Tail dependence in bivariate skew-Normal and skew-$t$ distributions

Paola Bortot
Department of Statistical Sciences - University of Bologna
paola.bortot@unibo.it

Abstract: Quantifying dependence between extreme values is a central problem in many theoretical and applied studies. The main distinction is between asymptotically independent and asymptotically dependent extremes, with important theoretical examples of these general limiting classes being the extremal behaviour of a bivariate Normal distribution, for asymptotic independence, and of the bivariate $t$ distribution, for asymptotic dependence. In this paper we study the tail dependence of skewed extensions of these two basic models, namely the bivariate skew-Normal and skew-$t$ distributions. We show that both distributions belong to the same limiting class as the generating family, the skew-Normal being asymptotically independent and the skew-$t$ being asymptotically dependent. However, within their respective limiting class, each provides a wider range of extremal dependence strength than the generating distribution. In addition, both the skew-Normal and the skew-$t$ distributions allow different upper and lower tail dependence.

Keywords: Asymptotic dependence, Asymptotic independence, Coefficient of tail dependence, Extreme values, Skew-Normal distribution, Skew-$t$ distribution.

1 Introduction

Quantifying extremal dependence is a central problem of multivariate extreme value theory and its statistical applications. Two forms of extremal dependence are possible: extremes are either asymptotically independent or asymptotically dependent. Theoretical examples of the two limiting behaviours are the bivariate Normal distribution, for asymptotic independence, and the bivariate $t$ distribution, for asymptotic dependence (see, for example, Ledford and Tawn, 1996, and Demarta and McNeil, 2005). In this work, we focus on the tail dependence of skewed extensions of these basic models, namely the bivariate skew-Normal and skew-$t$ distributions (Azzalini and Capitanio, 1999 and 2003).

In recent years, the skew-Normal and the skew-$t$ distributions, in their univariate and multivariate versions, have received considerable attention both in theoretical studies, for their numerous stochastic properties, and in applied studies, for the additional flexibility that they provide for modelling phenomena that depart from symmetry. For a comprehensive review see Azzalini (2005). Since these families have become standard tools in many areas of applications, including selective sampling, stochastic frontier and financial studies, it is of interest to
investigate the type of tail behaviour which is implied by their use. As a practical example, whenever multivariate financial return data are modelled through the skew-Normal distribution or the skew-t distribution (Adcock, 2004; Walls, 2005), it is crucial to know whether the adopted model admits the occurrence of simultaneous large losses (or large gains), i.e. whether returns are treated as asymptotically dependent or asymptotically independent. In addition, if the two distributions demonstrate to cover a wide range of asymptotic behaviours, they may provide useful models for extremal dependence as alternatives to those presently available in the literature (Ledford and Tawn, 1997; Bortot et al., 2000; Heffernan and Tawn, 2004).

We will show that both the skew-Normal distribution and the skew-t distribution inherit the asymptotic behaviour of the generating distribution, the skew-Normal being asymptotically independent and the skew-t being asymptotically dependent. However, within their respective limiting class, each provides a wider range of extremal dependence strength than the generating family. Moreover, both the skew-Normal and the skew-t distributions are shown to allow different upper and lower tail dependence.

The structure of the paper is as follows. Section 2 gives an overview of tail dependence measures proposed in the literature for asymptotically dependent and independent distributions. In Sections 3 and 4 respectively, the asymptotic characteristics of the bivariate skew-Normal distribution and of the bivariate skew-t distribution are illustrated. Technical proofs of the stated results are postponed to the Appendix. Finally, Section 5 contains some concluding remarks and ideas for future work.

2 Measures of asymptotic dependence

Let \((X_1, X_2)\) be a bivariate random variable with univariate marginal distribution functions \(\Pr(X_1 \leq x) = F_1(x)\) and \(\Pr(X_2 \leq x) = F_2(x)\).

Define

\[
\chi_U = \lim_{u \to 1} \chi_U(u),
\]

where

\[
\chi_U(u) = \Pr(F_1(X_1) > u | F_2(X_2) > u).
\]

The quantity \(\chi_U\) is known as the coefficient of upper tail dependence. It provides a measure of the asymptotic dependence in the upper tail of the distribution of \((X_1, X_2)\) by, loosely speaking, giving the probability that one variable is large given that the other variable is large. If \(\chi_U = 0\), then \(X_1\) and \(X_2\) are said to be asymptotically independent (AI) in the upper tail; if \(\chi_U > 0\), \(X_1\) and \(X_2\) are said to be asymptotically dependent (AD) in the upper tail, with larger values of \(\chi_U\) denoting stronger asymptotic dependence.

Definitions (1) and (2) for the upper tail can be extended to the lower tail by replacing \(\chi_U\) with its lower tail analogue

\[
\chi_L = \lim_{u \to 0} \chi_L(u),
\]

where

\[
\chi_L(u) = \Pr(F_1(X_1) < u | F_2(X_2) < u).
\]
Important examples of AI and AD distributions are, respectively, the bivariate Normal and the bivariate \( t \) distributions. Given their symmetry, \( \chi_U = \chi_L = \chi \) for both distributions, with \( \chi = 0 \), for the bivariate Normal variable with correlation coefficient \(|\rho| < 1\), while for the bivariate \( t \) distribution with \( \nu > 0 \) degrees of freedom and off-diagonal element of the correlation matrix equal to \( \rho > -1 \)

\[
\chi = 2T_1 \left( -\sqrt{(\nu + 1)(1 - \rho)}/\sqrt{1 + \rho}; \nu + 1 \right),
\]

where \( T_1(\cdot; \nu) \) denotes the distribution function of a univariate standard \( t \) variable with \( \nu \) degrees of freedom (Demarta and McNeil, 2005). Asymptotic independence for the bivariate \( t \) is obtained as the limiting case of \( \nu \to \infty \).

