Prof. P. Cheridito 13 August 2019

Quantitative Risk Management

Important:

- Put your student card on the table
- Begin each exercise on a new sheet of paper, and write your name on each sheet
- Only pen, paper and ten sides of summary are allowed

Please fill in the following table.

Last name	
First name	
Student number (if available)	

Please do not fill in the following table.

Question	Points	Control	Maximum
#1			11
#2			10
#3			8
#4			11
#5			10
Total			50

Question 1 (11 Pts)

a) Let X be a random variable with cdf

$$F(x) = \frac{1}{1+e^{-\frac{x-\mu}{\sigma}}}, \quad x \in \mathbb{R},$$

for parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. Calculate $\mathrm{VaR}_\alpha(X)$ and $\mathrm{ES}_\alpha(X)$ for $\alpha \in (0,1)$. (3 Pts)

b) Denote by $L^2(\Omega, \mathcal{F}, \mathbb{P})$ the space of all square-integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Which axioms of coherence does the mapping $\rho\colon L^2(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$, given by

$$\rho(X) := \mathrm{sd}(X) = \sqrt{\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]},$$

satisfy? Please, prove your statements. (4 Pts)

c) Construct a two-dimensional random vector (X_1, X_2) such that

(i) $X_i \sim \mathrm{Exp}(\lambda_i)$ for $\lambda_i > 0$, $i = 1, 2$, and

(ii) $\mathrm{VaR}_\alpha(X_1 + X_2) = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2)$ for all $\alpha \in (0,1)$. (4 Pts)

Solution 1

a) In order to compute $\mathrm{VaR}_\alpha(X)$, we simply invert the cdf and obtain

$$\mathrm{VaR}_\alpha(X) = q_X(\alpha) = \mu + \sigma \log\left(\frac{\alpha}{1-\alpha}\right).$$

As for $\mathrm{AVaR}_\alpha(X)$, we note that since X is continuously distributed, we have

$$\mathrm{ES}_\alpha(X) = \mathrm{AVaR}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)\,du = \mu + \sigma\frac{1}{1-\alpha}\int_\alpha^1 \log\left(\frac{u}{1-u}\right) du.$$

Changing the variable to $z = \frac{u}{1-u}$ and subsequently applying the integration by parts formula, we obtain

$$\begin{aligned}
\mathrm{ES}_\alpha(X) &= \mu + \sigma\frac{1}{1-\alpha}\int_\alpha^1 \log(z)\frac{1}{(1+z)^2}dz \\
&= \mu + \sigma\frac{1}{1-\alpha}\left(\left[-\log(z)\frac{1}{1+z}\right]_{z=\frac{\alpha}{1-\alpha}}^\infty + \int_{\frac{\alpha}{1-\alpha}}^\infty \frac{1}{z}\frac{1}{1+z}dz\right) \\
&= \mu + \sigma\frac{1}{1-\alpha}\left((1-\alpha)\log\left(\frac{\alpha}{1-\alpha}\right) + \left[\log\left(\frac{z}{1+z}\right)\right]_{z=\frac{\alpha}{1-\alpha}}^\infty\right) \\
&= \mu + \sigma\frac{1}{1-\alpha}\left((1-\alpha)\log\left(\frac{\alpha}{1-\alpha}\right) - \log(\alpha)\right) \\
&= \mu + \sigma\log\left(\frac{\alpha}{1-\alpha}\right) - \sigma\frac{\log(\alpha)}{1-\alpha}.
\end{aligned}$$

b) Let and $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. If $\lambda \in \mathbb{R}_+$, then

$$\mathrm{sd}(\lambda X) = \sqrt{\mathrm{Var}(\lambda X)} = \sqrt{\lambda^2\,\mathrm{Var}(X)} = \lambda \mathrm{sd}(X),$$

so standard deviation is positive homogeneous. Now, let $m \in \mathbb{R}$. We have that

$$\mathrm{sd}(X+m) = \sqrt{\mathrm{Var}(X+m)} = \sqrt{\mathrm{Var}(X)} = \mathrm{sd}(X),$$

