Tail dependence convergence rate of a skew-t and of a skew normal distribution

Thomas Fung*
Macquarie University, Sydney, Australia - thomas.fung@mq.edu.au
Honorary Associate, University of Sydney, Sydney, Australia

Eugene Seneta
University of Sydney, Sydney, Australia - eseneta@maths.usyd.edu.au

Abstract

We first examine the rate of decay to the limit of the lower tail dependence function i.e. the asymptotic tail dependence coefficient of a bivariate skew-t distribution. It is important to consider the correction term as the tail dependence function can be much different from its limit. We find that the rate is asymptotically a power-law. The results contain as a special case the usual bivariate symmetric t distribution, and hence the skew t distribution we consider here is an appropriate (skew) extension. We then discuss briefly the rate of convergence for the skew normal distribution under an equal-skewness condition.

Keywords: asymptotic tail dependence coefficient; power law; tail order.

1. Introduction

The coefficient of lower tail dependence of a random vector $\mathbf{X} = (X_1, X_2)^T$, the marginal distribution functions, F_1, F_2, of which are assumed continuous and strictly increasing, is defined as

$$\lambda_L = \lim_{u \to 0^+} \lambda_L(u), \quad \text{where} \quad \lambda_L(u) = P(X_1 \le F_1^{-1}(u) | X_2 \le F_2^{-1}(u)), \tag{1}$$

where F_1^{-1} and F_2^{-1} denote the marginal inverse distribution functions. The coefficient of asymptotic upper tail dependence of a random vector $\mathbf{X}$ is defined similarly as

$$\lambda_U = \lim_{u \to 1^-} \lambda_U(u), \quad \text{where} \quad \lambda_U(u) = P(X_1 \ge F_1^{-1}(u) | X_2 \ge F_2^{-1}(u)). \tag{2}$$

$\mathbf{X}$ is said to have asymptotic lower (resp. upper) tail dependence if λ_L (resp. λ_U) exists and is positive. If $\lambda_L = 0$ (resp. $\lambda_U = 0$) then $\mathbf{X}$ is said to be asymptotically independent in the lower (resp. upper) tail. These quantities provide insight on the tendency for the distribution to generate joint extreme event since they measure the strength of dependence (or association) in the tails of a bivariate distribution. It follows that $\lambda_L(u)$ can be expressed in terms of the copula of $\mathbf{X}$, $C(u_1, u_2)$, as

$$\lambda_L(u) = \frac{P(X_1 \le F_1^{-1}(u), X_2 \le F_2^{-1}(u))}{P(X_2 \le F_2^{-1}(u))} = \frac{C(u,u)}{u}, \tag{3}$$

and, similarly, the copula-based form of $\lambda_U(u)$ is

$$\lambda_U(u) = \frac{1 - 2u + C(u,u)}{1-u}. \tag{4}$$

The motivation for our investigation of the rate of convergence to λ_L for bivariate distributions initially arose from the following. Ramos and Ledford (2009) discussed a methodology initiated by Ledford and Tawn (1997) and studied intensively a family of bivariate distributions (which they characterised) which satisfied in particular the condition

$$\lambda_L(u) = u^{\frac{1}{\alpha}-1} L(u). \tag{5}$$

Here $L(u)$ is a slowly varying function as $u \to 0^+$, and $\alpha \in (0, 1]$, so that, in fact, the value of α could be used for comparison of the degree of tail dependence structure between members of the family.

Hua and Joe (2011) developed this idea further and defined $\kappa = 1/\alpha$ in (5) as the (lower) tail order of a copula. The case $1 < \kappa < 2$ is termed as intermediate tail dependence as it represents the copula has some level of positive dependence in the tail but not as strong as tail dependence with $\lambda_L = 0$. The tail order κ can be used to assess tail dependence strength when $\lambda_L = 0$.