The importance of AI distributions has been highlighted by various applications of multivariate extreme value modelling: especially in the analysis of environmental extremes, empirical evidence often suggests the presence of asymptotic independence. Coles et al. (1999) point out that, while for all distributions with \( \chi_U = 0 \) or \( \chi_L = 0 \) there is independence in the limit, at sub-asymptotic levels different degrees of dependence might be attainable. In other words, the speed of convergence of \( \chi_U(u) \) or \( \chi_L(u) \) to 0 may vary from one AI distribution to another. To measure the degree of sub-asymptotic dependence in the upper tail Coles et al. (1999) propose the index

\[
\tilde{\chi}_U = \lim_{u \to 1} \frac{2 \log \Pr(F_1(X_1) > u)}{\log \Pr(F_1(X_1) > u, F_2(X_2) > u)} - 1,
\]

which satisfies \(-1 < \tilde{\chi}_U \leq 1\). In particular, \( \tilde{\chi}_U = 1 \) if the distribution is asymptotically dependent in the upper tail and \(-1 < \tilde{\chi}_U < 1\) if the distribution is asymptotically independent in the upper tail, with larger values of \( \tilde{\chi}_U \) denoting stronger sub-asymptotic dependence.

The lower tail equivalent of \( \tilde{\chi}_U \) is given by

\[
\tilde{\chi}_L = \lim_{u \to 0} \frac{2 \log \Pr(F_1(X_1) < u)}{\log \Pr(F_1(X_1) < u, F_2(X_2) < u)} - 1
\]

which has the same properties as \( \tilde{\chi}_U \), but referred to the lower tail.

As an example, for a bivariate Normal distribution with correlation coefficient \( \rho \), \( \tilde{\chi}_U = \tilde{\chi}_L = \rho \).

Note that all the measures of extremal dependence discussed above are derived from probabilities evaluated on the variables \( F_1(X_1) \) and \( F_2(X_2) \), each of which has a Uniform\([0,1]\) distribution. Hence, they are copula-related measures which are invariant to any strictly increasing transformation of the marginal components \( X_1 \) and \( X_2 \).

3 Asymptotic independence of the skew-Normal distribution

Let \( (X_1, X_2) \) follow a bivariate skew-Normal distribution (SN) as defined in Azzalini and Capitanio (1999). As we will limit our attention to the extremal dependence measures introduced in Section 2, there is no loss of generality in assuming a standard SN with location parameters equal to 0 and scale parameters equal to 1. The joint density of \( (X_1, X_2) \) is then

\[
f(x_1, x_2; \alpha_1, \alpha_2, \rho) = 2\phi_2(x_1, x_2; \rho)\Phi_1(\alpha_1 x_1 + \alpha_2 x_2)
\]

(6)
where $\phi_2(\cdot, \cdot; \rho)$ denotes the density function of a standard bivariate Normal variable with correlation coefficient equal to $\rho$, $\Phi_1(\cdot)$ is the distribution function of a standard univariate Normal variable and $\alpha_1, \alpha_2 \in \mathbb{R}$ are the skewness parameters. For simplicity, we will refer to $\rho$ as the correlation coefficient, though in general it does not correspond to the actual correlation coefficient of the SN distribution which is a function of both $\rho$ and $\alpha_1, \alpha_2$ (see equation (4) of Azzalini and Capitanio, 1999).

In the following we will focus on the upper tail dependence of the SN distribution, since the property

$$f(x_1, x_2; \alpha_1, \alpha_2, \rho) = f(-x_1, -x_2; -\alpha_1, -\alpha_2, \rho),$$  \hspace{1cm} (7)

which can easily be verified from (6), implies that studying the behaviour of the lower tail of the SN distribution with skewness parameters $\alpha_1$ and $\alpha_2$ is equivalent to studying the behaviour of the upper tail with skewness parameters $-\alpha_1$ and $-\alpha_2$, respectively.

It can be shown that for any choice of $|\rho| < 1$ and of the skewness parameters, the SN distribution has $\chi_U = 0$, i.e. it is asymptotically independent in the upper tail. The proof is given in Appendix A.1. By (7) the lower tail is also asymptotically independent, i.e. $\chi_L = 0$.

![Figure 1: Plots of $\chi_U(u)$ against $u$. From top to bottom, the first line corresponds to $\alpha_1 = -2, \alpha_2 = 2$, and $\rho = 0.8$; the second line corresponds to $\alpha_1 = \alpha_2 = 0$ and $\rho = 0.8$, i.e. the standard bivariate Normal distribution with correlation coefficient 0.8; the third line corresponds to $\alpha_1 = \alpha_2 = 0.7$ and $\rho = 0.8$; the forth line corresponds to $\alpha_1 = \alpha_2 = -1$ and $\rho = 0.8$.](image)

Figure 1 shows plots of $\chi_U(u)$, evaluated numerically, against $u$ for $\rho = 0.8$ and for different choices of the skewness parameters, including $\alpha_1 = \alpha_2 = 0$ which corresponds to the standard bivariate Normal distribution. We see that, regardless of the choice of $\alpha_1$ and $\alpha_2$, the convergence of $\chi_U(u)$ to 0 is rather slow, with typically a sudden drop at the endpoint. However, differences
can be seen in the speed of convergence. It would appear, for example, that with negative $\alpha_1$ and $\alpha_2$ the decay to 0 is faster. This suggests that different strengths of sub-asymptotic dependence might be attainable by varying the degree of skewness. To investigate this aspect further, we derive $\bar{\chi}_U$ for the SN distribution.