so standard deviation is not translation invariant. Since

$$\rho(X,Y) \le 1,$$

by Cauchy–Schwarz inequality, we also have that

$$\begin{aligned}
\mathrm{sd}(X+Y) &= \sqrt{\mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X,Y)} \\
&= \sqrt{\mathrm{Var}(X) + \mathrm{Var}(Y) + 2\rho(X,Y)\mathrm{sd}(X)\mathrm{sd}(Y)} \\
&\le \sqrt{\mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{sd}(X)\mathrm{sd}(Y)} = \sqrt{\big(\mathrm{sd}(X)+\mathrm{sd}(Y)\big)^2} \\
&= \mathrm{sd}(X) + \mathrm{sd}(Y).
\end{aligned}$$

So standard deviation is subadditive. Lastly, let $X \sim \mathrm{Unif}(a,b)$ and $Y \sim \mathrm{Unif}(c,d)$ for some $a < b < c < d$, which guarantees that $X \le Y$ a.s. But it is obvious that if $b - a \ge d - c$, then will have $\mathrm{sd}(X) \ge \mathrm{sd}(Y)$, which disproves the monotonicity of standard deviation. In summary, standard deviation is positive homogeneous and subadditive, but not translation invariant and monotone.

c) The most direct solution is as follows. Let $U \sim \mathrm{Unif}(0,1)$ and let $q_i(u) = -\frac{1}{\lambda_i}\log(1-u)$, $i = 1,2$. Since, q_i is the quantile function of $\mathrm{Exp}(\lambda_i)$ distribution, we have by the quantile transformation lemma that $q_i(U) \sim \mathrm{Exp}(\lambda_i)$. Setting $X_i = q_i(U)$, we have $X_1 + X_2 = q_1(U) + q_2(U) = (q_1 + q_2)(U)$. The quantile transformation then implies that $q_1 + q_2$ is the quantile function of $X_1 + X_2$. We thus have by the definition of VaR that

$$\mathrm{VaR}_\alpha(X_1 + X_2) = (q_1 + q_2)(\alpha) = q_1(\alpha) + q_2(\alpha) = \mathrm{VaR}_\alpha(X_1) + \mathrm{VaR}_\alpha(X_2).$$

Alternatively, we know from the lecture that VaR is comonotone additive, so the copula of (X_1, X_2) needs to be the comonotonicity copula $M(u,v) = \min\{u,v\}$. We also know from the lecture that M is the copula of the random vector (U,U), where $U \sim \mathrm{Unif}(0,1)$. Since the quantile function of $\mathrm{Exp}(\lambda_i)$ distribution is given by $q_i(u) = -\frac{1}{\lambda_i}\log(1-u)$, we can set

$$(X_1, X_2) := \left(-\frac{1}{\lambda_1}\log(1-U), -\frac{1}{\lambda_2}\log(1-U)\right).$$

We then have that $X_i \sim \mathrm{Exp}(\lambda_i)$ by the quantile transformation lemma, and the comonotonicity of X_1 and X_2 follows from the invariance of copulas under strictly increasing transformations.

Lastly, one could describe the random vector by its distribution. Denoting the marginal cdfs of X_1 and X_2 by F_1 and F_2 respectively, it follows from Sklar's theorem that the cdf F of (X_1, X_2) reads

$$\begin{aligned}
F(x_1,x_2) &= M\big(F_1(x_1), F_2(x_2)\big) = \min\big\{1 - \exp(-\lambda_1 x_1), 1 - \exp(-\lambda_2 x_2)\big\} \\
&= 1 - \max\big\{\exp(-\lambda_1 x_1), \exp(-\lambda_2 x_2)\big\} = 1 - \exp\big(-\min\{\lambda_1 x_1, \lambda_2 x_2\}\big)
\end{aligned}$$

for all $x_1, x_2 > 0$.

