Problems and suggested solution

Question 1

(a) A friend of yours tosses a fair coin and lets you choose between two bets A and B. In bet A you win 10 CHF if *head* shows up, and you loose 10 CHF if *tail* shows up. In bet B you win 30 CHF if *head* shows up, and you loose 10 CHF if *tail* shows up.

(i) [**2 Points**] Quantify the risks of the two bets in terms of Value-at-Risk at level 0.9, $\mathrm{VaR}_{0.9}$.

(ii) [**2 Points**] Quantify the risks of the two bets in terms of standard deviation, sd.

(iii) [**1 Point**] Rank the risks of the two bets in terms of a coherent risk measure ρ.

Solution:

Denote by L_A and L_B the losses suffered from the respective bets. Then $\mathbb{P}[L_A = -10] = \mathbb{P}[L_A = 10] = 1/2$ and $\mathbb{P}[L_B = -30] = \mathbb{P}[L_B = 10] = 1/2$.

(i)

$$\mathrm{VaR}_{0.9}(L_A) = 10,$$
$$\mathrm{VaR}_{0.9}(L_B) = 10.$$

(ii)

$$\mathrm{sd}(L_A) = 10,$$
$$\mathrm{sd}(L_B) = 20.$$

(iii) It holds that $L_B \le L_A$ (almost surely – note that the gain and the loss happen for the same events, respectively, for the two bets). The coherence of ρ implies that it is *monotone*. Hence, $\rho(L_B) \le \rho(L_A)$ (which means that bet B is at most as risky as bet A).

(b) Suppose you own a portfolio consisting of one share of stock A with current value $S_t^A = 700$ and 3 shares of stock B with current value $S_t^B = 100$ per share (both in CHF). The monthly log-returns of the stocks in % over the last 4 months are given in the following table:

Lag k	3	2	1	0
log-return of stock A at lag k	-2.0	1.0	-1.0	0.0
log-return of stock B at lag k	1.0	1.5	-2.0	-1.0

(i) [**1 Point**] Express the loss L_{t+1} of the portfolio over the next month as a function of the risk factor changes X_{t+1}^A and X_{t+1}^B given by

$$X_{t+1}^A = \log(S_{t+1}^A) - \log(S_t^A) \qquad \text{and} \qquad X_{t+1}^B = \log(S_{t+1}^B) - \log(S_t^B).$$

(ii) **[1 Point]** Express the *linearized* loss L^{Δ}_{t+1} of the portfolio as a function of X^A_{t+1} and X^B_{t+1}.

(iii) **[4 Points]** Use historical simulation to estimate $\mathrm{VaR}_{0.6}(L^{\Delta}_{t+1})$, $\mathrm{ES}_{0.6}(L^{\Delta}_{t+1})$ and $\mathrm{AVaR}_{0.6}(L^{\Delta}_{t+1})$.

Solution:

(i)

$$\begin{aligned}
L_{t+1} &= -(V_{t+1} - V_t) \\
&= -V_t\Big[0.7\Big(e^{X^A_{t+1}} - 1\Big) + 0.3\Big(e^{X^B_{t+1}} - 1\Big)\Big] \\
&= -1000\Big[0.7\Big(e^{X^A_{t+1}} - 1\Big) + 0.3\Big(e^{X^B_{t+1}} - 1\Big)\Big]
\end{aligned}$$

(ii)