Appendix A.2 contains proofs of the following results.

- For $\alpha_1, \alpha_2 > 0$ and $\rho \in (0, 1)$, $\bar{\chi}_U = \rho$. This means that with positive skewness parameters and correlation coefficient, the degree of sub-asymptotic dependence of the SN distribution is essentially equivalent to that of the generating bivariate Normal distribution.

- For $\alpha_1 = \alpha_2 = \alpha < 0$ and $|\rho| < 1$,
  \[
  \bar{\chi}_U = \frac{1 + \rho}{1 + \alpha^2(1 - \rho^2)} - 1 < \rho
  \]  
  \(8\)

  In this case, the sub-asymptotic dependence of the SN is weaker than that of the generating bivariate Normal distribution, confirming the findings from Figure 1.

- For $\alpha_1 = -\alpha_2 = \alpha$ and $|\rho| < 1$,
  \[
  \bar{\chi}_U \geq \frac{2(1 - \rho^2)(1 + \alpha^2)}{2 + \alpha^2 - 2\rho\sqrt{1 + \alpha^2}} - 1,
  \]  
  \(9\)

  where
  \[
  \bar{\alpha}^2 = \frac{\alpha^2(1 - \rho)^2}{1 + \alpha^2(1 - \rho^2)}.
  \]

  It is easy to verify that the lower bound in equation (9) is strictly greater than $\rho$ for any $|\rho| < 1$ and $\alpha \neq 0$. In these cases the SN has stronger sub-asymptotic dependence than the generating Normal distribution.

Cases other than those considered above seem more difficult to derive. However, this lack of completeness does not deter from our main objective which was to show that the SN allows more flexible forms of asymptotic independence than the Normal distribution, by having $\bar{\chi}_U$ which is generally a function also of $\alpha_1$ and $\alpha_2$. In particular, for fixed $\rho$, the SN distribution can have stronger, weaker or equivalent sub-asymptotic dependence than the Gaussian model. The changes in strength of tail dependence induced by variations of the skewness parameters also imply that the SN allows different levels of sub-asymptotic dependence in the upper and lower tails. For instance, for $\alpha_1 = \alpha_2 > 0$ and $\rho > 0$ the upper tail behaves like that of a Normal distribution with correlation coefficient $\rho$, while the lower tail exhibits weaker sub-asymptotic dependence. A by-product of the developments contained in Appendix A is that the skewness parameters, together with the correlation coefficient, affect the rate of decay of the tails of the marginal distributions of a bivariate SN variable (see equations (14) and (15)) though the domain of attraction remains that of the Gumbel distribution, as for the Gaussian family.

As a final remark, observe that the multivariate SN is also asymptotically independent. This follows from the fact that the bivariate margins of a multivariate SN are also SN distributed (Azzalini and Capitanio, 1999, Proposition 2) and from Corollary 5.25 of Resnick (1987) by which the components of a multivariate distribution are asymptotically independent if and only if they are pairwise asymptotically independent.
4 Asymptotic dependence of the skew-$t$ distribution

Let $(X_1, X_2)$ follow a bivariate skew-$t$ distribution (St) as defined in Azzalini and Capitanio (2003). Again, for the present purposes there is no loss of generality in limiting the attention to a standard St distribution, with location parameters equal to 0 and scale parameters equal to 1. The joint density of $(X_1, X_2)$ is then

$$f(x_1, x_2; \alpha_1, \alpha_2, \rho, \nu) = 2t_2(x_1, x_2; \rho, \nu)T_1 \left( \frac{\nu + 2}{Q(x_1, x_2; \rho) + \nu}; \nu + 2 \right), \quad (10)$$

where $T_1$ is as in equation (4), $Q(x_1, x_2; \rho)$ is given by

$$Q(x_1, x_2; \rho) = \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{1 - \rho^2},$$

$\alpha_1, \alpha_2 \in \mathbb{R}$ are the skewness parameters and $t_2(\cdot, \cdot; \rho, \nu)$ denotes the density function of a bivariate standard $t$ distribution with $\nu > 0$ degrees of freedom and off-diagonal element of the dispersion matrix equal to $\rho$.

As with the SN variable, we can focus on the upper tail behaviour of the St distribution, since, by the property

$$f(x_1, x_2; \alpha_1, \alpha_2, \rho, \nu) = f(-x_1, -x_2; -\alpha_1, -\alpha_2, \rho, \nu),$$

the lower tail of a St variable with skewness parameters $\alpha_1, \alpha_2$ corresponds to the upper tail of a St variable with skewness parameters $-\alpha_1, -\alpha_2$, respectively.

In Appendix B we show that, for any St distribution having $\rho > -1$, $\chi_U > 0$, implying that both tails of the St distribution exhibit asymptotic dependence. However, as the following results show, the strength of asymptotic dependence may vary from one tail to the other.

