# Risk and Portfolio Analysis: Principles and Methods
# Solutions to exercises

Björn Löfdahl

January 8, 2015

# Contents

# 7 Empirical Methods

**Problem 7.1.** A unit within a bank is required to report an empirical estimate of $VaR_{0.01}(X)$, where $X$ is the portfolio value the next day from its trading activities. The empirical estimate $\widehat{VaR}_{0.01}(X)$ is based on market prices from the previous $n+1$ days that are transformed into a sample of size $n$ from the distribution of $X$, and the sample points are assumed to be independent and identically distributed. Compute the probability

$$P\Big(\widehat{VaR}_{0.01}(X) > VaR_{0.01}(X)\Big)$$

as a function of $n$ and determine its minimum and maximum for $n = 100, 101, \ldots, 300$.

**Solution.** We assume that effects from interest rates are negligible since we are dealing with a one-day horizon. Thus, we have $L = -X/R_0 = -X$. Recall the definition of Value-at-Risk,

$$VaR_p(X) = F_L^{-1}(1-p) = \min\{x : F(x) \geq 1-p\},$$

and that the empirical VaR estimator is given by

$$\widehat{VaR}_p(X) = F_{n,L}^{-1}(1-p) = L_{[np]+1,n}, \quad where \quad L_{1,n} \geq \ldots \geq L_{n,n}.$$

Now, let $Y_{F_L^{-1}(q)}$ be the number of sample points exceeding $F_L^{-1}(q)$, with $q = 1-p$. We obtain

$$\begin{aligned} P\Big(\widehat{VaR}_p(X) > VaR_p(X)\Big) &= P\Big(L_{[np]+1,n} > F_L^{-1}(1-p)\Big) \\ &= P\Big(Y_{F_L^{-1}(q)} \geq [np]+1\Big). \end{aligned}$$

Each sample point exceeds the $q$-quantile with probability $1-q$, independently of the other points. Thus, the number of sample points exceeding the q-quantile is binomially distributed, $Y_{F_L^{-1}(q)} \sim Bin(n, r)$ with

$$r = P\Big(L > F_L^{-1}(1-p)\Big) = 1 - F_L(F_L^{-1}(1-p)) = 1 - (1-p) = p,$$

if $F$ is continuous. Thus, we have

$$P\Big(Y_{F_L^{-1}(q)} \geq [np]+1\Big) = \sum_{k=[np]+1}^{n} \binom{n}{k} p^k (1-p)^{n-k}.$$

We find

$$\begin{aligned} \max_n P\Big(\widehat{VaR}_{0.01}(X) > VaR_{0.01}(X)\Big) &= 0.5926 \quad n = 199 \\ \min_n P\Big(\widehat{VaR}_{0.01}(X) > VaR_{0.01}(X)\Big) &= 0.2642 \quad n = 100 \end{aligned}$$

**Problem 7.2.** The tail conditional median $TCM_p(X) = median[L|L \geq VaR_p(X)]$, where $L = -X/R_0$, has been proposed as a more robust alternative to $ES_p(X)$ since $TCM_p(X)$ is not as sensitive as $ES_p(X)$ to the behaviour of the left tail of the distribution of $X$. Let $Y$ have a standard Student's t distribution with $\nu$ degrees of freedom, and set $X = e^{0.01Y} - 1$. Consider the empirical estimators $\widehat{TCM}_{0.01}(X)$ and $\widehat{ES}_{0.01}(X)$ based on a sample of size 1000 from the distribution of $L = -X$. Generate histograms based on samples of size $10^5$ from the distribtutions of $\widehat{TCM}_{0.01}(X)$ and $\widehat{ES}_{0.01}(X)$ for $\nu = 2$ and $\nu = 10$.

**Solution.** Recall that Expected shortfall is defined as

$$ES_p(X) = \frac{1}{p}\int_0^p VaR_u(X)du.$$

Since $[np]$ is an integer in this case, the empirical estimators are given by

$$\begin{aligned}\widehat{ES}_p(X) &= \frac{1}{p}\int_0^p \widehat{VaR}_u(X)du = \frac{1}{np}\sum_{k=1}^{np} L_{k,n} \\ \widehat{TCM}_p(X) &= median[L|L \geq \widehat{VaR}_p(X)] = median[L|L \geq L_{[np]+1,n}].\end{aligned}$$

The generated histograms of the estimators are presented below.

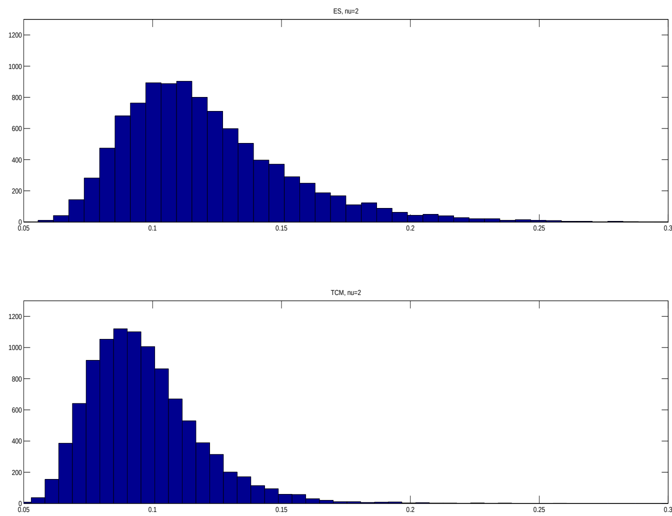


Figure 1: Expected shortfall and Tail conditional median for $\nu = 2$.

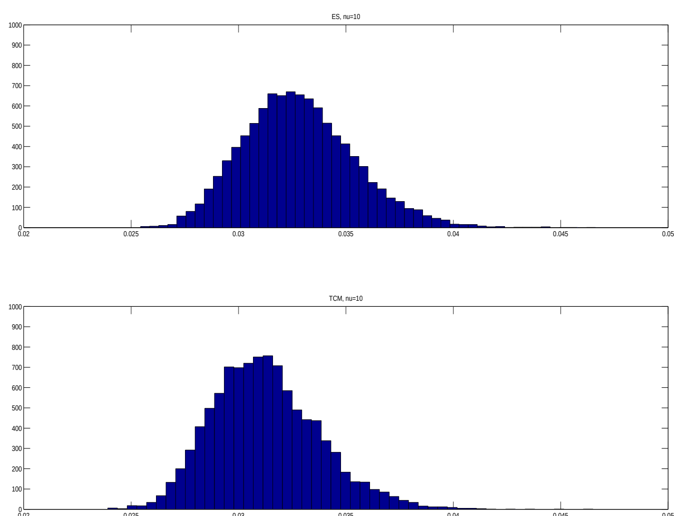


Figure 2: Expected shortfall and Tail conditional median for $\nu = 10$.