Expression (5) may be regarded as the rate of convergence of $\lambda_L(u)$ to λ_L when $\lambda_L = 0$. In this situation the right-hand side of (5) is a first order correction to the limit value. When α is close to 1, there is a substantial difference between $\lambda_L(u)$ and $\lambda_L = 0$, as the convergence is slow. And indeed, in practice it may be difficult to ascertain by computation of $\lambda_L(u)$ for ever smaller u whether in fact the limit is zero, and we have to resort to analytic proof as in Fung and Seneta (2011a). In the case of the bivariate normal with correlation coefficient ρ, $\alpha = \frac{1+\rho}{2}$ (see for example Fung and Seneta (2011b)), so convergence is slow when ρ is close to $+1$, which is consistent with strong positive correlation between variables.

On the other hand, when $\lambda_L(u) \to \lambda_L > 0$, which is also covered by (5) with $\alpha = 1$ and $L(u) \to \lambda_L$, the rate of convergence is more appropriately studied by considering the rate of convergence to 0 as $u \to 0^+$ of $|\lambda_L(u) - \lambda_L|$ i.e.

$$|\lambda_L(u) - \lambda_L| = u^\kappa L(u). \tag{6}$$

As a background example on account of its simplicity and ubiquity we consider the bivariate copula $C_A(u_1, u_2)$ given by

$$C_A(u_1, u_2) = \exp(A \log(u_1 u_2)), \tag{7}$$

where $A \in [1/2, 1]$ is a constant. This is an example of an extreme value copula in the sense of Joe (1997), p.140, and is taken from Capéraà *et al.* (2000). Consequently, $C_A(u, u) = u^{2A}$, whence, with self-explanatory notation

$$\lambda_L^A(u) = u^{2A-1}, \;\; \lambda_U^A(u) = (2A-1)(1 + A(1-u) + O((1-u)^2) \text{ as } u \to 1^-. \tag{8}$$

Thus both asymptotic tail independence, as an example of (5), and asymptotic tail dependence, with power-law rate of approach to the positive limit in the case of asymptotic tail dependence, are displayed.

The first object of our study is the bivariate skew-t distribution and the second is the skew normal distribution. A random vector $\mathbf{Z}$ is said to have a bivariate skew normal distribution (see Azzalini and Dalla Valle (1996)), denoted as $\mathbf{Z} \sim SN_2(\boldsymbol{\theta}, R)$, if it has density

$$f(\mathbf{z}) = 2\phi_2(\mathbf{z}, R)\Phi(\boldsymbol{\theta}^T \mathbf{z}) \tag{9}$$

where $\phi_2(\cdot, R)$ is the bivariate normal density with mean $\mathbf{0}$ and correlation matrix $R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\Phi(\cdot)$ is the cdf of $N(0,1)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$ is a vector that controls the asymmetry of the distribution. We then define the bivariate skew-t as that distribution resulting from variance-mixing of the bivariate skew normal, inversely with a gamma random variable $V \sim \Gamma(\frac{\eta}{2}, \frac{\eta}{2})$, with $\eta > 0$:

$$\mathbf{X} = V^{-\frac{1}{2}} \mathbf{Z}, \tag{10}$$

where $\mathbf{Z}$ is independently distributed of V. This skew distribution was originally introduced in multivariate form in Branco and Dey (2001) and studied extensively in Azzalini and Capitanio (2003). Some recent reviews on this area of study can be found in Azzalini and Genton (2008), Azzalini and Capitanio (2010) and in the book edited by Genton (2004). The case $\theta_1 = \theta_2 = 0$ reduces to the symmetric bivariate t distribution. In this sense, the bivariate skew-t distribution defined by (10) is a generalisation of the symmetric case.

The bivariate skew-t always satisfies $\lambda_L > 0$ (See Fung and Seneta (2010)). This was also considered in Bortot (2010) and Padoan (2011) with an approach initiated by Cheng and Genton (2007) which is quite different from that of Fung and Seneta (2010).