If $\alpha_1 = \alpha_2 = \alpha$, i.e. the margins are identically distributed, and $|\rho| < 1$ then the coefficient of upper tail dependence of the St variable is

$$\chi_U = K(\alpha, \nu, \rho)\chi_U(t_2), \quad (11)$$

where $\chi_U(t_2)$ is the coefficient of upper tail dependence for the generating bivariate $t$ distribution, whose expression is given by (4), and

$$K(\alpha, \nu, \rho) = \frac{T_1(2\alpha\sqrt{(\nu + 2)(1 + \rho)/2}; \nu + 2)}{T_1(\alpha(1 + \rho)\sqrt{\nu + 1}/\sqrt{\nu + \alpha^2(1 - \rho^2)}; \nu + 1)} \quad (12)$$

is the factor that differentiates the strength of tail dependence of the St distribution from that of the generating $t$ distribution. In particular,

- for $\alpha > 0$, $K(\alpha, \nu, \rho) > 1$, and the St distribution has stronger upper tail dependence than the generating $t$ distribution;
- for $\alpha < 0$, $K(\alpha, \nu, \rho) < 1$, and the St distribution has weaker asymptotic dependence than the generating $t$ distribution;
asymptotic independence is obtained as the limiting case of $\alpha \to -\infty$ for any $\nu > 0$, or of $\nu \to \infty$ for any $\alpha$, as for the bivariate $t$ distribution.

Similarly to the SN case, these results cover only a subset of all parameter configurations, but they are sufficient to demonstrate the increase in range of tail behaviours with respect to the generating family.

Since both tails of the St are asymptotically dependent we also have $\bar{\chi}_U = \bar{\chi}_L = 1$. In addition, by arguments similar to those applied to the multivariate SN, the multivariate St is asymptotically dependent.

## 5 Conclusions

The results in Sections 3 and 4 show that each of the bivariate SN and St distributions belongs to the same limiting class as that of its generating family. However, by having measures of tail dependence that are functions also of the skewness parameters, they exhibit more flexible forms of extremal dependence than the generating distributions. An interesting aspect that remains to be explored is whether this greater flexibility can be exploited for modelling extremal dependence as an alternative to the existing models for multivariate extremes. More precisely, the idea is to combine the censored likelihood approach of Ledford and Tawn (1997) and Bortot et al. (2000) with the tails of the bivariate SN or St distributions to model the extremal behaviour of an observed phenomenon. This approach would also have the advantage of admitting a straightforward extension to multivariate problems, since the multivariate versions of the SN and the St variables are easy to handle. Furthermore, we have seen that the two distributions allow different strengths of dependence in the lower tail and in the upper tail, suggesting each can be used to model data with different dependence characteristics at extremely high or extremely low levels.

## Appendix A

This appendix contains proofs of the asymptotic results for the SN distribution stated in Section 3.

We first consider the asymptotic behaviour of the marginal distributions of the bivariate SN which will be used in the following subsections. From Azzalini and Capitanio (1999), Proposition 2, the marginal distributions $F_1$ and $F_2$ of the bivariate SN variable $(X_1, X_2)$ defined in equation (6) are univariate SN with location parameter equal to 0, scale parameter equal to 1 and skewness parameters

\[
\bar{\alpha}_1 = \frac{\alpha_1 + \rho \alpha_2}{\sqrt{1 + \alpha_2^2(1 - \rho^2)}}, \quad \bar{\alpha}_2 = \frac{\alpha_2 + \rho \alpha_1}{\sqrt{1 + \alpha_1^2(1 - \rho^2)}},
\]

respectively. By developments similar to those of the well-known asymptotic expansion $1 - \Phi_1(x) \sim \frac{\phi_1(x)}{x}$, as $x \to \infty$, where $\phi_1(\cdot)$ denotes the density function of a standard univariate Normal variable, it is easy to verify that
• if $\bar{\alpha}_i < 0$, $i = 1, 2$,

$$1 - F_i(x) \sim \frac{\sqrt{2} \phi_1 \left( x \sqrt{1 + \bar{\alpha}_i^2} \right)}{\sqrt{\pi |\bar{\alpha}| x^2 (1 + \bar{\alpha}_i^2)}}$$

as $x \to \infty$;

(14)

• if $\bar{\alpha}_i > 0$, $i = 1, 2$,

$$1 - F_i(x) \sim \frac{\phi_1(x)}{x} \Phi_1 (\bar{\alpha}_i x),$$

as $x \to \infty$.

(15)

In the following $(Z_1, Z_2)$ will denote a bivariate Normal variable with standard margins and correlation coefficient $\rho$. In addition, throughout this Appendix we will work with $|\rho| < 1$.

A.1: Proof of the asymptotic independence of the skew-Normal distribution

To prove the asymptotic independence of the SN distribution we consider the following cases which cover all the possible parameter configurations.

(a) $\bar{\alpha}_1, \bar{\alpha}_2 \geq 0$

We assume $\bar{\alpha}_1 \leq \bar{\alpha}_2$, with the reversed inequality being proved in an analogous way. We have

$$\chi_U = \lim_{u \to 1} \Pr(F_1(X_1) > u | F_2(X_2) > u) = \lim_{u \to 1} \Pr(F_2(X_2) > u | F_1(X_1) > u) = \lim_{x \to \infty} \Pr(F_2(X_2) > F_1(x) | X_1 > x).$$

By equation (15) we see that, for large $x$, $F_1(x) \geq F_2(x)$, hence

$$\chi_U \leq \lim_{x \to \infty} \Pr(F_2(X_2) > F_2(x) | X_1 > x) =$$

$$= \lim_{x \to \infty} \Pr(X_2 > x | X_1 > x) = \lim_{x \to \infty} \frac{\Pr(X_1 > x, X_2 > x)}{\Pr(X_1 > x)}.$$

From $\Phi_1 (\alpha_1 x_1 + \alpha_2 x_2) \leq 1$, for $\alpha_1, \alpha_2, x_1, x_2 \in \mathbb{R}$, we have that $\Pr(X_1 > x, X_2 > x) \leq 2 \Pr(Z_1 > x, Z_2 > x)$. In addition, from Azzalini (2005) $\min(Z_1, Z_2)$ has a univariate SN distribution with skewness parameter equal to $-\sqrt{(1 - \rho)/(1 + \rho)} < 0$. Hence, by (14),

$$0 \leq \chi_U \leq \lim_{x \to \infty} \frac{2 \Pr(Z_1 > x, Z_2 > x)}{\Pr(X_1 > x)} = \lim_{x \to \infty} k \frac{\phi_1(x \sqrt{1 + (1 - \rho)/(1 + \rho)})}{x^2 \phi_1(x) \Phi_1 (\bar{\alpha}_1 x)} = 0,$$

where $k$ is a positive constant, which proves the asymptotic independence for $\bar{\alpha}_1, \bar{\alpha}_2 \geq 0$.