**Problem 7.3.** Let $\{Z_1, \ldots, Z_n\}$ be a sample of independent and identically distributed historical log returns that are distributed as the log return $\log \frac{S_T}{S_0}$ of an asset from today until time $T > 0$. Show that if the risk-free return over the investment period is 1, then the empirical estimator of $ES_p(S_T - S_0)$ is given by

$$\min_c -c + \frac{1}{np}\sum_{k=1}^{n}(c + S_0 - S_0 e^{Z_k})I\{Z_k \leq \log(1 + \frac{c}{S_0})\}.$$

**Solution.** Using proposition 6.5, ES has the representation

$$ES_p(X) = \min_c -c + \frac{1}{p}E[(c - \frac{X}{R_0})_+]. \tag{1}$$

The risk-free return over the investment period is 1, so $R_0 = 1$. Defining the loss $L = -X$, we rewrite (1) as

$$ES_p(X) = \min_c -c + \frac{1}{p}E[(c + L)I\{c + L \geq 0\}]. \tag{2}$$

The empirical estimator of (2) is

$$\widehat{ES}_p(X) = \min_c -c + \frac{1}{p}\hat{E}[(c + L)I\{c + L \geq 0\}], \tag{3}$$

where $\hat{E}$ denotes the expectation with respect to the empirical distribution of L, with $P(L = l_k) = \frac{1}{n}, \quad l = 1, \ldots n$. Expressing the loss in terms of log-returns, we obtain

$$L = -X = S_0 - S_T = S_0 - S_0 e^Z. \tag{4}$$

Inserting (4) into (3) and using the empirical distribution of L yields

$$\begin{aligned}
\widehat{ES}_p(X) &= \min_c -c + \frac{1}{np}\sum_{k=1}^{n}(c + L_k)I\{c + L_k \geq 0\} \\
&= \min_c -c + \frac{1}{np}\sum_{k=1}^{n}(c + S_0 - S_0 e^{Z_k})I\{Z_k \leq \log(1 + \frac{c}{S_0})\}
\end{aligned}$$

**Problem 7.4.** Let $\{Z_1, \ldots, Z_n\}$ be a sample of independent and identically distributed historical log returns that are distributed as the log return $\log \frac{S_T}{S_0}$ of an asset from today until time $T > 0$. Show that if the risk-free return over the investment period is 1 and if $\rho_\phi$ is a spectral risk measure with risk aversion function $\phi$, then the empirical estimator of $\rho_\phi(S_T - S_0)$ is given by

$$S_0 - S_0 \sum_{k=1}^{n} e^{Z_{k,n}} \int_{(n-k)/n}^{(n-k+1)/n} \phi(u)du.$$

**Solution.** Recall that a spectral risk measure $\rho_\phi$ is defined by

$$\rho_\phi(X) = -\int_0^1 \phi(u) F^{-1}_{X/R_0}(u)du, \tag{5}$$

where $\phi$ is decreasing, non-negative and integrates to 1. It is natural to estimate $\rho_\phi(X)$ using the empirical distribution of $X$. The empirical quantile function is given by $F^{-1}_{n,X}(p) = X_{[n(1-p)]+1,n}$. Moreover, since the risk-free return is 1, we have $L = -X/R_0 = -X$. Thus, we estimate $\rho_\phi(X)$ by

$$\hat{\rho}_\phi(X) = -\int_0^1 \phi(u) F^{-1}_{n,X}(u)du = \int_0^1 \phi(u) F^{-1}_{n,L}(1-u)du = \int_0^1 \phi(u) L_{[nu]+1,n}du. \tag{6}$$

$L_{[nu]+1,n}$ is constant between integer values of $[nu]$, which implies

$$\int_{u=(k-1)/n}^{k/n} \phi(u) L_{[nu]+1,n}du = L_{k,n} \int_{u=(k-1)/n}^{k/n} \phi(u)du, \quad k = 1, \ldots, n. \tag{7}$$

We have seen that we can express the loss $L$ as $L = S_0 - S_0 e^Z$. Now since $L$ is decreasing in $Z$, we must have $L_{k,n} = S_0 - S_0 e^{Z_{n-k+1,n}}$. Using this fact and inserting (7) into (6) yields

$$\begin{aligned}
\hat{\rho}_\phi(X) &= \sum_{k=1}^{n} \int_{u=(k-1)/n}^{k/n} \phi(u) L_{[nu]+1,n}du = \sum_{k=1}^{n} L_{k,n} \int_{u=(k-1)/n}^{k/n} \phi(u)du \\
&= \sum_{k=1}^{n} (S_0 - S_0 e^{Z_{n-k+1,n}}) \int_{u=(k-1)/n}^{k/n} \phi(u)du \\
&= S_0 - S_0 \sum_{k=1}^{n} e^{Z_{n-k+1,n}} \int_{u=(k-1)/n}^{k/n} \phi(u)du.
\end{aligned}$$

# 8 Parametric Models and Their Tails

**Problem 8.1.** The distribution function $F(x) = p\Phi(x/\sigma_1)+(1-p)\Phi(x/\sigma_2)$ of a mixture of the two normal distributions $\mathrm{N}(0, \sigma_1^2)$ and $\mathrm{N}(0, \sigma_2^2)$ corresponds to drawing a value with propability $p$ from the $\mathrm{N}(0, \sigma_1^2)$-distribution and with propability $1 - p$ from the $\mathrm{N}(0, \sigma_2^2)$-distribution.

(a) Use maximum likelihood to estimate the parameters $p, \sigma_1, \sigma_2$ based on the sample $\{t_4^{-1}(k/201) : k = 1, \ldots, 200)\}$.

(b) Plot the density function of the mixture distribution with the parameters estimated in (a) and compare it to the density function of the standard Student's $t$ distribution with four degrees of freedom.

(c) Plot the quantiles of the Student's $t$ distribution with four degrees of freedom against the quantiles of the mixture distribution with the parameters estimated in (a).

(d) Determine the asymptotic behavior of $F(x)$ as $x \to -\infty$ in terms of an explicitly given function $G$ such that $\lim_{x\to-\infty} F(x)/G(x) = 1$.

**Solution.** The maximum likelihood estimates of $p, \sigma_1, \sigma_2$ are the values that maximise the log-likelihood function $l(p, \sigma_1, \sigma_2)$ defined by

$$l(p, \sigma_1, \sigma_2) = \sum_{k=1}^{n} \log f(x_k|p, \sigma_1, \sigma_2), \tag{8}$$

where $x_1, \ldots, x_n$ is an i.i.d. sample from some distribution. The density of the normal mixture is given by

$$f(x|p, \sigma_1, \sigma_2) = \frac{d}{dx}F(x) = \frac{p}{\sigma_1}\phi(\frac{x}{\sigma_1}) + \frac{1-p}{\sigma_2}\phi(\frac{x}{\sigma_2}). \tag{9}$$

Inserting (9) into (8), we obtain

$$\sum_{k=1}^{n} \log\Big(\frac{p}{\sigma_1}\phi(\frac{x_k}{\sigma_1}) + \frac{1-p}{\sigma_2}\phi(\frac{x_k}{\sigma_2})\Big). \tag{10}$$

We maximize (10) numerically and obtain the parameter estimates $(\hat{p}, \hat{\sigma}_1, \hat{\sigma}_2) = (0.6270, 0.8663, 1.7917)$. We plot the densities for the normal mixture and Student's $t$ distributions.

They appear almost identical, which is confirmed by a qq-plot.

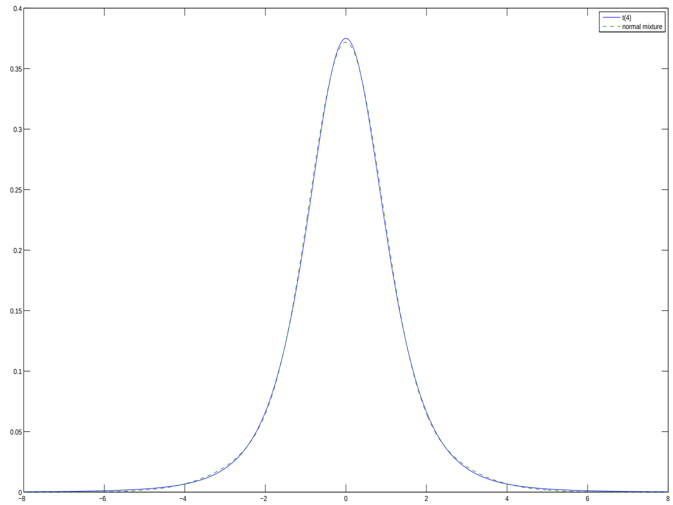


Figure 3: Densities for the normal mixture and Student's $t$ distributions.