$$L^{\Delta}_{t+1} = -1000\Big[0.7X^A_{t+1} + 0.3X^B_{t+1}\Big]$$

(iii) The historically simulated losses are

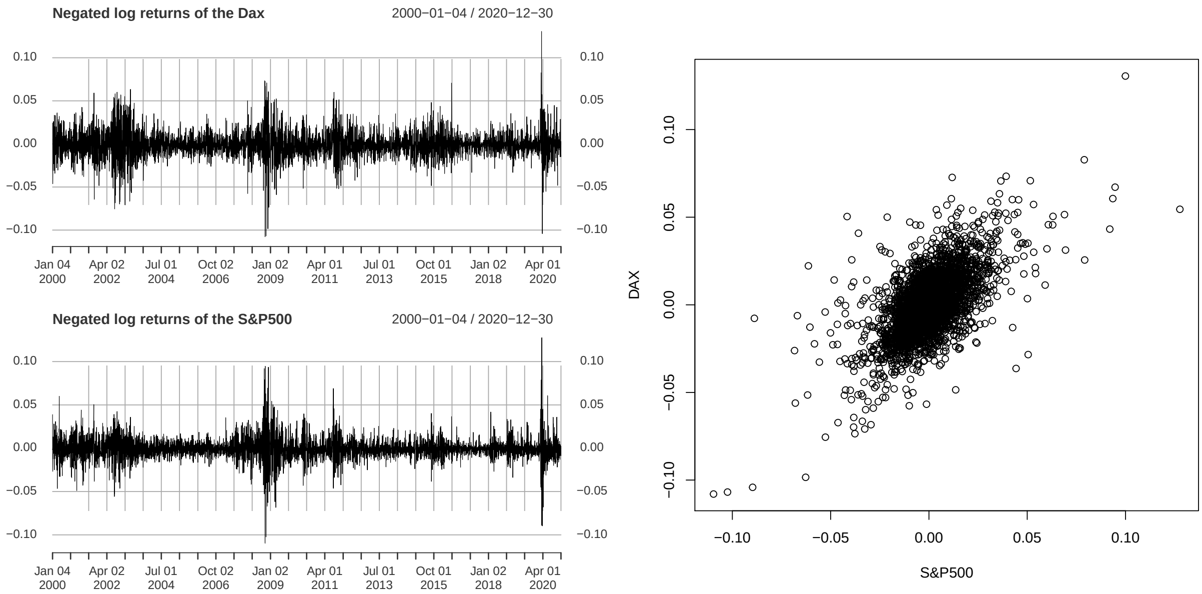
$$11, -11.5, 13, 3.$$

For $\mathrm{VaR}_{0.6}(L^{\Delta}_{t+1})$, we pick the $\lceil 0.6 \times 4 \rceil\,\mathrm{rd} = 3\,\mathrm{rd}$ smallest value, which is $\underline{11}$.
For $\mathrm{ES}_{0.6}(L^{\Delta}_{t+1})$, we need to take the average over all observations exceeding or being equal to the corresponding VaR, which is $(11 + 13)/2 = \underline{12}$.
For $\mathrm{AVaR}_{0.6}(L^{\Delta}_{t+1})$, we obtain

$$\begin{aligned}
\frac{1}{1-0.6}\int_{0.6}^{1} \mathrm{VaR}_u(L^{\Delta}_{t+1})\,\mathrm{d}u &= \frac{1}{0.4}[0.15 \times 11 + 0.25 \times 13] \\
&= \frac{3}{8} \times 11 + \frac{5}{8} \times 13 = \underline{12.25}.
\end{aligned}$$

Question 2

(a) [**4 Points**] The following pictures show 5195 daily negative log-returns of the S&P500 and Dax from the start of 2000 to the end of 2020 (left column) along with a scatter plot of these negative log-returns (right column).



Describe four stylized facts of univariate/multivariate daily financial log-return series and relate them to the pictures.

Solution:

One point for each correct stylized fact along with a reasonable description of the figure.

The stylized facts of univariate time series are:

(U1) Return series are not iid although they show little serial correlation:
In the time series plots on the left, we see volatility clusters, which defies the independence assumption. On the other hand, the past returns do not provide any information about the sign of the next return, which implies the little serial correlation.

(U2) Series of absolute or squared returns show profound serial correlation:
We can see this in the time series plots on the left and it is implied by the visible volatility clusters. In case of a high absolute return at day t (e.g., around Jan 2009) the return on the next day $t+1$ tends to be higher as well. Similarly, times of low volatility (e.g., between 2004 and 2006) seem also to be persistent.

(U3) Conditional expected returns are close to zero:
We can see this in the time series plots on the left. No matter what information about the past is given, the next return fluctuates around 0.

(U4) Volatility (conditional standard deviation) appears to vary over time:
See explanation for (U2).

(U5) Extreme returns appear in clusters:
See explanation for (U2). Also observe that (U5) implies (U4).

(U6) Return series are leptokurtic or heavy-tailed (power-like tail):
This is harder to judge from the provided plots. It would be easier to have histograms or – even better – QQ-Plots.

The stylized facts of multivariate time series are:

(M1) Multivariate return series show little evidence of cross-correlation, except for contemporaneous returns (i.e. at the same time t):
We can see evidence for (positive) contemporaneous cross-correlation in the scatter plot on the right.

(M2) Multivariate series of absolute returns show profound cross-correlation:
We see this by jointly considering the two time series plots. A high absolute return on day t in one time series tends to be associated with an absolute high return in the other time series the next day.

(M3) Correlations between contemporaneous returns vary over time:
This is not so easily visible. A plot of the empirical cross-correlation function would help.

(M4) Extreme returns in one series often coincide with extreme returns in several other series (i.e. tail dependence):
This is nicely visible both in the joint time series plot and also in the scatter plot. E.g., we have extreme returns in both time series in late 2008 / early 2009 (start of the financial crisis) and a pronounced volatility spike at the start of the first Covid wave (March / April 2020). In the scatter plot, we can see points in the upper right and lower left corner, which also underline this stylized fact.