(b) $\alpha_1, \alpha_2 < 0$

With this choice of the skewness parameters, the only possible configurations for the marginal parameters are $\bar{\alpha}_1, \bar{\alpha}_2 < 0$, or $\bar{\alpha}_1 < 0, \bar{\alpha}_2 > 0$, or $\bar{\alpha}_1 > 0, \bar{\alpha}_2 < 0$. We assume that $\bar{\alpha}_1 < \bar{\alpha}_2$, as the reversed inequality is proved in a similar manner. Note that under this assumption we have $\bar{\alpha}_1 < 0$. As in case (a),

$$\chi_U \leq \lim_{x \to \infty} \frac{\Pr(X_1 > x, X_2 > x)}{\Pr(X_1 > x)}.$$
The inequality \( \Phi_1(\alpha_1 x_1 + \alpha_2 x_2) \leq \Phi_1(x(\alpha_1 + \alpha_2)) \), valid for \( x_1, x_2 \geq x \) and \( \alpha_1, \alpha_2 < 0 \), implies that

\[
\chi_u \leq \lim_{x \to \infty} \frac{2 \Pr(Z_1 > x, Z_2 > x) \Phi_1(x(\alpha_1 + \alpha_2))}{\Pr(X_1 > x)} = 
\lim_{x \to \infty} k \frac{\phi_1(x\sqrt{1 + (1 - \rho)/(1 + \rho)}) \phi_1(x(\alpha_1 + \alpha_2))}{x \phi_1(x\sqrt{1 + \alpha_1^2})},
\]

where \( k \) is a positive constant. The latter limit is equal to 0 if

\[
\frac{2}{1 + \rho} + (\alpha_1 + \alpha_2)^2 - 1 - \alpha_1^2 > 0,
\]

which is true for any \(|\rho| < 1\).

(c) \( \bar{\alpha}_1, \bar{\alpha}_2 < 0 \) with \( \alpha_1 < 0, \alpha_2 \geq 0 \) or \( \alpha_1 \geq 0, \alpha_2 < 0 \)

We focus on the case \( \alpha_1 < 0, \alpha_2 \geq 0 \), as the other setting can be derived similarly. Under this assumption we have \( \bar{\alpha}_1 < \bar{\alpha}_2 \) and from (14), \( F_2 \left( x \sqrt{1 + \alpha_1^2/1 + \alpha_2^2} \right) \leq F_1(x) \), for large \( x \). Then,

\[
\chi_u = \lim_{x \to \infty} \Pr(F_2(X_2) > F_1(x)|X_1 > x) \leq \lim_{x \to \infty} \Pr(F_2(X_2) > F_2 \left( x \sqrt{1 + \alpha_1^2/1 + \alpha_2^2} \right)|X_1 > x) = 
\lim_{z \to \infty} \Pr \left( X_2 > z | X_1 > x \right) \frac{\Pr \left( X_1 > z \sqrt{1 + \alpha_1^2/1 + \alpha_2^2}, X_2 > z \right)}{\Pr \left( X_1 > z \sqrt{1 + \alpha_1^2/1 + \alpha_2^2} \right)}. \tag{16}
\]

We now apply l'Hôpital’s rule to the latter limit in (16). For the derivative of the denominator, \( D_1(z) \), we have

\[
D_1(z) \sim k_1 \frac{\exp \left\{ -\frac{1}{2} z^2 (1 + \alpha_2^2) \right\}}{z} \quad \text{as } z \to \infty
\]

where \( k_1 \) is a negative constant. By the total derivative rule, the derivative of the numerator, \( D_2(z) \), is such that

\[
D_2(z) = D_{21}(z) + D_{22}(z)
\]

where

\[
D_{21}(z) = \int_z^{\infty} -2 \sqrt{1 + \alpha_2^2} \phi_1(z^*, x_2) \Phi_1(\alpha_1 z^* + \alpha_2 x_2) dx_2,
\]

with

\[
z^* = z \sqrt{1 + \alpha_2^2/1 + \alpha_1^2},
\]

and

\[
D_{22}(z) = \int_{z^*}^{\infty} -2 \phi_1(x_1, z) \Phi_1(\alpha_1 x_1 + \alpha_2 z) dx_1.
\]

By integration by parts and knowing that \( \alpha_1 z^* + \alpha_2 z \to -\infty \), as \( z \to \infty \), we have that

\[
D_{21}(z) \sim k_2 \frac{\phi_1(z^*) \phi_1(\alpha_1 z^* + \alpha_2 z)}{(z - \rho z^*)|\alpha_1 z^* + \alpha_2 z|},
\]

\[
D_{22}(z) \sim k_2 \frac{\phi_1(z^*) \phi_1(\alpha_1 z^* + \alpha_2 z)}{(z - \rho z^*)|\alpha_1 z^* + \alpha_2 z|}.
\]

9
where \( k_2 \) is a negative constant and

\[
\phi^*(z) = \exp \left\{ -\frac{1}{2(1 - \rho^2)} (z^* + z^2 - 2\rho z^* z) \right\}.
\]