Question 2 (10 Pts)

a) Let X be a d-dimensional random vector with a $t_d(\nu, 0, \Sigma)$-distribution for $d \ge 2$, $\nu > 0$ and a positive definite $d \times d$-matrix Σ. Are the components of X exchangeable? (3 Pts)

b) Let $X \sim S_d(\psi)$ for some $d \ge 2$. Show that all univariate marginal distributions of X are equal. (3 Pts)

c) Denote by $\mathbb{S}_+^2$ the set of all positive semidefinite symmetric 2×2-matrices, and let X be a two-dimensional random vector with a $N_2(\mu, \Sigma)$-distribution for a fixed mean vector $\mu \in \mathbb{R}^2$ and a covariance matrix Σ in the set

$$S = \left\{ \Sigma \in \mathbb{S}_+^2 : \Sigma_{ii} = \sigma_i^2 \text{ for } i = 1, 2, \text{ and } \underline{\rho} \le \frac{\Sigma_{12}}{\sigma_1 \sigma_2} \le \overline{\rho} \right\},$$

where $\sigma_i > 0$, $i = 1, 2$, and $-1 \le \underline{\rho} \le \overline{\rho} \le 1$ are given constants. The set S models correlation uncertainty between the components of X. Consider a vector $w \in \mathbb{R}_+^2$ and a probability level $\alpha \in (1/2, 1)$. Compute the worst-case value-at-risk

$$\sup_{\Sigma \in S} \mathrm{VaR}_\alpha \left(- \sum_{i=1}^{2} w_i X_i \right)$$

of the portfolio loss $-\sum_{i=1}^{2} w_i X_i$. (4 Pts)

Solution 2

a) We say that a random vector $X = (X_1, \dots, X_d)$ is exchangeable if

$$(X_1, \dots, X_d) \overset{(d)}{=} (X_{\pi(1)}, \dots, X_{\pi(d)})$$

for any permutation π of $\{1, \dots, d\}$. Since every permutation of $\{1, \dots, d\}$ can be represented by a $d \times d$-matrix P with

$$P_{ij} = \mathbb{1}_{\{\pi(i)=j\}}$$

for all $i, j \in \{1, \dots, d\}$ and since we know from the lecture that t distribution is (as a normal variance mixture) closed under affine transformations, we have that $Y := PX \sim t_d(\nu, 0, P\Sigma P^\top)$. It is therefore enough to find a positive definite $d \times d$-matrix Σ, such that $P\Sigma P^\top \neq \Sigma$. Without loss of generality, take Σ diagonal with $\Sigma_{11} > \Sigma_{22}$ and P a permutation matrix corresponding to a permutation that swaps the first and second component of X (that is a matrix obtained by swapping the first and the second row of the $d \times d$ identity matrix). In that case it follows that $(P\Sigma P^\top)_{11} = \Sigma_{22}$ and $(P\Sigma P^\top)_{22} = \Sigma_{11}$, that is, $P\Sigma P^\top \neq \Sigma$.

Alternatively, assuming $\nu > 2$ so that $\mathbb{E}\left[X^2\right] < \infty$, we know from the class that

$$\mathrm{Cov}\left(X\right) = \frac{\nu}{\nu - 2} \Sigma.$$

Considering again a permutation that only swaps X_i and X_j, $i \neq j$, we get that $\mathrm{Var}\left(Y_i\right) = \mathrm{Var}\left(X_j\right) = \frac{\nu}{\nu-2} \Sigma_{jj}$ and $\mathrm{Var}\left(Y_j\right) = \mathrm{Var}\left(X_i\right) = \frac{\nu}{\nu-2} \Sigma_{ii}$ such that if $\Sigma_{ii} \neq \Sigma_{jj}$, the marginal distributions of X and Y differ, which means that we cannot have equality in distribution of X and Y.