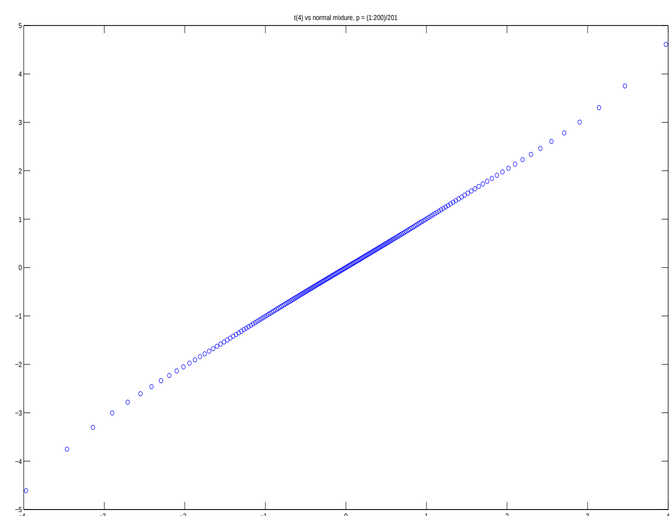


Figure 4: qq-plot of the normal mixture vs Student's $t$ distributions.

However, as we look further out in the tail, it becomes obvious that the Student's $t$ distribution has a heavier tail than the normal mixture. This illustrates the fact that it might be dangerous to draw conclusions about the tail of a distribution from data obtained in the center of the distribution.

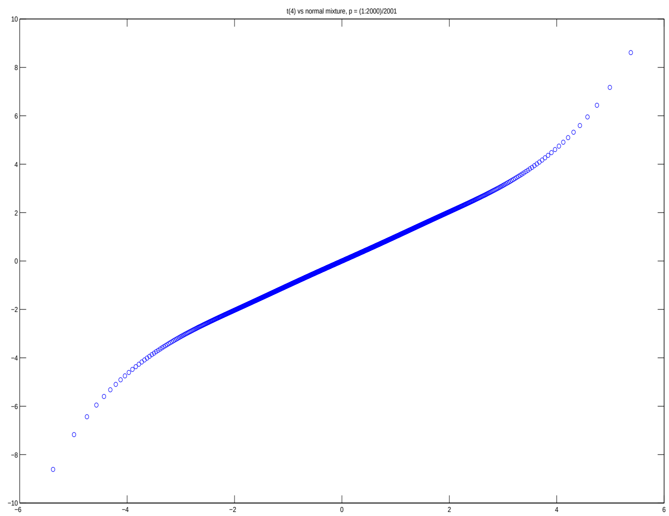


Figure 5: qq-plot of the normal mixture vs Student's $t$ distributions.

To determine the asymptotic behavior of $F(x)$ as $x \to -\infty$, it suffices to find a function $G$ such that

$$\lim_{q \to 0} \frac{F(G^{-1}(q))}{q} = 1, \tag{11}$$

since this implies that

$$\lim_{x \to -\infty} \frac{F(x)}{G(x)} = \lim_{q \to 0} \frac{F(G^{-1}(q))}{G(G^{-1}(q))} = \lim_{q \to 0} \frac{F(G^{-1}(q))}{q} = 1. \tag{12}$$

It is natural to assume that the distribution with the fatter tail will dominate. In this case, it is the distribution with the higher $\sigma$. From now on, we will assume that $\sigma_1 > \sigma_2$, otherwise we can simply rearrange the order. This would imply that

$$F(x) = p\Phi(x/\sigma_1) + (1-p)\Phi(x/\sigma_2) \sim p\Phi(x/\sigma_1). \tag{13}$$

Thus, we assume that $G(x) = p\Phi(x/\sigma_1)$, which is equivalent to $G^{-1}(q) = \sigma_1 \Phi^{-1}(q/p)$. We obtain

$$\lim_{q \to 0} \frac{F(G^{-1}(q))}{q} = \lim_{q \to 0} \frac{p\Phi(\sigma_1 \frac{\Phi^{-1}(q/p)}{\sigma_1}) + (1-p)\Phi(\sigma_1 \frac{\Phi^{-1}(q/p)}{\sigma_2})}{q}. \tag{14}$$

It easily seen that the first term equals 1. If we can show that the second term vanishes, then we have the desired result. Let $z = \Phi^{-1}(q/p)$, or equivalently, $q = p\Phi(z)$.

Then, as $q \to 0$, $z \to -\infty$. For the standard normal distribution function it holds that

$$\Phi(x/\sigma) \sim \frac{\sigma}{-x}\phi(x/\sigma), \tag{15}$$

as $x \to -\infty$, see Example 8.1 for details. We have

$$\begin{aligned}
\lim_{q\to 0}\frac{(1-p)\Phi(\sigma_1\frac{\Phi^{-1}(q/p)}{\sigma_2})}{q} &= \lim_{z\to-\infty}\frac{(1-p)\Phi(\frac{z}{\sigma_2/\sigma_1})}{p\Phi(z)} \sim \frac{(1-p)\frac{\sigma_2/\sigma_1}{-z}\phi(\frac{z}{\sigma_2/\sigma_1})}{p\frac{1}{-z}\phi(z)} && (16)\\
&= C\exp\Big(-\frac{z^2}{2(\sigma_2/\sigma_1)^2}+\frac{z^2}{2}\Big) && (17)\\
&= C\exp\Big(\frac{z^2}{2}(1-\frac{\sigma_1^2}{\sigma_2^2})\Big) \to 0, && (18)
\end{aligned}$$

since by assumption $\sigma_1 > \sigma_2$. We conclude that $G(x) = p\Phi(x/\sigma_1)$.

**Problem 8.2.** Consider the Student's $t$ location-scale family with parameter vector $(\mu, \sigma, \nu)$.

(a) Determine the log-likelihood function and estimate the parameters based on the sample $\{t_4^{-1}(k/201) : k = 1, \dots, 200\}$.

Simulate 3,000 samples of size 200 from the standard Student's $t$ distribution with four degrees of freedom.

(b) For each sample compute the maximum-likelihood estimate of the parameter vector $(\mu, \sigma, \nu)$. Make a scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ and interpret the plot.

(b) For each sample compute the least-squares estimate of the parameter vector $(\mu, \sigma, \nu)$. Make a scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$, interpret the plot, and compare the plot to that in (b).

(b) For each sample compute the sample standard deviation and divide the sample by the sample standard deviation. Consider each rescaled sample to be a sample from a Student's $t$ distribution with unit variance and estimate the degrees-of-freedom parameter by maximum likelihood. Transform the estimates into estimates of the parameter pair $(\sigma, \nu)$ for a centered Student's $t$ distribution with scale parameter $\sigma$. Make a scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$, interpret the plot, and compare the plot to that in (b).

**Solution.** The density of the location-scale Student's $t$ distribution is given by

$$f(x|\mu, \sigma, \nu) = \frac{\Gamma((\nu+1)/2)}{\sigma\sqrt{\nu\pi}\Gamma(\nu/2)}\Big(1 + \frac{(x-\mu)^2}{\nu\sigma^2}\Big)^{-(\nu+1)/2}.$$

The log-likelihood function becomes

$$l(\mu, \sigma, \nu) = \sum_{k=1}^{200} \log\left(\frac{\Gamma((\nu+1)/2)}{\sigma\sqrt{\nu\pi}\Gamma(\nu/2)}\Big(1 + \frac{(x_k-\mu)^2}{\nu\sigma^2}\Big)^{-(\nu+1)/2}\right),$$

where $x_k = t_4^{-1}(k/201)$. Maximizing $l$ numerically gives the parameter estimates $(\hat{\mu}, \hat{\sigma}, \hat{\nu}) = (0, 1.0349, 5.3090)$.