Manner and Segers (2011) and Chicheportiche and Bouchaud (2012) considered (6) for the symmetric t case (i.e. $\boldsymbol{\theta} = \mathbf{0}$). They found that

$$\lambda_L(u) = \lambda_L + \gamma u^{\frac{2}{\eta}} + O(u^{\frac{4}{\eta}}), \tag{11}$$

where

$$\gamma = \frac{2}{(2/\eta + 1)} f_{t_{\eta+1}}\left(-\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}}\right)\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}}\,\frac{\eta}{2}\left(\frac{\sqrt{\pi}\,\Gamma(\frac{\eta}{2})}{\Gamma(\frac{\eta+1}{2})\eta^{\frac{\eta}{2}-1}}\right)^{\frac{2}{\eta}} \tag{12}$$

with $f_{t_{\eta+1}}(\cdot)$ as the density of the (symmetric) t distribution with $\eta+1$ degrees of freedom. The authors argue that at any non-zero value of u, $\lambda_L(u)$ may be much different from its limit, so it is important to consider the correction term, in the same degree-of-approximation sense as the right-hand side of (5) when $\lambda_L = 0$. These two papers, provides a second motivation for us to consider the rate of convergence aspect, specifically the first and second order correction terms, as a generalisation, for the more general skew-t distribution.
The remainder of this paper is set out as follows: In Section 2, we summarise the rate of convergence in the form of (6) for the skew-t distribution. In Section 3, we discuss some notable features of our results. In Section 4, result for the the skew normal distribution distribution in the form of (5) under an equal-skewness condition is discussed.

2. The skew-t distribution

The starting point of our method is to apply L'Hôpital's rule to (3), we have

$$\lambda_L = \lim_{q\to 0^+} \frac{C(q,q)}{q} = \lim_{q\to 0^+} \frac{dC(q,q)}{dq}. \tag{13}$$

After some basic properties of the derivative and conditional distribution of a copula C (See Nelsen (2006) pp 13, 41), we have

$$\lambda_L = \lim_{u\to 0^+} \lambda_L(u) = \lim_{y\to -\infty} [P(X_2 \le F_2^{-1}(F_1(y))|X_1 = y) + P(X_1 \le F_1^{-1}(F_2(y))|X_2 = y)].$$

Fung and Seneta (2010) showed that if $\mathbf{X} = (X_1, X_2)^T$ is a random vector defined by (10), then

$$\lim_{y\to-\infty} P(X_2 \le F_2^{-1}(F_1(y))|X_1 = y) = \int_{-\infty}^{-a_{2.1}} f_{t_{\eta+1}}(z) \frac{F_{t_{\eta+2}}\left(\left(\theta_2\sqrt{\frac{(1-\rho^2)}{\eta+1}}z - (\theta_1 + \rho\theta_2)\right)\sqrt{\frac{\eta+2}{1+\frac{z^2}{\eta+1}}}\right)}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})}\, dz, \tag{14}$$

where $a_{2.1} = \left(\left(\frac{F_{t_{\eta+1}}(-\lambda_2\sqrt{\eta+1})}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})}\right)^{\frac{1}{\eta}} - \rho\right)\sqrt{\frac{\eta+1}{1-\rho^2}}$; and the result is similar for $\lim_{y\to-\infty} P(X_1 \le F_1^{-1}(F_2(y))|X_2 = y)$ by interchanging the index 1 and 2 in (14).
Our main result is summarised into the following theorem. The proof is omitted here but can be found in Fung and Seneta (2014).

Theorem 1 *For the bivariate skew-t distribution:*

$$\lambda_L(u) = \lambda_L + \mathcal{K}(\eta, R, \boldsymbol{\theta})u^{\frac{2}{\eta}} + O(u^{\frac{4}{\eta}}) \tag{15}$$

as $u \to 0^+$*, where* $\mathcal{K}(\eta, R, \boldsymbol{\theta})$ *is a constant and is defined as*

$$\mathcal{K}(\eta, R, \boldsymbol{\theta}) = \frac{k^*_{2.1} + k^*_{1.2}}{2/\eta + 1} \tag{16}$$

