a3)*dpois(xx/yy,lambda=a1)}
e6=n*ttt


#case 2
```

```
f2=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{ttt=0
if (x[i]-y[i]-1>0) k=seq(y[i]+1,x[i],1)
if (x[i]-y[i]-1>0)
  {ttt=sum(dpois(x[i]-k,lambda=a1)*dpois(y[i],lambda=a2)*dpois(k,lambda=a3))}
ttt=ttt+dpois(x[i]-y[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
ttt=ttt+dpois(x[i]-y[i],lambda=a1)*dpois(y[i],lambda=a3)*(1-ppois(y[i],lambda=a2))
tt=tt-log(ttt)}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}


tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f2,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est2=ee

#cat("Case 2 ",est4$minimum,"\n")
```

```
a1=est2$estimate[1]
a2=est2$estimate[2]
a3=est2$estimate[3]
ttt=0
if (xx-yy-1>0) k=seq(yy+1,xx,1)
if (xx-yy-1>0) ttt=sum(dpois(xx-k,lambda=a1)*dpois(yy,lambda=a2)*dpois(k,lambda=a3))
ttt=ttt+dpois(xx-yy,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
ttt=ttt+dpois(xx-yy,lambda=a1)*dpois(yy,lambda=a3)*(1-ppois(yy,lambda=a2))
e2=n*ttt


#case 7
f7=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{ttt=0
if (x[i]+1<=y[i]-1)
  {ttt=dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)
    *(ppois(y[i]-1,lambda=a3)-ppois(x[i]+1,lambda=a3))}
if (y[i]>x[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (y[i]>x[i]) ttt=ttt+dpois(x[i],lambda=a1)*ppois(y[i]-1,lambda=a2)*dpois(y[i],lambda=a3)
if (x[i]<y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+dpois(x[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
if (x[i]==y[i]) ttt=ttt+dpois(x[i],lambda=a1)*ppois(y[i]-1,lambda=a2)*dpois(x[i],lambda=a3)
if (x[i]<y[i])
```

```
  {ttt=ttt+(1-ppois(x[i],lambda=a1))*dpois(y[i],lambda=a2)*dpois(x[i],lambda=a3)}
if (x[i]==y[i])
  {ttt=ttt+(1-ppois(x[i],lambda=a1))*dpois(x[i],lambda=a2)*dpois(x[i],lambda=a3)}
if (x[i]==y[i])
  {ttt=ttt+(1-ppois(x[i],lambda=a1))*ppois(x[i]-1,lambda=a2)*dpois(x[i],lambda=a3)}
tt=tt-log(ttt)}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}


tt=1.0e20
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f7,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est7=ee

#cat("Case 7 ",est7$minimum,"\n")


a1=est7$estimate[1]
a2=est7$estimate[2]
a3=est7$estimate[3]
ttt=NA
if (xx+1<=yy-1)
  {ttt=dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*(ppois(yy-1,lambda=a3)-ppois(xx+1,lambda=a3))}
if (yy>xx) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (yy>xx) ttt=ttt+dpois(xx,lambda=a1)*ppois(yy-1,lambda=a2)*dpois(yy,lambda=a3)
if (xx<yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
```

```
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
if (xx==yy) ttt=ttt+dpois(xx,lambda=a1)*ppois(yy-1,lambda=a2)*dpois(xx,lambda=a3)
if (xx<yy) ttt=ttt+(1-ppois(xx,lambda=a1))*dpois(yy,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+(1-ppois(xx,lambda=a1))*dpois(xx,lambda=a2)*dpois(xx,lambda=a3)
if (xx==yy) ttt=ttt+(1-ppois(xx,lambda=a1))*ppois(xx-1,lambda=a2)*dpois(xx,lambda=a3)
e7=n*ttt