Next, we simulate 3,000 samples of size 200 from the standard Student's $t$ distribution with four degrees of freedom. For each sample, we compute the maximum likelihood estimates $(\hat{\mu}, \hat{\sigma}, \hat{\nu})$. A scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ is presented below.

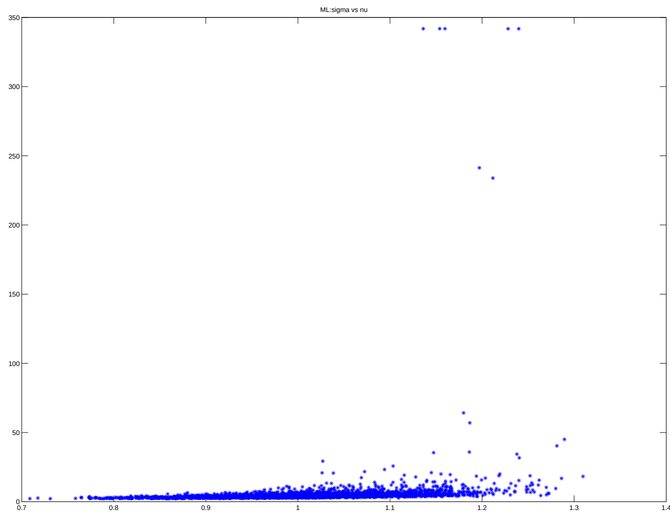


Figure 6: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using ML.

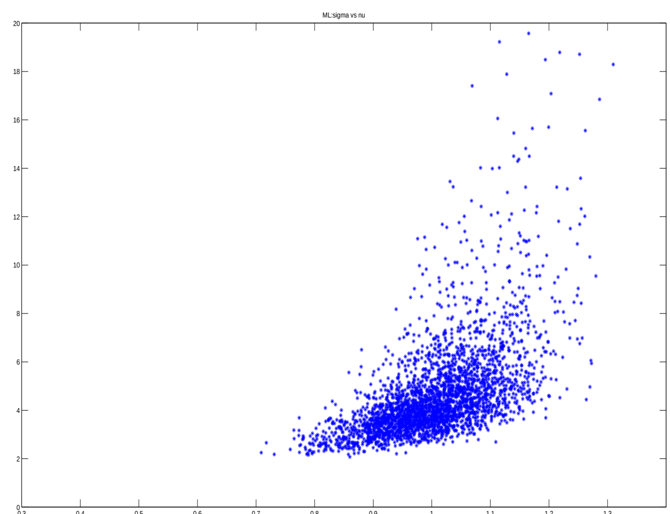


Figure 7: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using ML.

A likelihood surface for $\hat{\sigma}$ vs $\hat{\nu}$ for one sample is plotted below.

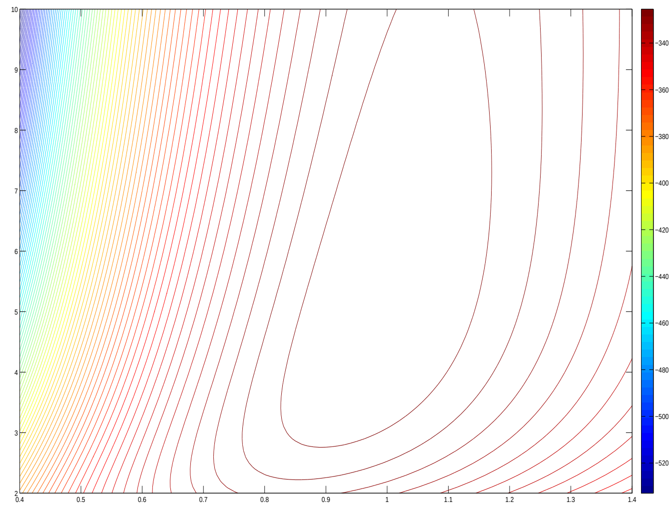

Figure 8: Likelihood surface for $\hat{\sigma}$ vs $\hat{\nu}$ for one sample.

We see that the likelihood surface seems rather flat in the center. You may get quite different optimal values of $(\hat{\sigma}, \hat{\nu})$ for different numerical algorithms.

The least-squares estimates of $(\mu, \sigma, \nu)$ are the values that minimize the sum of the squared deviations between the empirical quantiles and the quantiles of a chosen parametric distribution, formally

$$\sum_{k=1}^{n} \Big( z_{k,n} - F^{-1}\big(\frac{n-k+1}{n+1}\big)\Big)^2. \tag{19}$$

Recall that the distribution function of the location-scale Student's $t$ distribution is given by

$$F(x) = t_\nu\big(\frac{x-\mu}{\sigma}\big),$$

where $t_\nu(x)$ is the standard Student's $t$ distribution function. It follows that the quantile function is given by

$$F^{-1}(p) = \mu + \sigma t_\nu^{-1}(p), \tag{20}$$

where $t_\nu^{-1}(p)$ is the standard Student's $t$ quantile function. Inserting (20) into (19), we obtain the following expression for the sum of squared deviations:

$$\sum_{k=1}^{n} \Big( z_{k,n} - \mu - \sigma t_\nu^{-1}\big(\frac{n-k+1}{n+1}\big)\Big)^2. \tag{21}$$

Minimizing (21) w.r.t. $(\mu, \sigma, \nu)$ gives the least-squares estimates $(\hat{\mu}, \hat{\sigma}, \hat{\nu})$.

We simulate 3,000 samples of size 200 from the standard Student's $t$ distribution with four degrees of freedom. For each sample, we compute the least-squares estimates $(\hat{\mu}, \hat{\sigma}, \hat{\nu})$. A scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ is presented below.

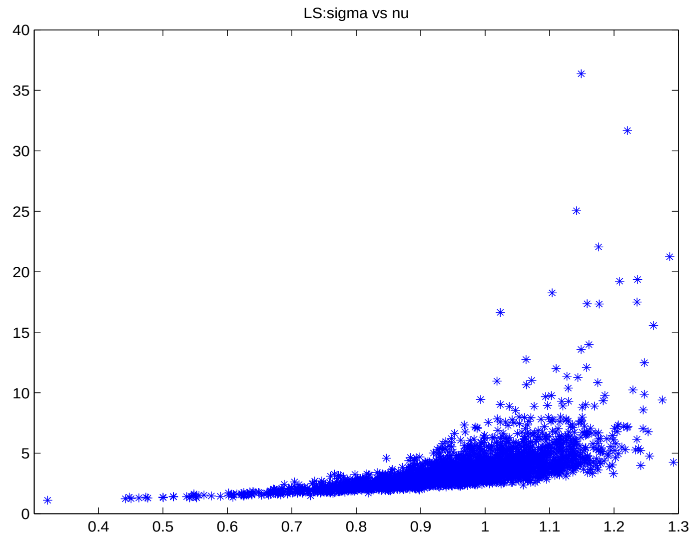


Figure 9: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using LS.

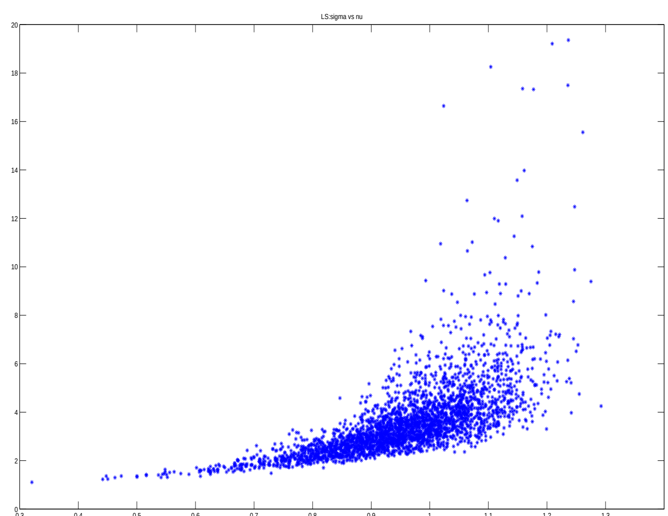

Figure 10: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using LS.