with

$$k^*_{2.1} = \left(\frac{(\pi\eta)^{\frac{1}{2}}\Gamma(\frac{\eta}{2})\eta}{2\Gamma(\frac{\eta+1}{2})\eta^{\frac{\eta+1}{2}} F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})}\right)^{\frac{2}{\eta}} \Bigg\{ -\int_{-\infty}^{L_1} f_{t_{\eta+1}}(z) \Big\{ \frac{f_{t_{\eta+2}}(a(z)+b(z))b(z)\frac{\eta}{2}}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})} + \lambda_1\sqrt{\eta+1}\frac{\eta}{2}$$
$$+ \frac{f_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})F_{t_{\eta+2}}(a(z)+b(z))}{(F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1}))^2} \Big\} dz + \frac{1}{(\frac{1-\rho^2}{\eta+1})^{\frac{1}{2}}} \Big\{ \left(\frac{F_{t_{\eta+1}}(-\lambda_2\sqrt{\eta+1})}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})}\right)^{\frac{1}{\eta}} \left(\frac{d_1 - d_2(\frac{c_1}{c_2})^{\frac{2}{\eta}}}{\eta}\right)$$
$$+ \frac{\eta}{2}\Big(\left(\frac{F_{t_{\eta+1}}(-\lambda_2\sqrt{\eta+1})}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})}\right)^{\frac{1}{\eta}} - \rho\Big) \Big\} \times f_{t_{\eta+1}}(L_1) \frac{F_{t_{\eta+2}}(a(L_1)+b(L_1))}{F_{t_{\eta+1}}(-\lambda_1\sqrt{\eta+1})} \Bigg\}.$$

and $k^*_{1.2}$ *is defined similarly.*

Here the constant $\mathcal{K}(\eta, R, \boldsymbol{\theta})$ is a function of all the parameters, $\eta, \rho_1, \rho_2, \theta_1, \theta_2$, even though, surprisingly, the convergence rate $u^{\frac{2}{\eta}}$, is, still, determined solely by η. A further feature is that in both (8) and (15), when the coefficient of asymptotic tail dependence is positive, the convergence rate is strictly power-law.

3. Discussion and Simplifications

1. **Goodness of approximation** We have plotted the exact $\lambda_L(u)$ and $\lambda_U(u)$, and λ_L, λ_U plus their $2/\eta-$power correction against u, in Figures 1 and 2. As for the plots for (11) in the papers cited for the symmetric case, we see that the correction term has significant effect as soon as u or $1-u$, respectively, become small. For example, for $\lambda_L(u)$ in Figure 1, we have $k^*_{2.1} + k^*_{1.2} = 0.246$ and the first order correction is 0.0279 at $u = 0.01$ while $\lambda_L = 0.1071$. Thus the correction can be quite substantial. Overall, the convergence rate correction is reasonably good and it gets more accurate as η becomes smaller.

2. **Simplification 1.** Next, if $\theta_1 = \theta_2 = 0$ (i.e. the symmetric t case), then $\mathcal{K}(\eta, R, \boldsymbol{\theta})$ in (16) reduces to γ in (12) as $L_1 = -\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}}$ and $a(z) = b(z) = \lambda_1 = \lambda_2 = d_1 - d_2(\frac{c_1}{c_2})^{\frac{2}{\eta}} = 0$. This consistency further supports the proposal that (10) is a proper skew extension to the symmetric multivariate t distribution.

3. As $\eta \to \infty$, the skew-t approaches the skew-normal, for which the asymptotic tail dependence coefficients λ_L, λ_U are both zero (Lysenko, Roy and Waeber (2009) and Bortot (2010)). Although some simplification of the constant term $\mathcal{K}(\eta, R, \boldsymbol{\theta})$ is possible when $\theta_1 = \theta_2$, even in the case $\boldsymbol{\theta} = 0$, the explicit expression (12) is complex. Recall that this is the case of a bivariate normal, where in fact the convergence rate is a regularly varying function. On the other hand from (12), we have

$$
\begin{aligned}
\mathcal{K}(\eta, R, \mathbf{0}) &= \frac{2}{(2/\eta+1)} f_{t_{\eta+1}}\Big(-\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}}\Big)\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}}\,\frac{\eta}{2}\left(\frac{\sqrt{\pi}\Gamma(\frac{\eta}{2})}{\Gamma(\frac{\eta+1}{2})\eta^{\frac{\eta}{2}-1}}\right)^{\frac{2}{\eta}} \\
&\sim \frac{1}{\sqrt{2\pi}}\left(\frac{2}{1+\rho}\right)^{-(\frac{\eta+2}{2})}\sqrt{\frac{(\eta+1)(1-\rho)}{1+\rho}} \to 0
\end{aligned}
$$

as $\eta \to \infty$, so the correction term in the limit does not provide useful information about the correction term of the limit even when $\theta_1 = \theta_2 = 0$.