b) Let $1_j \in \mathbb{R}^d$ for a $j \in \{1, \dots, d\}$ denote a vector whose components are all 0 except for the j-th component that is equal to 1. Then the j-th margin of X is given by $X_j = 1_j^\top X$. Let us show that the characteristic function of X_j is independent of j. Since we know that $X \sim S_d(\psi)$, we know from the lecture that the characteristic function $\phi_X \colon \mathbb{R}^d \to \mathbb{C}$ of X is given by

$$\phi_X(u) = \psi\left(\|u\|^2\right).$$

Analogously to before, let $u_j \in \mathbb{R}^d$ denote a vector whose components are all 0 except for the j-th component that is equal to $u \in \mathbb{R}$. We then have that

$$\phi_{X_j}(u) = \mathbb{E}\left[\exp\left(iu1_j^\top X\right)\right] = \mathbb{E}\left[\exp\left(iu_j^\top X\right)\right] = \phi_X(u_j) = \psi\left(\|u_j\|^2\right) = \psi(u^2),$$

which is what we wanted to show. Since the characteristic function of a random variable uniquely determines its distribution, we are done.

Alternatively, this can be proved by proving the contraposition, that is proving that if all univariate margins of X are not the same, then there exists no characteristic generator ψ such that $X \sim S_d(\psi)$. Without loss of generality, assume that

$$X_2 \overset{(d)}{=} X_3 \overset{(d)}{=} \dots \overset{(d)}{=} X_d \quad \text{and} \quad X_1 \overset{(d)}{\neq} X_2.$$

Then for any permutation π of $\{1, \dots, d\}$ such that $\pi(1) \neq 1$ when have that

$$(X_1, \dots, X_d) \overset{(d)}{\neq} (X_{\pi(1)}, \dots, X_{\pi(d)})$$

because if the marginal distributions are not the same, the joint distribution cannot be either. But since π can be represented by an orthogonal matrix P_π, we have found an orthogonal matrix P_π such that

$$P_\pi X \overset{(d)}{\neq} X,$$

which means that there exists no characteristic generator ψ such that $X \sim S_d(\psi)$.

The simplest way to do this, however, is directly. Let $X \sim S_d(\psi)$. Then we have $UX = X$ in distribution for any orthogonal $d \times d$-matrix U. Since every permutation matrix is orthogonal, considering the permutation matrix P which swaps the i-th and j-th element, we must clearly have $X_i = X_j$ is distribution since $PX = X$ in distribution. This can be repeated for every pair of the components of X.

c) Since we know that $X \sim N_2(\mu, \Sigma)$ for some $\Sigma \in S$, we know from the class that $-w^\top X \sim N(-w^\top \mu, w^\top \Sigma w)$. We thus have

$$\mathrm{VaR}_\alpha\left(-w^\top X\right) = -w^\top \mu + \sqrt{w^\top \Sigma w}\, \Phi^{-1}(\alpha) = -w^\top \mu + \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \Sigma_{12}}\, \Phi^{-1}(\alpha),$$

where Φ^{-1} is the quantile function of $N(0,1)$. Because of the above, we can write

$$\begin{aligned}
\sup_{\Sigma \in S} \mathrm{VaR}_\alpha\left(-w^\top X\right) &= \sup_{\underline{\rho} \le \rho \le \overline{\rho}} \left\{ -w^\top \mu + \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 \rho}\, \Phi^{-1}(\alpha) \right\} \\
&= -w^\top \mu + \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 \sup_{\underline{\rho} \le \rho \le \overline{\rho}} \{\rho\}}\, \Phi^{-1}(\alpha) \\
&= -w^\top \mu + \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 \overline{\rho}}\, \Phi^{-1}(\alpha) \\
&= -w^\top \mu + \sqrt{w^\top \overline{\Sigma} w}\, \Phi^{-1}(\alpha) = \mathrm{VaR}_\alpha\left(-w^\top Y\right),
\end{aligned}$$

where the second equality uses standard properties of the supremum (the function $x \mapsto \sqrt{x}$ is increasing and $\Phi^{-1}(\alpha) > 0$ since we assume $\alpha \in (1/2, 1)$), $\overline{\Sigma}$ is a 2×2-matrix with $\overline{\Sigma}_{ii} = \sigma_i^2$ for $i = 1, 2$ and $\overline{\Sigma}_{12} = \overline{\Sigma}_{21} = \sigma_1 \sigma_2 \overline{\rho}$ and $Y \sim N_2(\mu, \overline{\Sigma})$. That is, the worst-case value-at-risk is in this case attained when the correlation between X_1 and X_2 are as large as the uncertainty set S allows.