For each sample, we compute the sample standard deviation $s$, and divide the sample by $s$. We consider each rescaled sample to be a sample from a Student's $t$ distribution with unit variance. Recall that a random variable $Y$ with the location-scale Student's $t$ distribution has the representation

$$Y \stackrel{d}{=} \mu + \sigma Z,$$

where $Z$ has a standard Student's $t$ distribution. To obtain a distribution with unit variance, we must have

$$1 = Var(Y) = Var(\mu + \sigma Z) = \sigma^2 Var(Z) = \sigma^2 \frac{\nu}{\nu - 2}$$

which yields $\sigma = \sqrt{\frac{\nu-2}{\nu}}$. Using this, the log-likelihood function becomes

$$l(\mu, \nu) = \sum_{k=1}^{200} \log \left( \frac{\Gamma((\nu+1)/2)}{\sqrt{(\nu-2)\pi}\Gamma(\nu/2)} \left(1 + \frac{(x_k - \mu)^2}{\nu - 2}\right)^{-(\nu+1)/2} \right).$$

Maximizing $l$ yields the degrees-of-freedom estimate $\hat{\nu}$. To find the estimate of the scale parameter $\sigma$, consider again

$$Var(Y) = \sigma^2 \frac{\nu}{\nu - 2},$$

which is equivalent to

$$\sigma = \sqrt{\frac{\nu - 2}{\nu} Var(Y)}.$$

We estimate $\sigma$ with

$$\hat{\sigma} = s\sqrt{\frac{\hat{\nu} - 2}{\hat{\nu}}}$$

for each sample. A scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ is presented below.

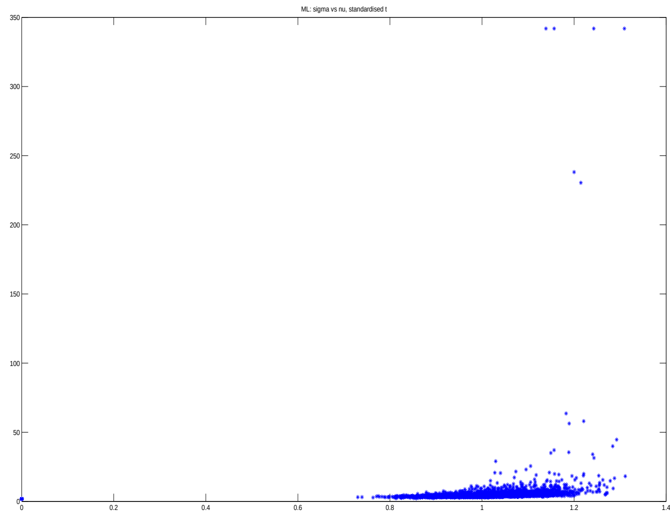


Figure 11: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using ML.

There is something strange with this picture: we have some observations near the point $(0, 0)$. If we zoomed in, we would see that these points had $\hat{\nu} < 2$. Since we have $\sigma = \sqrt{\frac{\nu-2}{\nu}}$, this should be impossible. We must take care that $\hat{\nu}$ only takes values larger than 2 in our optimization procedure. Maximizing $l$ with the constraint $\hat{\nu} > 2$ for each sample yields the following scatter plot.

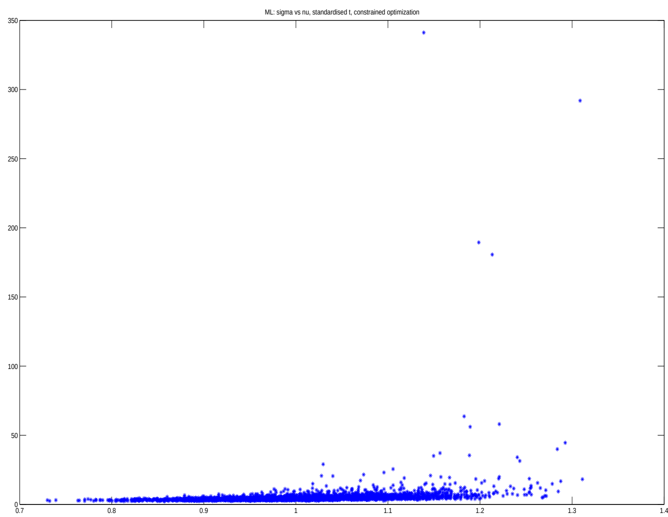


Figure 12: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using ML.

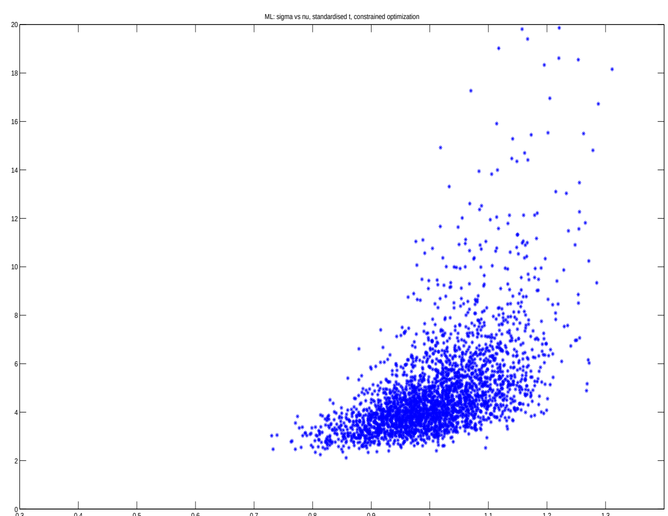


Figure 13: Scatter plot of the 3,000 parameter estimates $(\hat{\sigma}, \hat{\nu})$ using ML.

Clearly, our numerical problem is gone. It is however not so clear whether this two-step fitting algorithm gave any improvement over standard maximum likelihood.

**Problem 8.3.** Let $X$ be $\mathrm{LN}(\mu, \sigma^2)$-distributed.

(a) Show that, as $x \to \infty$,

$$P(X > x) \sim \frac{\sigma}{\sqrt{2\pi}(\log x - \mu)} \exp\Big( - \frac{(\log x - \mu)^2}{2\sigma^2}\Big)$$

(b) Use the result in (a) to show that, for any $\lambda, \alpha > 0$,

$$\lim_{x\to\infty} \frac{P(X > x)}{e^{-\lambda x}} = \infty \qquad \text{and} \qquad \lim_{x\to\infty} \frac{P(X > x)}{x^{-\alpha}} = 0.$$

**Solution.** $X$ has the representation

$$X \stackrel{d}{=} \exp(\mu + \sigma Z), \qquad Z \sim N(0, 1).$$

Using this, we have

$$\begin{aligned} P(X > x) &= 1 - P(X \leq x) = 1 - P(\exp(\mu + \sigma Z) \leq x) = 1 - P(Z \leq \frac{\log x - \mu}{\sigma}) \\ &= 1 - \Phi(\frac{\log x - \mu}{\sigma}) = \Phi(-\frac{\log x - \mu}{\sigma}) \end{aligned}$$

For the standard normal distribution function it holds that

$$\Phi(x) \sim \frac{1}{-x}\phi(x), \tag{22}$$

as $x \to -\infty$, see Example 8.1 for details.