Hence,

\[
\frac{D_{21}(z)}{D_1(z)} = O \left( \frac{z\psi(z)}{(z - \rho z^*) |\alpha_1 z^* + \alpha_2 z|} \right),
\]

where

\[
\psi(z) = \exp \left\{ -\frac{1}{2} z^2 \left[ \left( \frac{1}{1 + \alpha_2^2} \right) \left( \frac{1}{1 - \rho^2} + \alpha_2^2 \right) + \left( \frac{\rho^2}{1 - \rho^2} + \alpha_2^2 - \alpha_2^2 \right) + 2 \frac{1}{1 + \alpha_2^2} \left( \alpha_1 \alpha_2 - \frac{\rho}{1 - \rho^2} \right) \right] \right\}.
\]

It is easy to verify that the term in square brackets in equation (17) is strictly positive on the range \( \alpha_1 < 0, \alpha_2 \geq 0 \) and \( \bar{\alpha}_1, \bar{\alpha}_2 < 0 \). Therefore, \( D_{21}(z)/D_1(z) \to 0 \), as \( z \to \infty \).

Again, applying integration by parts to \( D_{22}(z) \), we have that

\[
\frac{D_{22}(z)}{D_1(z)} = O \left( \frac{z\psi(z)}{z^* (1/(1 - \rho^2) + \alpha_2^2) + z(\alpha_1 \alpha_2 - \rho/(1 - \rho^2)) |\alpha_1 z^* + \alpha_2 z|} \right),
\]

where \( \psi(z) \) is as in (17). Hence, \( D_{22}(z)/D_1(z) \to 0 \), as \( z \to \infty \) and this completes the proof of case (c).

(d) \( \bar{\alpha}_1 < 0, \bar{\alpha}_2 \geq 0 \) or \( \bar{\alpha}_1 \geq 0, \bar{\alpha}_2 < 0 \)

We focus on the first setting, i.e. \( \bar{\alpha}_1 < 0, \bar{\alpha}_2 \geq 0 \), as the second one can be treated in a similar manner. Note that the selected configuration of marginal skewness parameters can be obtained with \( \alpha_1 < 0, \alpha_2 \geq 0 \), or \( \alpha_1, \alpha_2 > 0 \) and \( \rho < 0 \), or \( \alpha_1, \alpha_2 < 0 \) and \( \rho < 0 \). By (14) and (15) we see that, for large \( x \), \( F_1(x) > F_2(x \sqrt{1 + \alpha_2^2}) \). Then,

\[
\chi_U = \lim_{x \to \infty} \Pr(F_2(X_2) > F_1(x) | X_1 > x) \leq \lim_{x \to \infty} \Pr(F_2(X_2) > F_2(x \sqrt{1 + \alpha_2^2}) | X_1 > x) = \lim_{x \to \infty} \frac{\Pr(X_1 > z/\sqrt{1 + \alpha_2^2}, X_2 > z)}{\Pr(X_1 > z/\sqrt{1 + \alpha_2^2})} \leq \lim_{x \to \infty} \frac{2\Pr(Z_1 > z/\sqrt{1 + \alpha_1^2}, Z_2 > z)}{\Pr(X_1 > z/\sqrt{1 + \alpha_2^2})},
\]

where the latter inequality comes from \( \Phi_1(\alpha_1 x_1 + \alpha_2 x_2) \leq 1 \), for \( x_1, x_2, \alpha_1, \alpha_2 \in \mathbb{R} \). We can apply Savage’s approximation (Savage, 1962) to \( \Pr(Z_1 > z/\sqrt{1 + \alpha_1^2}, Z_2 > z) \) to derive

\[
\Pr(Z_1 > z/\sqrt{1 + \alpha_1^2}, Z_2 > z) \sim \frac{\phi_2(z/\sqrt{1 + \alpha_1^2}, z)}{z^2(1/(1 + \alpha_1^2) - \rho)(1 - \rho/\sqrt{1 + \alpha_1^2})}, \quad \text{as } z \to \infty. \quad (18)
\]

Note that the above approximation requires that Savage’s condition is satisfied which, in the present context, reduces to the following inequalities

\[
\frac{1}{\sqrt{1 + \alpha_1^2}} - \rho > 0 \quad \text{and} \quad 1 - \frac{\rho}{\sqrt{1 + \alpha_1^2}} > 0.
\]
In this appendix we derive $\bar{\alpha}$. Derivation of $\bar{\alpha}$ knowing that $D$ it is readily checked that the term on the left-hand side of (20) is equal to

$$\frac{\Pr(Z_1 > z/\sqrt{1+\bar{\alpha}_1^2}, Z_2 > z)}{\Pr(X_1 > z/\sqrt{1+\bar{\alpha}_1^2})} = O \left( \exp \left\{ -\frac{1}{2} z^2 \left( \frac{1 - \rho \sqrt{1 + \bar{\alpha}_1^2}}{(1 + \bar{\alpha}_1^2)(1 - \rho^2)} \right)^2 \right\} \right), \quad \text{as } z \to \infty. \tag{20}$$

The factor multiplying $z^2$ within the exponential function is strictly negative, which proves $\chi_U = 0$ for case (d).

Observe that Savage’s approximation cannot be applied to case (c) as it generally does not satisfy Savage’s condition.

**A.2: Derivation of $\bar{\chi}_U$ for the skew-Normal distribution**

In this appendix we derive $\bar{\chi}_U$ for the cases considered in Section 3.