(b) [**2 Points**] Mention a stylized fact of univariate financial log-return time series GARCH(1,1)-processes can replicate well, and explain briefly how they do so.

Solution:

GARCH(1,1)-processes a tailored to exhibit a time varying volatility, which produces volatility clusters.

They do so since their squared volatility at time t, σ_t^2, is an affine function of the squared volatility at time $t-1$, σ_{t-1}^2, and of the squared return at time $t-1$.

(Alternatively, it is also okay if the answers refer to (U1), (U2), or (U3) with a correct explanation.)

(c) [**4 Points**] Discuss if the following statement is true or false: "If the random variables X_1 and X_2 both follow a standard normal distribution with known correlation ρ, then it is possible to calculate $\mathrm{VaR}_\alpha(v_1X_1 + v_2X_2)$ for any $\alpha \in (0,1)$ and for any $v_1, v_2 \in \mathbb{R}$."

Solution:

This is false.

Indeed, knowing $\mathrm{VaR}_\alpha(v_1X_1 + v_2X_2)$ for each $\alpha \in (0,1)$ is equivalent to knowing the distribution of $v_1X_1 + v_2X_2$.

However, knowing the distribution of $v_1X_1 + v_2X_2$ for any $v_1, v_2 \in \mathbb{R}$ is equivalent to knowing the distribution of the random vector (X_1, X_2), using characteristic functions.

On the other hand, only knowing the marginal distributions of X_1 and X_2 along with their correlation does not uniquely determine their joint distribution. E.g., they could be jointly normal. But on the other hand, their copula could also be a mixture of the co- and counter-monotonicity copula.

Remark: The previous point is true if and only if $\rho \in (-1, 1)$. It was not expected to state that explicitly.

As an alternative solution, it would also acceptable to provide an argument for a specific choice of $v_1, v_2 \in \mathbb{R}$ (and also $\alpha \in (0, 1)$) where where $\mathrm{VaR}_\alpha(v_1X_1 + v_2X_2)$ is not uniquely determined by the two standard normal marginal distributions and the correlation.

Question 3

Let X be a random random variable with cumulative distribution function

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

(a) **[1 Point]** Does X have a density? If no, explain why it cannot have a density. If yes, derive the density.

Solution:

Yes, the density is

$$f(x) = F'(x) = \frac{e^x}{(e^x+1)^2} = \frac{1}{e^x + 2 + e^{-x}}.$$

(b) **[2 Points]** Find all $k \in \mathbb{N} = \{1, 2, \dots\}$ such that $\mathbb{E}[|X|^k] < \infty$, providing an explanation.

Solution:

We claim that $\mathbb{E}[|X|^k] < \infty$ for all $k \in \mathbb{N}$.

Indeed, f is an even function. So it is sufficient to check that $\int_0^\infty f(x)x^k dx < \infty$ for all $k \in \mathbb{N}$. Since $f(x) \le e^{-x}$ for $x \ge 0$, it suffices to check that $\int_0^\infty e^{-x}x^k dx < \infty$ for all $k \in \mathbb{N}$. This follows inductively: Indeed $\int_0^\infty e^{-x}dx = 1$ and $\int e^{-x}x^k dx = k \int e^{-x}x^{k-1}dx$ for $k \ge 1$.

An alternative way to argue that all moments are finite is to use the solution of problem c). If a distribution is in the maximum domain of attraction of H_0, then all moments exist.

(c) **[4 Points]** Does F belong to the maximum domain of attraction $\mathrm{MDA}(H_\xi)$ for a standard GEV distribution H_ξ? If yes, what is ξ and what are the normalizing sequences?

Hint: You may use that for any sequence $(w_n)_{n\in\mathbb{N}}$ converging to $w \in \mathbb{R}$ it holds that

$$\lim_{n\to\infty} \left(1 + \frac{w_n}{n}\right)^n = \exp(w).$$

Solution:

We have seen that all moments exist. Moreover, the tail decays exponentially. Indeed,

$$\bar{F}(x) = e^{-x}\frac{e^x}{e^x + 1},$$

where $\frac{e^x}{e^x+1}$ is a slowly varying function.

Alternatively, we can also argue with the fact that all moments exist (see b)), but that $x_F = \infty$.

Hence, the only possible standard GEV distribution is the Gumbel case, $\xi = 0$, where

$$H_0(x) = \exp(-e^{-x}),$$

though we cannot directly apply a characterization result here.