Now, as $x \to \infty$, $-\frac{\log x - \mu}{\sigma} \to -\infty$. It follows that

$$\Phi(-\frac{\log x - \mu}{\sigma}) \sim \frac{1}{-(-\frac{\log x - \mu}{\sigma})}\phi(-\frac{\log x - \mu}{\sigma}) = \frac{\sigma}{\sqrt{2\pi}(\log x - \mu)} \exp\Big( - \frac{(\log x - \mu)^2}{2\sigma^2}\Big).$$

Using this result,

$$\begin{aligned} \lim_{x\to\infty} \frac{P(X > x)}{e^{-\lambda x}} &= \lim_{x\to\infty} \frac{\frac{\sigma}{\sqrt{2\pi}(\log x - \mu)} \exp(-\frac{(\log x - \mu)^2}{2\sigma^2})}{\exp(-\lambda x)} \\ &= \lim_{x\to\infty} C \exp\Big( - \frac{(\log x - \mu)^2}{2\sigma^2} + \lambda x\Big)(\log x - \mu)^{-1} \\ &= \lim_{x\to\infty} C \exp\Big( - \frac{(\log x - \mu)^2}{2\sigma^2} + \lambda x - \log(\log x - \mu)\Big). \end{aligned}$$

Now, it is well known that $x$ dominates $\log x$, which implies that $x$ also dominates $\log(\log x)$. Further, to see that $x$ dominates $(\log x)^2$, let $y = \log x$, and recall that $e^y$ dominates $y^2$. Thus, the expression in the exponent goes to $\infty$, and it follows that

$$\lim_{x\to\infty} \frac{P(X > x)}{e^{-\lambda x}} = \infty.$$

Again using the result from (a),

$$\begin{aligned}
\lim_{x\to\infty} \frac{P(X>x)}{x^{-\alpha}} &= \lim_{x\to\infty} \frac{\sigma}{\sqrt{2\pi}(\log x-\mu)} \exp\Big(-\frac{(\log x-\mu)^2}{2\sigma^2}\Big) x^{\alpha} \\
&= \lim_{x\to\infty} C \exp\Big(-\frac{(\log x-\mu)^2}{2\sigma^2} + \alpha\log x - \log(\log x-\mu)\Big).
\end{aligned}$$

The dominating term is $(\log x)^2$. It follows that the exponent goes to $-\infty$, and

$$\lim_{x\to\infty} \frac{P(X>x)}{x^{-\alpha}} = 0.$$

Thus, we have shown that the log-normal tail is heavier than every exponential tail, but lighter than any polynomial tail.

# 9 Multivariate Models

**Problem 9.1.**

**Solution.** Let $a = (h_1, \dots, h_d, -1)$ and $Y = (X_1, \dots, X_d, L)$. $Y$ has an elliptical distribution, that is

$$Y = \mu + AZ, \quad a^T Y = a^T \mu + a^T AZ \stackrel{d}{=} a^T \mu + \sqrt{a^T \Sigma a} Z_1,$$

where $\Sigma = AA^T$ and $Z$ has a spherical distribution. From Proposition 3.2, the portfolio weights that minimize $E[(h_0 + a^T Y)^2]$ must satisfy $E[h_0 + a^T Y] = h_0 + a^T \mu = 0$. Thus, we have

$$\begin{aligned} E[(h_0 + a^T Y)^2] &= (E[h_0 + a^T Y])^2 + Var(h_0 + a^T Y) \\ &= Var(\sqrt{a^T \Sigma a} Z_1) = a^T \Sigma a Var(Z_1). \end{aligned}$$

We see that the optimal quadratic hedge is the vector that minimizes $a^T \Sigma a$. Now, for any positive homogeneous risk measure $\rho$, we have

$$\rho(h_0 + a^T Y) = \rho(h_0 + a^T \mu + a^T AZ) = \rho(a^T AZ) = \sqrt{a^T \Sigma a} \rho(Z_1).$$

Thus, the vector $a$ that minimizes $E[(h_0 + a^T Y)^2]$ also minimizes $\rho(h_0 + a^T Y)$.

**Problem 9.2.**

**Solution.** $X$ and $Y$ have the representations

$$X \stackrel{d}{=} R_0\mathbf{1} + W_x AZ, \quad Y \stackrel{d}{=} \mathbf{1}R_0 + W_y AZ,$$

where $AA^T = \Sigma$ is a common dispersion matrix, $Z \sim N_d(0, I)$ and $W_x$ and $W_y$ are non-negative random variables. The portfolio values at the end of the investment period, denoted $V_X(w)$ and $V_Y(w)$, can be written as

$$V_X(w) = w^T X \stackrel{d}{=} w^T(R_0\mathbf{1} + W_x AZ) = V_0 R_0 + W_x w^T AZ \stackrel{d}{=} V_0 R_0 + W_x \sqrt{w^T AA^T w} Z_1,$$

and similar for $V_Y(w)$. Thus, for a positive homogeneous risk measure $\rho$ and a positive semi-definite dispersion matrix $AA^T$, we have

$$\frac{\rho(V_X(w) - V_0R_0)}{\rho(V_Y(w) - V_0R_0)} = \frac{\rho(V_0R_0 + W_x\sqrt{w^T AA^T w}Z_1 - V_0R_0)}{\rho(V_0R_0 + W_y\sqrt{w^T AA^T w}Z_1 - V_0R_0)} = \frac{\sqrt{w^T AA^T w}\rho(W_x Z_1)}{\sqrt{w^T AA^T w}\rho(W_y Z_1)} = \frac{\rho(W_x Z_1)}{\rho(W_y Z_1)}.$$

If, in particular, $X$ has a Student's $t$ distribution with four degrees of freedom, $Y$ has a normal distribution, and $\rho$ is given by $\mathrm{VaR}_p$, then

$$\frac{VaR_p(V_X(w) - V_0R_0)}{VaR_p(V_Y(w) - V_0R_0)} = \frac{VaR_p(W_x Z_1)}{VaR_p(Z_1)} = \frac{t_4^{-1}(p)}{\Phi^{-1}(p)}.$$

**Problem 9.3.**

**Solution.** The Gaussian copula for the pair $(X_1, X_2)$, with common distribution function $t_4$, can be written

$$C_\rho^{Ga}(F_1(x_1), F_2(x_2)) = \Phi_\rho^2(\Phi^{-1}(t_4(x_1)), \Phi^{-1}(t_4(x_2))),$$

where $\rho$ is the linear correlation. Note that,under the Gaussian copula, the pair $(\Phi^{-1}(t_4(X_1)), \Phi^{-1}(t_4(X_2)))$ has a bivariate normal distribution. Using, in turn, the probability and quantile transforms, we obtain

$$\lim_{x\to\infty} P(X_2 > x | X_1 > x) = \lim_{x\to\infty} P(\Phi^{-1}(t_4(X_2)) > \Phi^{-1}(t_4(x)) | \Phi^{-1}(t_4(X_1)) > \Phi^{-1}(t_4(x))).$$

As $x \to \infty$, $\Phi^{-1}(t_4(x)) \to \infty$, so we may rewrite the above as

$$\lim_{z\to\infty} P(\Phi^{-1}(t_4(X_2)) > z | \Phi^{-1}(t_4(X_1)) > z).$$

It follows from the symmetry of elliptical distributions that

$$\lim_{z\to\infty} P(\Phi^{-1}(t_4(X_2)) > z | \Phi^{-1}(t_4(X_1)) > z) = \lim_{z\to-\infty} P(\Phi^{-1}(t_4(X_2)) \le z | \Phi^{-1}(t_4(X_1)) \le z).$$

Finally, using Proposition 9.5, we have

$$\lim_{x\to\infty} P(X_2 > x | X_1 > x) = \lim_{z\to-\infty} P(\Phi^{-1}(t_4(X_2)) \le z | \Phi^{-1}(t_4(X_1)) \le z) = 0.$$