- $\alpha_1, \alpha_2 > 0$ and $\rho \in (0, 1)$

Note that with this parameter setting $\bar{\alpha}_1, \bar{\alpha}_2 > 0$. We assume $\bar{\alpha}_1 \leq \bar{\alpha}_2$ as the reversed condition can be proved similarly. From (5)

$$\bar{\chi}_U = \lim_{u \to 1} \frac{2 \log \Pr(X_1 > u_1)}{\log \Pr(X_1 > u_1, X_2 > u_2)} - 1 \quad \tag{19}$$

where $u_1 = F_1^{-1}(u)$ and $u_2 = F_2^{-1}(u)$. We denote by $D_1(u)$ and $D_2(u)$ the numerator and the denominator, respectively, of the fraction in (19). From the assumption $\bar{\alpha}_1 \leq \bar{\alpha}_2$ it follows that, as $u \to 1$, $u_1 \leq u_2$ and

$$\log \Pr(X_1 > u_2, X_2 > u_2) \leq D_2(u) \leq \log \Pr(X_1 > u_1, X_2 > u_1).$$

Furthermore, the inequalities $\Phi_1(u_2(\alpha_1 + \alpha_2)) \leq \Phi_1(\alpha_1 x_1 + \alpha_2 x_2)$, valid for $x_1, x_2 \geq u_2$ and $\alpha_1, \alpha_2 > 0$, and $\Phi_1(\alpha_1 x_1 + \alpha_2 x_2) \leq 1$ lead to

$$\log [2\Pr(Z_1 > u_2, Z_2 > u_2)\Phi_1(u_2(\alpha_1 + \alpha_2))] \leq D_2(u) \leq \log [2\Pr(Z_1 > u_1, Z_2 > u_1)].$$

Hence,

$$\lim_{u \to 1} \frac{D_1(u)}{\log [2\Pr(Z_1 > u_1, Z_2 > u_1)]} - 1 \leq \bar{\chi}_U \leq \lim_{u \to 1} \frac{D_1(u)}{\log [2\Pr(Z_1 > u_2, Z_2 > u_2)\Phi_1(u_2(\alpha_1 + \alpha_2))]} - 1. \tag{20}$$

Since $\min(Z_1, Z_2)$ has a univariate SN distribution with skewness parameter $-\sqrt{(1 - \rho)/(1 + \rho)}$, it is readily checked that the term on the left-hand side of (20) is equal to $\rho$. Similarly, knowing that $D_1(u)$ can be expressed also as

$$D_1(u) = 2 \log \Pr(X_2 > u_2),$$

the term on the right-hand side of (20) is equal to $\rho$, which proves the result.
• $\alpha_1 = \alpha_2 = \alpha < 0$ and $|\rho| < 1$

In this case $\tilde{\alpha}_1 = \tilde{\alpha}_2 = \tilde{\alpha}$ with

$$\tilde{\alpha} = \frac{\alpha(1 + \rho)}{\sqrt{1 + \alpha^2(1 - \rho^2)}},$$

and $\bar{\chi}_U$ reduces to

$$\bar{\chi}_U = \lim_{x \to \infty} \frac{2 \log \Pr(X_1 > x)}{\log \Pr(X_1 > x, X_2 > x)} - 1.$$

We have

$$\Pr(X_1 > x, X_2 > x) \sim \int_x^\infty \int_x^\infty 2\phi_2(x_1, x_2; \rho) \frac{\phi_1(\alpha(x_1 + x_2))}{-\alpha(x_1 + x_2)} dx_1 dx_2, \quad \text{as } x \to \infty. \quad (22)$$

By applying integration by parts to the double integral in (22) and neglecting terms of smaller order we obtain

$$\Pr(X_1 > x, X_2 > x) \sim k\psi(x) x^3 \left(2\alpha^2 + \frac{1}{1+\rho}\right)^2, \quad \text{as } x \to \infty$$

where $k$ is a positive constant and

$$\psi(x) = \exp \left\{ -x^2 \left(2\alpha^2 + \frac{1}{1+\rho}\right) \right\}.$$

Hence, denoting by $k_1$ a positive constant, we have

$$\bar{\chi}_U = \lim_{x \to \infty} \frac{2 \log \left[ k_1 \phi_1 \left( x\sqrt{1 + \alpha^2} / x^2 \right) \right]}{\log \left[ \psi(x) / \left\{ x^3 \left(2\alpha^2 + \frac{1}{1+\rho}\right)^2 \right\} \right]} - 1 = \frac{1 + \rho}{1 + \alpha^2(1 - \rho^2)} < \rho.$$

• $\alpha_1 = -\alpha_2 = \alpha$ and $|\rho| < 1$

We focus on $\alpha < 0$ as the reversed inequality can be proved similarly. We then have

$$\tilde{\alpha}_1 = -\tilde{\alpha}_2 = \frac{\alpha(1 - \rho)}{\sqrt{1 + \alpha^2(1 - \rho^2)}} = \alpha^*,$$

with $\tilde{\alpha}_1 < 0$ and $\tilde{\alpha}_2 > 0$. As in case (d) of Appendix A.1, for large $x$, $F_1(x) > F_2(x\sqrt{1 + \alpha^*^2})$ and

$$\bar{\chi}_U \geq \lim_{z \to \infty} \frac{2 \log \Pr(X_1 > z/\sqrt{1 + \alpha^*^2})}{\log \Pr(X_1 > z/\sqrt{1 + \alpha^*^2}, X_2 > z)} - 1 \geq \lim_{z \to \infty} \frac{2 \log \Pr(Z_1 > z/\sqrt{1 + \alpha^*^2}, Z_2 > z)}{\log 2 \Pr(Z_1 > z/\sqrt{1 + \alpha^*^2}, Z_2 > z)} - 1.$$ 