Hence, we need to prove the convergence explicitly by showing that $F^n(c_n x + d_n) \to H_0(x)$ as $n \to \infty$ for any $x \in \mathbb{R}$ and for some normalizing sequences (c_n) and (d_n).

We use $c_n = 1$ and $d_n = \log(n)$. (Note that these sequences are not unique, other choices can also do the job.)

Then we obtain

$$F^n(c_n x + d_n) = \left(1 - \frac{e^{-x}}{n + e^{-x}}\right)^n = \left(1 + \frac{w_n}{n}\right)^n,$$

where

$$w_n := -\frac{e^{-x}}{1 + e^{-x}/n} \overset{n\to\infty}{\to} -e^{-x}.$$

This shows the claim, exploiting the hint.

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

Solution:

It holds for $x \geq 0$ that

$$F_u(x) = \frac{F(x+u) - F(u)}{1 - F(u)}.$$

This can be simplified to

$$\begin{aligned} F_u(x) &= (e^u + 1)\left(\frac{e^{x+u}}{e^{x+u} + 1} - \frac{e^u}{e^u + 1}\right) \\ &= e^u + 1 - \frac{e^u + 1}{e^{x+u} + 1} - e^u \\ &= 1 - \frac{e^u + 1}{e^{x+u} + 1}. \end{aligned}$$

(e) **[4 Points]** Does there exist a parameter $\xi \in \mathbb{R}$ and a function $\beta\colon \mathbb{R} \to (0, \infty)$ such that

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

where $G_{\xi,\beta}$ denotes the cumulative distribution function of a GPD? If yes, for which ξ and β does this hold?

Solution:

Since $x_F = \infty$, Pickands–Balkema–de Haan Theorem implies that there exists a measurable function $\beta\colon \mathbb{R} \to (0, \infty)$ such that

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

if and only if $F \in \mathrm{MDA}(H_\xi)$.

We have shown in (c) that $F \in \mathrm{MDA}(H_0)$, thus (1) holds for $\xi = 0$ and for some function $\beta(u)$, which needs to be shown that it is constant 1.

Indeed, for the pointwise limit, we have that for all $x \in \mathbb{R}$

$$\lim_{u \to \infty} F_u(x) = G_{0,1}(x) = 1 - e^{-x}$$

To show the uniform convergence, observe that

$$\sup_{x>0} |F_u(x) - G_{0,1}(x)| = \sup_{x>0} \left| \frac{e^{-x} - 1}{e^{x+u} + 1} \right| \leq \sup_{x>0} \frac{1}{e^{x+u} + 1} = \frac{1}{e^u + 1} \overset{u \to \infty}{\longrightarrow} 0.$$

Question 4

Let (X_1, X_2) be a two-dimensional random vector with cumulative distribution function

$$F(x_1, x_2) = \begin{cases} \exp\left(-\left(x_1^{-\theta} + x_2^{-\theta}\right)^{1/\theta}\right), & \text{if } x_1 > 0 \text{ and } x_2 > 0, \\ 0, & \text{else} \end{cases}$$

for a parameter $\theta \in [1, \infty)$.

(a) **[2 Points]** Derive the two marginal cumulative distribution functions F_1 and F_2.

Solution:

The marginal distributions F_1 and F_2 coincide due to symmetry. It holds that

$$\begin{aligned} F_1(x_1) &= \lim_{x_2 \to \infty} F(x_1, x_2) \\ &= \begin{cases} \exp\left(-\left(x_1^{-\theta} + \lim_{x_2 \to \infty} x_2^{-\theta}\right)^{1/\theta}\right), & x_1 > 0 \\ 0, & \text{else} \end{cases} \\ &= \begin{cases} \exp\left(-1/x_1\right), & x_1 > 0, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

(b) **[2 Points]** Derive the copula of F.

Solution:

The quantile function of F_j, $j = 1, 2$, is given by $q_j(u_j) = 1/(-\log u_j)$ for $u_j \in [0, 1]$ (meaning that $q_j(0) = 0$ and $F_j^{\leftarrow}(1) = \infty$).

Using Sklar's Theorem, we obtain for $(u_1, u_2)' \in [0, 1]^2$

$$C(u_1, u_2) = F\Big(q_1(u_1), q_2(u_2)\Big) = \exp\left(-\left((-\log u_1)^{\theta} + (-\log u_2)^{\theta}\right)^{1/\theta}\right).$$

(c) (i) **[2 Points]** Assume $\theta = 2$. Compute the probability that X_1 and X_2 both exceed their $\mathrm{VaR}_{0.95}$.