The Student's $t$ copula for the pair $(X_1, X_2)$, with common distribution function $t_4$, can be written

$$C_{\nu,\rho}^t(F_1(x_1), F_2(x_2)) = t_{6,\rho}^2(t_6^{-1}(t_4(x_1)), t_6^{-1}(t_4(x_2))).$$

Note that, under the Student's $t$ copula, the pair $(t_6^{-1}(t_4(X_1)), t_6^{-1}(t_4(X_2)))$ has a bivariate Student's $t$ distribution with $\nu = 6$ degrees of freedom. Using, in turn, the probability and quantile transforms, we obtain

$$\lim_{x\to\infty} P(X_2 > x | X_1 > x) = \lim_{x\to\infty} P(t_6^{-1}(t_4(X_2)) > t_6^{-1}(t_4(x)) | t_6^{-1}(t_4(X_1)) > t_6^{-1}(t_4(x))).$$

Again using the symmetry of elliptical distributions, we may rewrite the above with $z = t_6^{-1}(t_4(x))$ as

$$\lim_{z\to-\infty} P(t_6^{-1}(t_4(X_2)) \le z | t_6^{-1}(t_4(X_1)) \le z).$$

Since the $t_6$-distribution is regularly varying with tail index $\alpha = 6$, in follows from Proposition 9.6 that

$$\begin{aligned}\lim_{x\to\infty} P(X_2 > x | X_1 > x) &= \lim_{z\to-\infty} P(t_6^{-1}(t_4(X_2)) \le z | t_6^{-1}(t_4(X_1)) \le z) \\ &= \frac{\int_{(\pi/2 - arcsin\rho)/2}^{\pi/2} cos^6 t dt}{\int_0^{\pi/2} cos^6 t dt} \approx 0.17\end{aligned}$$

**Problem 9.4.**

**Solution.** For comonotone random variables $X_1$ and $X_2$ with distribution functions $F_1$ and $F_2$ we can write

$$(X_1, X_2) = (X_1, F_2^{-1}(F_1(X_1))).$$

Thus, we have

$$VaR_p(X_1 + X_2) = -F^{-1}_{X_1+X_2}(p) = -F^{-1}_{X_1+F_2^{-1}(F_1(X_1))}(p). \tag{23}$$

The function $x + F_2^{-1}(F_1(x))$ is non-decreasing in $x$, and if we assume that $F_1$ and $F_2$ are continuous, it follows from Proposition 6.3 that (23) equals

$$-(F_1^{-1}(p) + F_2^{-1}(F_1(F_1^{-1}(p))) = -F_1^{-1}(p) - F_2^{-1}(p) = VaR_p(X_1) + VaR_p(X_2),$$

which shows that $VaR_p$ is additive for comonotone random variables.

Using this result, we have, for any spectral risk measure $\rho_\phi$,

$$\begin{aligned}
\rho_\phi(X_1 + X_2) &= -\int_0^1 \phi(u) F^{-1}_{X_1+X_2}(u)du = -\int_0^1 \phi(u)(F_1^{-1}(u) + F_2^{-1}(u))du \\
&= -\int_0^1 \phi(u)(F_1^{-1}(u))du - \int_0^1 \phi(u)(F_2^{-1}(u))du = \rho_\phi(X_1) + \rho_\phi(X_2).
\end{aligned}$$

**Problem 9.5.**

**Solution.** Let $(U_1', U_2')$ be an independent copy of $(U_1, U_2)$. Recall that Kendall's tau is defined as

$$\tau(U_1, U_2) = P((U_1 - U_1')(U_2 - U_2') > 0) - P((U_1 - U_1')(U_2 - U_2') < 0).$$

If $(U_1, U_2)$ does not have a point mass anywhere, this expression simplifies to

$$\tau(U_1, U_2) = 2P((U_1 - U_1')(U_2 - U_2') > 0) - 1.$$

Further,

$$P((U_1 - U_1')(U_2 - U_2') > 0) = P(U_1 - U_1' > 0, U_2 - U_2' > 0) + P(U_1 - U_1' < 0, U_2 - U_2' < 0).$$

We have

$$\begin{aligned} P(U_1 - U_1' < 0, U_2 - U_2' < 0) &= P(U_1 < U_1', U_2 < U_2') = \int P(U_1 \le u_1, U_2 \le u_2) dC(u_1, u_2) \\ &= \int C(u_1, u_2) dC(u_1, u_2) = E[C(U_1, U_2)]. \end{aligned}$$

Similarly,

$$\begin{aligned} P(U_1 - U_1' > 0, U_2 - U_2' > 0) &= 1 - P(U_1 < U_1') - P(U_2 < U_2') + P(U_1 < U_1', U_2 < U_2') \\ &= 1 - 0.5 - 0.5 + E[C(U_1, U_2)] = E[C(U_1, U_2)], \end{aligned}$$

and it follows that

$$\tau(U_1, U_2) = 2(E[C(U_1, U_2)] + E[C(U_1, U_2)]) - 1 = 4E[C(U_1, U_2)] - 1.$$

Now, recall that the expected value of a random variable $X$ on $[0, 1]$ can be written

$$E[X] = \int_0^1 x dF(x) = \int_0^1 \int_0^x dt dF(x) = \int_0^1 \int_t^1 dF(x) dt = \int_0^1 P(X \ge t) dt.$$

Using this relation, we obtain

$$\begin{aligned} E[C(U_1, U_2)] &= \int_0^1 P(C(U_1, U_2) > t) dt = \int_0^1 (1 - t + \frac{\Psi^{-1}(t)}{(\Psi^{-1})'(t)}) dt \\ &= 1 - \frac{1}{2} + \int_0^1 \frac{\Psi^{-1}(t)}{(\Psi^{-1})'(t)} dt, \end{aligned}$$

which yields

$$\tau(U_1, U_2) = 4E[C(U_1, U_2)] - 1 = 4(\frac{1}{2} + \int_0^1 \frac{\Psi^{-1}(t)}{(\Psi^{-1})'(t)} dt) - 1 = 1 + 4\int_0^1 \frac{\Psi^{-1}(t)}{(\Psi^{-1})'(t)} dt.$$

For the special case of the Clayton copula, we have from Example 9.16 that

$$\psi^{-1}(u) = u^{-\theta} - 1, \qquad (\psi^{-1})'(u) = -\theta u^{-\theta-1}.$$

It follows that

$$\tau(U_1, U_2) = 1 + 4\int_0^1 \frac{\Psi^{-1}(t)}{(\Psi^{-1})'(t)} dt = 1 + 4\int_0^1 \frac{t^{-\theta} - 1}{-\theta t^{-\theta-1}} dt = 1 - \frac{2}{\theta} + \frac{4}{\theta(\theta+2)} = \frac{\theta}{\theta+2}.$$

**Problem 9.6.**

**Solution.** The distribution function can be obtained from Table 4.1 simply by summing up the cells, e.g.

$$P(X_1 \leq 1, X_2 \leq 3) = P(X_1 = 1, X_2 = 1) + P(X_1 = 1, X_2 = 2) + P(X_1 = 1, X_2 = 3).$$

Repeating this for all cells gives the distribution function on matrix form as

| $x_1 \backslash x_2$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 0.098736 | 0.099792 | 0.099842 | 0.099842 |
| 2 | 0.731454 | 0.830309 | 0.849379 | 0.850300 |
| 3 | 0.796051 | 0.938708 | 0.976856 | 0.980003 |
| 4 | 0.800633 | 0.950533 | 0.995117 | 1 |

Table 1: Distribution function $F(x_1, x_2)$.