By applying Savage’s approximation (18) to the denominator of the latter limit we obtain

$$\bar{\chi}_U \geq \lim_{z \to \infty} \frac{k_1 - z^2}{k_2 - z^2 \left( \frac{1}{\sqrt{1 - \rho^2}} - \frac{2\rho}{\sqrt{1 + \alpha^*^2}} \right)} - 1 = \frac{2(1 - \rho^2)(1 + \alpha^*^2)}{2 + \alpha^*^2 - 2\rho\sqrt{1 + \alpha^*^2}} - 1, \quad (23)$$

where $k_1, k_2$ are positive constants. It is then easy to verify that the term on the right-hand side of equation (23) is greater than $\rho$ for any $|\rho| < 1$ and $\alpha \neq 0$. 

12
Appendix B

This appendix contains proofs of the asymptotic results for the St distribution stated in Section 4.

From Azzalini and Capitanio (2003), the marginal distributions $F_1$ and $F_2$ of the bivariate St variable $(X_1, X_2)$ defined in equation (10) are univariate St with $\nu$ degrees of freedom, location parameter equal to 0, scale parameter equal to 1 and skewness parameters $\bar{\alpha}_1$, $\bar{\alpha}_2$, respectively, given by equation (13). It is easy to verify that

$$1 - F_i(x) \sim 2T_1(\bar{\alpha}_i\sqrt{\nu + 1}; \nu + 1) (1 - T_1(x; \nu)),$$

as $x \to \infty$. (24)

In the following we denote by $(Y_1, Y_2)$ a bivariate standard $t$ variable with $\nu$ degrees of freedom and off-diagonal element of the dispersion matrix equal to $\rho$.

To prove the asymptotic dependence of the bivariate St distribution with $\rho > -1$ we assume $\bar{\alpha}_1 \leq \bar{\alpha}_2$, as the reversed inequality can be treated similarly. Under this assumption, for large $x$, $F_1(x) \geq F_2(x)$; hence

$$\chi_U = \lim_{x \to \infty} \frac{\Pr(F_1(X_1) > u, F_2(X_2) > u)}{\Pr(F_2(X_2) > u)} \geq$$

$$\geq \lim_{x \to \infty} \frac{\Pr(X_1 > x, X_2 > x)}{\Pr(X_2 > x)} = \lim_{x \to \infty} \frac{2\Pr(Y_1 > x, Y_2 > x)T_1(\alpha_1 + \alpha_2)\sqrt{(\nu+2)(1+\rho)}; \nu + 2}{2T_1(\bar{\alpha}_2\sqrt{\nu + 1}; \nu + 1) (1 - T_1(x; \nu))} =$$

$$= \frac{T_1\left(\alpha_1 + \alpha_2)\sqrt{(\nu+2)(1+\rho)}; \nu + 2\right)}{T_1(\bar{\alpha}_2\sqrt{\nu + 1}; \nu + 1)},$$

(25)

where $\chi_U(t_2)$ is the value of $\chi_U$ for the generating bivariate $t$ variable $(Y_1, Y_2)$. Since (25) is strictly greater than zero for any $\nu \in \mathbb{R}$ and $\rho > -1$ the result is proved.

To derive the exact value of $\chi_U$ we concentrate on the case of identical margins, i.e. $\alpha_1 = \alpha_2 = \alpha$, and take $|\rho| < 1$. Under this assumption, the coefficient of upper tail dependence reduces to

$$\chi_U = \lim_{x \to \infty} \frac{\Pr(X_1 > x, X_2 > x)}{\Pr(X_1 > x)}.$$

It follows that

$$\chi_U = \lim_{x \to \infty} \frac{T_1\left(2\alpha\sqrt{(\nu+2)(1+\rho)}; \nu + 2\right)\Pr(Y_1 > x, Y_2 > x)}{T_1(\bar{\alpha}\sqrt{\nu + 1}; \nu + 1) (1 - T_1(x; \nu))},$$

where $\bar{\alpha}$ is the marginal skewness parameter, common to $F_1$ and $F_2$, given by (21). Hence,

$$\chi_U = K(\alpha, \nu, \rho)\chi_U(t_2),$$

with

$$K(\alpha, \nu, \rho) = \frac{T_1\left(2\alpha\sqrt{(\nu+2)(1+\rho)}; \nu + 2\right)}{T_1\left(\bar{\alpha}\sqrt{\nu + 1}; \nu + 1\right)}.$$
For $\alpha < 0$,
\[ 2\alpha \sqrt{\frac{(\nu + 2)(1 + \rho)}{2}} < \bar{\alpha} \sqrt{\nu + 1} < 0 \]
while, for $\alpha > 0$,
\[ 2\alpha \sqrt{\frac{(\nu + 2)(1 + \rho)}{2}} > \bar{\alpha} \sqrt{\nu + 1} > 0 \]
Taking into account that $T_1(x; \nu + 2) < T_1(x; \nu + 1)$, for $x < 0$, and $T_1(x; \nu + 2) > T_1(x; \nu + 1)$, for $x > 0$, we have $K(\alpha, \nu, \rho) < 1$ for $\alpha < 0$ and $K(\alpha, \nu, \rho) > 1$ for $\alpha > 0$. Finally, $K(\alpha, \nu, \rho) \to 0$ as $\alpha \to -\infty$, which completes the proof.

**Acknowledgments**

I would like to thank Antonella Capitanio for helpful discussions on a preliminary version of the paper. This work was funded by PRIN 2006, grant No. 2006132978, from MIUR, Italy.

**References**