4. The skew normal distribution under an equal-skewness condition

While we were able to treat in a unified way the skew-t with θ_1 not necessarily equal to θ_2, the skew normal requires an essentially different approach even in the equal-skewness case $\theta_1 = \theta_2 = \theta$. The results for the tail dependence convergence rate are summarised into the following theorem.

Theorem 2 *Let $\boldsymbol{X} \sim SN_2(\boldsymbol{\theta}, R)$ with $\theta_1 = \theta_2 = \theta$. As $u \to 0^+$,*

(a) if $\theta > 0$,

$$
\lambda_L(u) \sim u^{\beta^2} \frac{\alpha^3}{\pi\lambda^4\beta(1+\beta^2)^2}\sqrt{\frac{2}{\pi}}(2\pi\lambda)^{1+\beta^2}\Big(\frac{1+\lambda^2}{2}\Big)^{\frac{3}{2}}\left[-\log u\right]^{\beta^2-\frac{1}{2}} \tag{17}
$$

with $\lambda = \frac{\theta(1+\rho)}{\sqrt{1+\theta^2(1-\rho^2)}}$, $\alpha = \frac{\theta(1+\rho)}{\sqrt{1+2\theta^2(1+\rho)}}$ and $\beta = \sqrt{\frac{(1-\rho)(1+2\theta^2(1+\rho))}{1+\rho}}$;

(b) if $\theta < 0$,

$$
\lambda_L(u) \sim u^{\frac{1-\rho}{1+\rho}} \times \frac{1+\rho}{2}\sqrt{\frac{1+\rho}{1-\rho}}(-\pi\log u)^{-\frac{\rho}{1+\rho}}. \tag{18}
$$

The theorem shows that when $\theta < 0$ there is minimal difference between symmetric and skew normal in terms of the intermediate tail dependence as they share the same regular varying index which is $\frac{1-\rho}{1+\rho}$. On the other hand, when $\theta > 0$, there is a larger regular varying index by a factor of $(1+2\theta^2(1+\rho))$ when compared to the normal case and therefore skew normal has smaller intermediate tail dependence than the normal in the lower tail.

References

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J.R. Statist. Soc. B.*, **65**, 367–389.

Azzalini, A. and Capitanio, A. (eds.) (2010). *Skew-symmetric and flexible distributions.* Special issue of *Metron*, **68**.

Azzalini, A. and Genton, M.G. (2008). Robust likelihood methods based on the skew-t and related distribution. *International Statistical Review*, **76**, 106–129.

Azzalini, A. and Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**, 715–726.

Bortot, P. (2010). Tail dependence in bivariate skew-normal and skew-t distribution. http://www2.stat.unibo.it/bortot/ricerca/paper-sn-2.pdf.

Branco, M.D. and Dey, D.K. (2001). A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, **79**, 99–113.

Capéraà, P., Fougères, A.-L. and Genest, C. (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, **72**, 30–49.

Chang, S.-M. and Genton, M.G. (2007). Extreme value distributions for the skew-symmetric family of distributions. *Communications in Statistics - Theory and Methods*, **36**, 1705–1717.

Chicheportiche, R. and Bouchaud, J.-P. (2012). The joint distribution of stock returns is not elliptical. *International Journal of Theoretical and Applied Finance*, **15**, 1250019 (23 pages).

Fung, T. and Seneta, E. (2010). Tail dependence for two skew t distributions. *Statistics and Probability Letters*, **80**, 784–791.

Fung, T. and Seneta, E. (2011a). Tail dependence and skew distributions. *Quantitative Finance*, **11**, 327–333.

Fung, T. and Seneta, E. (2011b). The bivariate normal copula function is regularly varying. *Statistics and Probability Letters*, **81**, 1670–1676.