To obtain the copula $C$ defined by $C(F_1(x_1), F_2(x_2)) = F(x_1, x_2)$, simply change the axis values from $(x_1, x_2)$ to $(F_1(x_1), F_2(x_2))$, e.g. $C(F_1(2), F_2(3)) = F(2, 3)$. This gives the copula in matrix form as

| $F_1(x_1) \backslash F_2(x_2)$ | 0.800633 | 0.950533 | 0.995117 | 1 |
|---|---|---|---|---|
| 0.099842 | 0.098736 | 0.099792 | 0.099842 | 0.099842 |
| 0.850300 | 0.731454 | 0.830309 | 0.849379 | 0.850300 |
| 0.980003 | 0.796051 | 0.938708 | 0.976856 | 0.980003 |
| 1 | 0.800633 | 0.950533 | 0.995117 | 1 |

Table 2: The copula $C(F_1(x_1), F_2(x_2))$.

The above copula can be approximated by a Gaussian copula, and the correlation parameter $\rho$ is estimated using least-squares, that is $\rho$ is chosen as to minimize

$$\sum_{(u,v)} (\Phi_\rho^2(\Phi^{-1}(u), \Phi^{-1}(v)) - C(u, v))^2.$$

The estimated linear correlation is $\rho = 0.5984$.

**Problem 9.7.**

**Solution.** We seek a function $g$ such that $P(X_k = 1|g(Y) = \theta) = \theta$. We have

$$\begin{aligned}P(X_k = 1|g(Y) = \theta) &= P(X_k = 1|Y = g^{-1}(\theta)) = P(\sqrt{\rho}g^{-1}(\theta) + \sqrt{1-\rho}Y_k \leq \Phi^{-1}(p)) \\ &= P(Y_k \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}g^{-1}(\theta)}{\sqrt{1-\rho}}) = \Phi(\frac{\Phi^{-1}(p) - \sqrt{\rho}g^{-1}(\theta)}{\sqrt{1-\rho}}).\end{aligned}$$

Setting this expression equal to $\theta$ and substituting $\theta$ for $g(Y)$ yields

$$g(Y) = \Phi(\frac{\Phi^{-1}(p) - \sqrt{\rho}Y}{\sqrt{1-\rho}}).$$

To find the $q$-quantile of $g(Y)$, we first note that $g$ is decreasing. Propositions 6.3-6.4 yield

$$\begin{aligned}F^{-1}_{g(Y)}(q) &= -F^{-1}_{-g(Y)}(1-q) = -(-g(F^{-1}_Y(1-q))) = g(\Phi^{-1}(1-q)) \\ &= \Phi(\frac{\Phi^{-1}(p) - \sqrt{\rho}\Phi^{-1}(1-q)}{\sqrt{1-\rho}}) = \Phi(\frac{\Phi^{-1}(p) + \sqrt{\rho}\Phi^{-1}(q)}{\sqrt{1-\rho}}).\end{aligned}$$

Consider the aforementioned portfolio of $n$ =1,000 loans, and define the number of defaults $D_n = \sum_{k=1}^n X_k$. Then, the one-year profit $S_n$ of the portofolio is

$$S_n = 10,000(n - D_n) - 0.25 \cdot 1,000,000 S_n = 10,000n - 260,000D_n.$$

Further, the one-year Expected Shortfall is given by

$$ES_p(S_n) = \frac{1}{0.01}\int_0^{0.01} VaR_u(S_n)du,$$

with

$$VaR_u(S_n) = VaR_u(10,000n - 260,000D_n) = -\frac{10,000n}{R_0} + 260,000VaR_u(-D_n).$$

To evaluate the above expression, we must resort to simulations or approximations. We choose the latter, and consider the case where $n$ is large. Indeed, it follows from the conditional law of large numbers that, conditional on $Y$,

$$\frac{D_n}{n} \to P(X_k = 1|Y) = g(Y) \quad a.s.$$

Thus, we may, for large $n$, approximate $D_n$ by

$$D_n \approx ng(Y).$$

Using this approximation,

$$VaR_u(-D_n) = F^{-1}_{D_n/R_0}(1-u) \approx F^{-1}_{ng(Y)/R_0}(1-u) = \frac{n}{R_0}F^{-1}_{g(Y)}(1-u) = \frac{n}{R_0}\Phi(\frac{\Phi^{-1}(p) + \sqrt{\rho}\Phi^{-1}(1-u)}{\sqrt{1-\rho}}).$$

Finally, we obtain an approximate $ES_p(S_n)$ as

$$ES_p(S_n) \approx -\frac{10,000n}{R_0} + \frac{260,000n}{0.01R_0} \int_0^{0.01} \Phi(\frac{\Phi^{-1}(p) + \sqrt{\rho}\Phi^{-1}(1-u)}{\sqrt{1-\rho}})du,$$

which can be integrated numerically. We find that $ES_p(S_n) \approx 46.8$ millions, or $4,68\%$ of the capital.

**Problem 9.8.**

**Solution.** The portfolio weights $w_1$ and $w_2$ satisfy the following system of equations

$$\begin{aligned} w_1 + w_2 &= V_0 \\ w_1 E[R_1] + w_2 E[R_2] &= 1.06 V_0, \end{aligned}$$

where $R_0$ and $R_1$ denote the return on the bond and stock portoflios, respectively, and $V_0$is the initial capital. The system admits the solution $w_1 = w_2 = \frac{1}{2}$. Let $w = (w_1, w_2, -1)$ and $X = (R_1, R_2, L)$. By assumption, $X$ has a multivariate Student's $t$ distribution with $\nu = 4$, and it follows that

$$A - L = w^T X \stackrel{d}{=} w^T(\mu + AZ) \stackrel{d}{=} w^T\mu + \sqrt{w^T A A^T w} Z_1,$$

where $Z$ has a multivariate standard Student's $t$ distribution. Denoting $AA^T$ by $\Sigma$, we have, under the assumption that the risk-free return $R_0 = 1$, that

$$\begin{aligned} VaR_{0.005}(A - L) &= VaR_{0.005}(w^T\mu + \sqrt{w^T\Sigma w} Z_1) \\ &= -w^T\mu + \sqrt{w^T\Sigma w} VaR_{0.005}(Z_1) \\ &= -w^T\mu + \sqrt{w^T\Sigma w} t_4^{-1}(0.995). \end{aligned}$$

The dispersion matrix is given by

$$\Sigma_{i,j} = Cor(X_i, X_j)\frac{\nu - 2}{\nu}\sqrt{Var(X_i)Var(X_j)}.$$

We evaluate the risk numerically and obtain $VaR_{0.005}(A-L) \approx -920,000$, which means that the insurer is solvent.

Next, we consider an instantaneous decline of 15% in the value of the stock market portfolio. Immediately after the shock, the portfolio weights $w$ are $(\frac{V_0}{2}, 0.85\frac{V_0}{2}, -1) = (0.5V_0, 0.425V_0, -1)$. Re-evaluating the risk numerically yields $VaR_{0.005}(A-L) \approx 13,000$, which means that the insurer is no longer solvent. To achieve solvency, the insurer wishes to rebalance the portfolio with weights $\tilde{w}_1$ and $\tilde{w}_2$ so that

$$VaR_{0.005}(A - L) = 0,$$

under the constraint $\tilde{w}_1 + \tilde{w}_2 = \tilde{V}_0$, where $\tilde{V}_0 = \frac{V_0}{2} + 0.85\frac{V_0}{2}$. Solving numerically for $\tilde{w}$, we obtain

$$(\tilde{w}_1, \tilde{w}_2) = (0.5132V_0, 0.4118V_0),$$

so the insurer should reduce the exposure to the stock market in favour of the bond market. The expected return of the adjusted asset portfolio is

$$\frac{E[A]}{\tilde{V}_0} = \frac{\tilde{w}_1 E[R_1] + \tilde{w}_2 E[R_2]}{\tilde{V}_0} = 1.0556,$$

slightly lower than the initial target return of 1.06.

# References

[1] H. Hult et. al., Risk and Portfolio Analysis; Principles and Methods.