Question 3 (8 Pts)

a) Compute the lower tail dependence coefficient λ_l of the two-dimensional copulas $W(u,v) = (u + v - 1)^+$ and $M(u,v) = \min\{u, v\}$. (3 Pts)

b) Let (X_1, X_2) be a two-dimensional random vector with joint distribution given by the cdf

$$F(x_1, x_2) = \exp\left(-(-x_1 - x_2)^{1/\beta}\right) \quad \text{for } x_1, x_2 \leq 0 \quad \text{and} \quad \beta \geq 1.$$

Calculate the marginal distributions and the copula of (X_1, X_2). (5 Pts)

Solution 3

a) We know from the lecture that

$$\lambda_l = \lim_{\alpha \downarrow 0} \frac{C(\alpha, \alpha)}{\alpha},$$

so using the fact that $(2\alpha - 1)^+ = 0$ for all $\alpha \leq 0.5$ we obtain

$$\lambda_l(M) = \lim_{\alpha \downarrow 0} \frac{\min\{\alpha, \alpha\}}{\alpha} = \lim_{\alpha \downarrow 0} \frac{\alpha}{\alpha} = 1 \quad \text{and} \quad \lambda_l(W) = \lim_{\alpha \downarrow 0} \frac{(2\alpha - 1)^+}{\alpha} = 0.$$

b) We first compute the marginal cdfs F_{X_1} and F_{X_2} of X_1 and X_2, respectively. The marginal distributions are easily computed as

$$F_{X_1}(x_1) = F(x_1, \infty) = F(x_1, 0) = \exp\left(-(-x_1)^{1/\beta}\right),$$
$$F_{X_2}(x_2) = F(\infty, x_2) = F(0, x_2) = \exp\left(-(-x_2)^{1/\beta}\right).$$

We now want to use Sklar's theorem, which states that we can compute the copula of X as $C(u, v) = F\left(q_{X_1}(u), q_{X_2}(v)\right)$. We thus need to compute the quantile function of X_1 (the margins are identical). By inverting F_{X_1}, we get that

$$q_{X_1}(u) = q_{X_2}(u) = -(-\log(u))^{\beta}.$$

The copula is therefore given by

$$C(u, v) = \exp\left(-\left((-\log(u))^{\beta} + (-\log(v))^{\beta}\right)^{1/\beta}\right),$$

which is the Gumbel copula.

Question 4 (11 Pts)

Let X be a random variable with cdf

$$F(x) = 1 - x^{-\alpha}, \quad x \geq 1,$$

for a parameter $\alpha > 0$.

a) Does X have a density? If yes, derive it. (1 Pts)

b) Find all $k = 1, 2, \ldots$ such that $\mathbb{E}\left[|X|^k\right] < \infty$. (2 Pts)

c) Does F belong to MDA(H_ξ) for a generalized extreme value distribution H_ξ? If yes, what is ξ and what are the normalizing sequences? (3 Pts)

d) Calculate the excess distribution function $F_u(x) = \mathbb{P}\left[X - u \leq x \mid X > u\right]$, $x \geq 0$. (2 Pts)

e) Does there exist a parameter $\xi \in \mathbb{R}$ and a positive measurable function β such that

$$\lim_{u\to\infty} \sup_{x>0} \left|F_u(x) - G_{\xi,\beta(u)}(x)\right| = 0,$$

where $G_{\xi,\beta}$ is a generalized Pareto distribution? If yes, what are ξ and β? (3 Pts)

Solution 4

a) The density exists and can be computed as follows.