Hint: You may use that $\exp\left(-\left((-\log 0.95)^2 + (-\log 0.95)^2\right)^{1/2}\right) \approx 0.93$.

(ii) **[2 Points]** Show that this probability is approximately 12 times larger than the probability of the same event if X_1 and X_2 were independent.

Solution:

(i) If $(X_1, X_2) \sim F$, then

$$\begin{aligned}\mathbb{P}[X_1 > q_1(0.95), X_2 > q_2(0.95)] &= \mathbb{P}[F_1(X_1) > 0.95, F_2(X_2) > 0.95] \\ &= C(1,1) - C(0.95,1) - C(1,0.95) + C(0.95,0.95) \\ &= 1 - 0.95 - 0.95 + 0.93 \\ &= 0.03.\end{aligned}$$

(ii) If $X_1 \sim F_1$, $X_2 \sim F_2$ and X_1 and X_2 are independent, then

$$\begin{aligned}\mathbb{P}[X_1 > q_1(0.95), X_2 > q_2(0.95)] &= \mathbb{P}[F_1(X_1) > 0.95, F_2(X_2) > 0.95] \\ &= \Pi(1,1) - \Pi(0.95,1) - \Pi(1,0.95) + \Pi(0.95,0.95) \\ &= 1 - 0.95 - 0.95 + 0.95^2 \\ &= (1 - 0.95)^2 \\ &= 0.0025.\end{aligned}$$

Of course, an alternative way for the solution is to exploit the independence directly. That is

$$\begin{aligned}\mathbb{P}[X_1 > q_1(0.95), X_2 > q_2(0.95)] &= \mathbb{P}[F_1(X_1) > 0.95, F_2(X_2) > 0.95] \\ &= \mathbb{P}[F_1(X_1) > 0.95]\mathbb{P}[F_2(X_2) > 0.95] \\ &= (1 - 0.95)^2 \\ &= 0.0025.\end{aligned}$$

Finally, it holds that

$$0.03/0.0025 = \frac{3}{25} \times 10^2 = \frac{12}{100} \times 10^2 = 12.$$

Question 5

(a) **[2 Points]** Describe the notions of risk and uncertainty, clearly pointing out the difference between them.

(b) **[2 Points]** Where would you place financial markets in the spectrum between risk and uncertainty? Briefly justify your answer.

(c) **[4 Points]** Let $L^2(\mathbb{P})$ be the space of all square-integrable random variables on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider the standard deviation mapping $\mathrm{sd}\colon L^2(\mathbb{P}) \to \mathbb{R}$ given by

$$\mathrm{sd}(X) = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}, \quad X \in L^2(\mathbb{P}).$$

Which properties of a coherent risk measure does sd have? Please, justify your answers.

Solution:

(a) The term "risk" describes the situation of a random quantity with known probability distribution.

The term "uncertainty" refers to the situation of a random quantity with unknown probability distribution.

(b) Financial markets are somewhere between the two notions. Using historical data, it is possible to get a reasonable estimate about the probability distribution of a random loss. However, the assumption of stationarity is definitely violated since there occur regular change points in the distribution. E.g., in the situation of a major and unforeseen crises such as the covid crisis or the current war in Europe.

(c)
- The monotonicity does not hold in general. To this end, consider, e.g., L_1 with distribution $\mathbb{P}(L_1 = -11) = \mathbb{P}(L_1 = -9) = 1/2$ and L_2 which is almost surely 0. Clearly, $L_1 \leq L_2$ almost surely. However, $\mathrm{sd}(L_1) = 1 > \mathrm{sd}(L_2) = 0$.
- The translation property is violated since $\mathrm{sd}(L + c) = \mathrm{sd}(L)$ for any $c \in \mathbb{R}$.
- The positive homogeneity, $\mathrm{sd}(\lambda L) = \lambda \mathrm{sd}(L)$ for all $\lambda > 0$, holds due to the linearity of the expectation.
- The subadditivity holds. Indeed, we can check this for variance and recall that the square root function is monotone,

$$\begin{aligned}\mathrm{sd}(L_1 + L_2)^2 &= \mathrm{Var}(L_1 + L_2) = \mathrm{Var}(L_1) + 2\mathrm{Cov}(L_1, L_2) + \mathrm{Var}(L_2) \\ &\leq \mathrm{Var}(L_1) + 2\mathrm{sd}(L_1)\mathrm{sd}(L_2) + \mathrm{Var}(L_2) = \Big(\mathrm{sd}(L_1) + \mathrm{sd}(L_2)\Big)^2.\end{aligned}$$

(The inequality follows from the Cauchy–Schwarz inequality – or from the fact that the correlation is bounded by 1.)