Fung, T. and Seneta, E. (2014). Convergence rate to a lower tail dependence coefficient of a skew-t distribution. *Journal of Multivariate Analysis*, **128**, 62–72.

Genton, M.G.(ed.) (2004). *Skew-Elliptical Distributions and Their Applications*, Boca Raton, FL: Chapman & Hall/CRC.

Hua, L. and Joe, H. (2011). Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis* **102**, 1454–1471.

Joe, H.(1997) *Multivariate Models and Dependence Concepts.* London: Chapman and Hall.

Ledford, A.W. and Tawn, J.A. (1997). Modelling dependence with joint tail regions. *J. R. Statist. Soc. B*, **59**, 475–499.

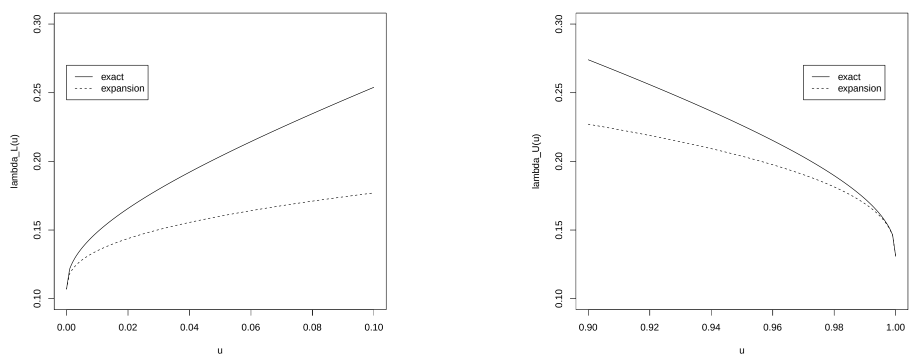
Lysenki, N., Roy, P. and Waeber, R. (2009). Multivariate extremes of generalised skew-normal distributions. *Statistics and Probability Letters*, **79**, 525–533.

Manner, H. and Segers, J. (2011). Tails of correlation mixtures of elliptical copulas. *Insurance: Mathematics and Economics*, **48**, 153–160.

Nelsen, R.B. (2006). *An Introduction to Copulas.* 2nd ed. New York: Springer.

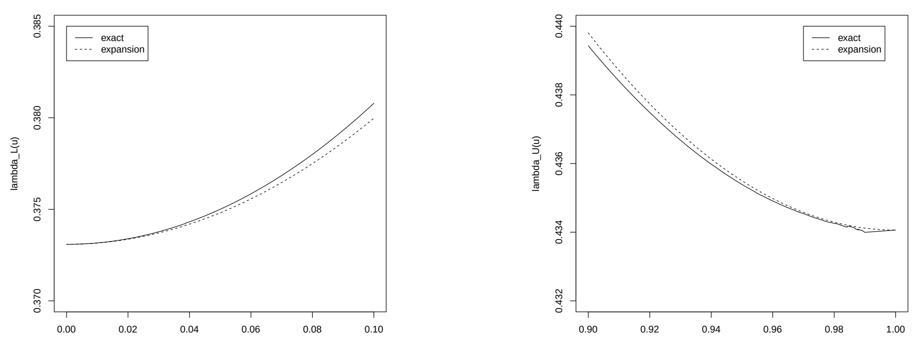
Padoan, S.A. (2011). Multivariate extreme models based on underlying skew t and skew-normal distributions. *Journal of Multivariate Analysis*, **102**, 977–991.

Ramos, A. and Ledford, A. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc. B*, **71**, 219–241.



(a) (b)

Figure 1: Exact and approximation with correction term for $\lambda_L(u)$ and $\lambda_U(u)$ against u for the skew t distribution with $\rho = 0.3$, $\boldsymbol{\theta} = (0.1, 0.3)^T$ and $\eta = 5$.



(a) (b)

Figure 2: Exact and approximation with correction term for $\lambda_L(u)$ and $\lambda_U(u)$ against u for the skew t distribution with $\rho = 0.3$, $\boldsymbol{\theta} = (0.1, 0.3)^T$ and $\eta = 1$.