$$f_X(x) = \frac{d}{dx} F_X(x) = \alpha x^{-\alpha-1} \mathbb{1}_{\{x\geq 1\}}.$$

b) Using the density from the previous exercise, we directly compute

$$\mathbb{E}\left[|X|^k\right] = \mathbb{E}\left[X^k\right] = \int_1^\infty x^k f_X(x)dx = \int_1^\infty \alpha x^{k-\alpha-1} dx.$$

The above integral converges if and only if

$$k - \alpha - 1 < -1 \iff \quad k < \alpha,$$

so $\mathbb{E}\left[|X|^k\right] < \infty$ for all $k \in \{1, 2, \ldots\}$ with $k < \alpha$.

Alternatively, we could use the fact that F_X is in the MDA of Fréchet distribution with parameter $\xi = 1/\alpha$ (this will be shown in the solution to next exercise), which then gives by a result that we have seen in the class that $\mathbb{E}\left[|X|^k\right] < \infty \iff k < 1/\xi = \alpha$.

c) It can be seen from the previous exercise that the tail of the given distribution decays like a power and also that $\mathbb{E}\left[|X|^k\right] < \infty \iff k < \alpha$. We would thus expect $F_X \in \text{MDA}(H_\xi)$ for $\xi = 1/\alpha > 0$, that is the Fréchet distribution. This observation helps with constructing the normalizing sequences as

$$c_n = \frac{1}{\alpha}\left(\frac{1}{n}\right)^{-1/\alpha} \quad \text{and} \quad d_n = \left(\frac{1}{n}\right)^{-1/\alpha}.$$

We then have for all $x \in \mathbb{R}$ satisfying $c_n x + d_n > 1$ for all but finitely many $n \in \mathbb{N}$ that

$$F_X^n(c_n x + d_n) = \left(1 - (c_n x + d_n)^{-\alpha}\right)^n = \left(1 - \frac{(1 + \frac{1}{\alpha}x)^{-\alpha}}{n}\right)^n \to \exp\left(-\left(1 + \frac{1}{\alpha}x\right)^{-\alpha}\right)$$

as $n \to \infty$. For any other $x \in \mathbb{R}$ there exists an $N \in \mathbb{N}$ such that $c_n x + d_n \leq 1$ for all $n \geq N$, which means that $F_X^n(c_n x + d_n) = 0$ for all $n \geq N$ and therefore $\lim_{n\to\infty} F_X^n(c_n x + d_n) = 0$. With our choice of sequences this means that we need for all but finitely many $n \in \mathbb{N}$ that

$$\frac{1}{\alpha}\left(\frac{1}{n}\right)^{-1/\alpha} x + \left(\frac{1}{n}\right)^{-1/\alpha} > 1 \iff \frac{1}{\alpha}x + 1 > \left(\frac{1}{n}\right)^{1/\alpha} \iff x > -\alpha + \alpha\left(\frac{1}{n}\right)^{1/\alpha}.$$

This clearly holds for $x \in \mathbb{R}$ with $x > -\alpha$. So the limiting cdf is indeed that of GEV distribution with parameter $\xi = 1/\alpha$.

d) As seen in the lecture, we have that

$$F_u(x) = \mathbb{P}\left[X - u \leq x \,|\, X > u\right] = \frac{\mathbb{P}\left[u < X \leq x + u\right]}{\mathbb{P}\left[X > u\right]} = \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)},$$

which leads to

$$F_u(x) = \frac{1 - (x+u)^{-\alpha} - 1 + u^{-\alpha}}{u^{-\alpha}} = 1 - \frac{(x+u)^{-\alpha}}{u^{-\alpha}} = 1 - \left(1 + \frac{x}{u}\right)^{-\alpha}.$$

e) Pickands–Balkema–de Haan theorem gives us that

$$\lim_{u\to\infty} \sup_{x\geq 1} \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0 \tag{1}$$

if and only if $F_X \in \mathrm{MDA}(H_\xi)$. We have shown that $F_X \in \mathrm{MDA}(H_{1/\alpha})$, thus (1) holds for $\xi = 1/\alpha$ and for some measurable function $\beta\colon [1,\infty) \to (0,\infty)$ yet to be determined. We have

$$\lim_{u\to\infty} \sup_{x>0} \left| F_u(x) - G_{1/\alpha,\beta(u)}(x) \right| = \lim_{u\to\infty} \sup_{x>0} \left| \left(1 + \frac{x}{\alpha\beta(u)}\right)^{-\alpha} - \left(1 + \frac{x}{u}\right)^{-\alpha} \right|.$$