#case 3
f3=function (p)
{a1=p[1]
a2=p[2]
a3=p[3]
tt=1.0e20
if (a1>0&a2>0&a3>0)
{tt=0
for (i in 1:n)
{ttt=0
if (min(x[i],y[i]-1)>=0) k=seq(0,min(x[i],y[i]-1),1)
if (min(x[i],y[i]-1)>=0)
  {ttt=sum(dpois(x[i]-k,lambda=a1)*dpois(yy,lambda=a2)*dpois(k,lambda=a3))}
ttt=ttt+dpois(x[i]-y[i],lambda=a1)*dpois(y[i],lambda=a2)*dpois(y[i],lambda=a3)
ttt=ttt+dpois(x[i]-y[i],lambda=a1)*dpois(y[i],lambda=a3)*ppois(y[i]-1,lambda=a2)
tt=tt-log(ttt)}}
if (is.na(tt)||abs(tt)>1.0e20) tt=1.0e20
return(tt)}



tt=1.0e20
```

```
for (i1 in 1:5) for (i2 in 1:5) for (i3 in 1:5)
{est=nlm(f3,p=c(i1,i2,i3))
if (est$minimum<tt)
{tt=est$minimum
ee=est}}
est3=ee

#cat("Case 3 ",est3$minimum,"\n")



a1=est3$estimate[1]
a2=est3$estimate[2]
a3=est3$estimate[3]
ttt=0
if (min(xx,yy-1)>=0) k=seq(0,min(xx,yy-1),1)
if (min(xx,yy-1)>=0) ttt=sum(dpois(xx-k,lambda=a1)*dpois(yy,lambda=a2)*dpois(k,lambda=a3))
ttt=ttt+dpois(xx-yy,lambda=a1)*dpois(yy,lambda=a2)*dpois(yy,lambda=a3)
ttt=ttt+dpois(xx-yy,lambda=a1)*dpois(yy,lambda=a3)*ppois(yy-1,lambda=a2)
e3=n*ttt



cat(obs[j]," & ",formatC(e0,digits=1,format="f")," & ",
formatC(eA,digits=1,format="f")," & ",
formatC(eB,digits=1,format="f")," & ",
formatC(eC,digits=1,format="f")," & ",
formatC(e1,digits=1,format="f")," & ",
formatC(e2,digits=1,format="f")," & ",
formatC(e3,digits=1,format="f")," & ",
#formatC(e4,digits=1,format="f")," & ",
formatC(e5,digits=1,format="f")," & ",
```

```
formatC(e6,digits=1,format="f")," & ",
formatC(e7,digits=1,format="f"),
"\\\\","\n")

if (!is.na(e0)) ee0=ee0+(obs[j]-e0)**2/e0
if (!is.na(eA)) eeA=eeA+(obs[j]-eA)**2/eA
if (!is.na(eB)) eeB=eeB+(obs[j]-eB)**2/eB
if (!is.na(eC)) eeC=eeC+(obs[j]-eC)**2/eC
if (!is.na(e1)) ee1=ee1+(obs[j]-e1)**2/e1
if (!is.na(e2)) ee2=ee2+(obs[j]-e2)**2/e2
if (!is.na(e3)) ee3=ee3+(obs[j]-e3)**2/e3
#if (!is.na(e4)) ee4=ee4+(obs[j]-e4)**2/e4
if (!is.na(e5)) ee5=ee5+(obs[j]-e5)**2/e5
if (!is.na(e6)) ee6=ee6+(obs[j]-e6)**2/e6
if (!is.na(e7)) ee7=ee7+(obs[j]-e7)**2/e7
}

cat(formatC(ee0,digits=1,format="f")," & ",
formatC(eeA,digits=1,format="f")," & ",
formatC(eeB,digits=1,format="f")," & ",
formatC(eeC,digits=1,format="f")," & ",
formatC(ee1,digits=1,format="f")," & ",
formatC(ee2,digits=1,format="f")," & ",
formatC(ee3,digits=1,format="f")," & ",
#formatC(ee4,digits=1,format="f")," & ",
formatC(ee5,digits=1,format="f")," & ",
formatC(ee6,digits=1,format="f")," & ",
formatC(ee7,digits=1,format="f"),
"\\\\","\n")
```