Picking $\beta(u) = \frac{u}{\alpha}$, we clearly obtain

$$\lim_{u\to\infty} \sup_{x>0} \left| F_u(x) - G_{1/\alpha,\beta(u)}(x) \right| = \lim_{u\to\infty} \sup_{x>0} \left| \left(1 + \frac{x}{u}\right)^{-\alpha} - \left(1 + \frac{x}{u}\right)^{-\alpha} \right| = 0.$$

So we can indeed take $\beta(u) = \frac{u}{\alpha}$.

Question 5 (10 Pts)

a) Assuming that you can simulate $U \sim \mathrm{Unif}(0,1)$ and $Z \sim N_d(0, I_d)$, where I_d denotes the d-dimensional identity matrix, describe an algorithm for simulating $X \sim M_d(\mu, \Sigma, \hat{F}_W)$. (3 Pts)

b) Name three methods for estimating a copula C from data. (3 Pts)

c) Explain what principal component analysis is. (4 Pts)

Solution 5

a) Let q_W denote a quantile function of the random variable W.

(1) Simulate $Z \sim N_d(0, I_d)$;

(2) Simulate W independent from Z by simulating $U \sim \mathrm{Unif}(0,1)$ and setting $W = q_W(U)$;

(3) Compute the Cholesky decomposition $\Sigma = AA^\top$;

(4) Return $X = \mu + \sqrt{W} AZ$.

b) Method of moments using rank correlation, maximum likelihood estimation, inference functions for margins estimator, maximum pseudo-likelihood estimator, non-parametric estimator by forming a pseudo-sample from the copula.

c) The goal of PCA is dimension reduction. The main idea is that if $X_1, \dots, X_n \in \mathbb{R}^d$ almost lie in a lower dimensional subspace A with $\dim(A) = p < d$, then the projections of $X_1, \dots, X_n$ onto A constitute a good approximation of the original observations $X_1, \dots, X_n$ and we can effectively work in a p-dimensional setting without losing much information.

PCA works as follows. Let $X \in \mathbb{R}^d$ be a random vector with $\mathbb{E}\left[X^2\right] < \infty$ and denote $\mu = \mathbb{E}\left[X\right]$ and $\Sigma = \mathrm{Cov}\left(X\right)$. Σ is symmetric and positive semidefinite; we can write $\Sigma = U\Lambda U^\top$, where $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_d)$ is the diagonal matrix of eigenvalues of Σ that are, without loss of generality, ordered so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d \geq 0$ and U is an orthogonal matrix whose columns are given by the eigenvectors of Σ in the order corresponding to the order of $\lambda_1, \ldots, \lambda_d$ in the matrix Λ. One then defines the principal component transform $Y = U^\top (X - \mu)$. We have

$$\mathbb{E}[Y] = U^\top \mathbb{E}[X] - U^\top \mu = 0$$

and

$$\mathrm{Cov}(Y) = U^\top \mathrm{Cov}(X - \mu) U = U^\top \Sigma U = U^\top U \Lambda U^\top U = \Lambda,$$

where the last equality follows from the orthogonality of U. The above means that the components $Y_1, \ldots, Y_d$ of Y are uncorrelated and $\mathrm{Var}(Y_i) = \lambda_i$. If we therefore have that λ_i is low for some $i \in \{1, \ldots, d\}$, it means that if "drop them", we can reduce the dimension of X without much loss of information.