

Problems and suggested solution

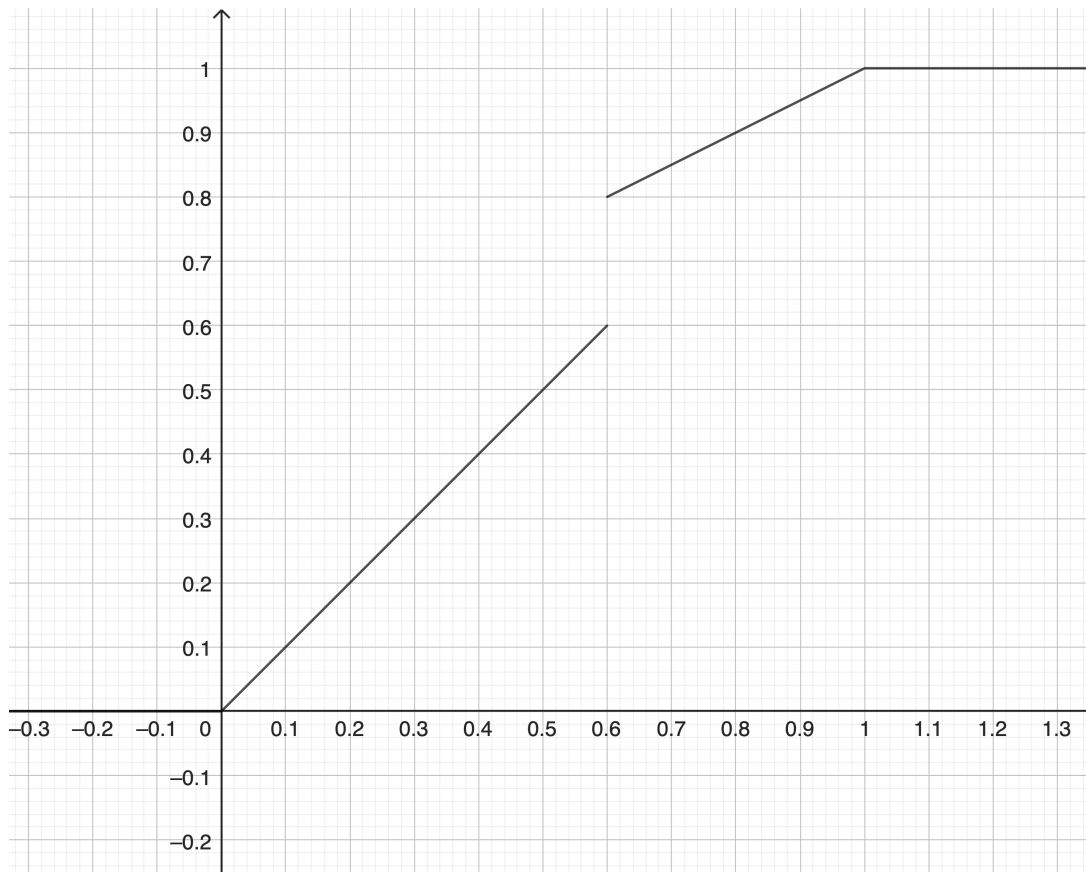
Question 1

- (a) [3 Points] Consider a random variable X with cdf F given by

$$F(x) = \begin{cases} 0 & ; x < 0 \\ x & ; 0 \leq x < 0.6 \\ 0.5x + 0.5 & ; 0.6 \leq x < 1 \\ 1, & ; x \geq 1. \end{cases}$$

The corresponding graph is shown below.

Compute $\text{VaR}_\alpha(X)$, $\text{AVaR}_\alpha(X)$ and $\text{ES}_\alpha(X)$ at level $\alpha = 0.8$.



Solution:

$$\text{VaR}_{0.8}(X) = 0.6$$

$$\text{AVaR}_{0.8}(X) = \frac{1}{1-0.8}(0.2 \cdot 0.6 + 0.5 \cdot 0.2 \cdot 0.4) = 0.8$$

$$\text{ES}_{0.8}(X) = \mathbb{E}[X \mid X \geq 0.6] = \frac{1}{2}(0.6 + \text{AVaR}_{0.8}(X)) = 0.7$$

- (b) [3 Points] What does it mean for the cdf of a non-constant random variable Y that $\text{VaR}_{0.9}(Y) = \text{AVaR}_{0.9}(Y)$? Explain by explicitly referring to the definition of $\text{AVaR}_\alpha(Y)$ and by providing a sketch of the cdf of such a Y .

Solution:

$\text{AVaR}_\alpha(Y) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_u(Y) du = \text{VaR}_\alpha(Y)$, or an equivalent explanation in words

The equality $\text{VaR}_{0.9}(Y) = \text{AVaR}_{0.9}(Y)$ holds if and only if $\text{VaR}_u(Y)$ is constant for all $u \in (\alpha, 1)$.

If the argument is partially correct, or lacks precision and rigor, only 0.5P.

The sketch of the cumulative distribution function must not be degenerate and at the 0.9-quantile it must jump to 1.

(c) [**3 Points**] Let ρ_1 and ρ_2 be two coherent risk measures. Decide without explanation if the following combinations are again coherent or not.

- (i) $\rho_1 - \rho_2$; (ii) $\rho_1 + \rho_2$; (iii) $\max(\rho_1, \rho_2)$

Solution:

(i) $\rho_1 - \rho_2$ not coherent

(ii) $\rho_1 + \rho_2$ not coherent

(iii) $\max(\rho_1, \rho_2)$ coherent

Question 2

- (a) [1 Point] What is the definition of strong stationarity of a time series $(X_t)_{t \in \mathbb{Z}}$?

Solution:

For each $s, t, h \in \mathbb{Z}$, $s < t$, the random vector $(X_s, X_{s+1}, \dots, X_t)$ has the same distribution as $(X_{s+h}, X_{s+1+h}, \dots, X_{t+h})$.

- (b) In the figure below, you may see three time series. One shows daily values of the DAX, S_t , between the start of 2000 and the end of 2020, one shows its daily absolute returns, $S_t - S_{t-1}$, and another one its daily log-returns, $\log(S_t) - \log(S_{t-1})$.

- (i) [1 Point] Which plot belongs to which time series? (No need to justify.)

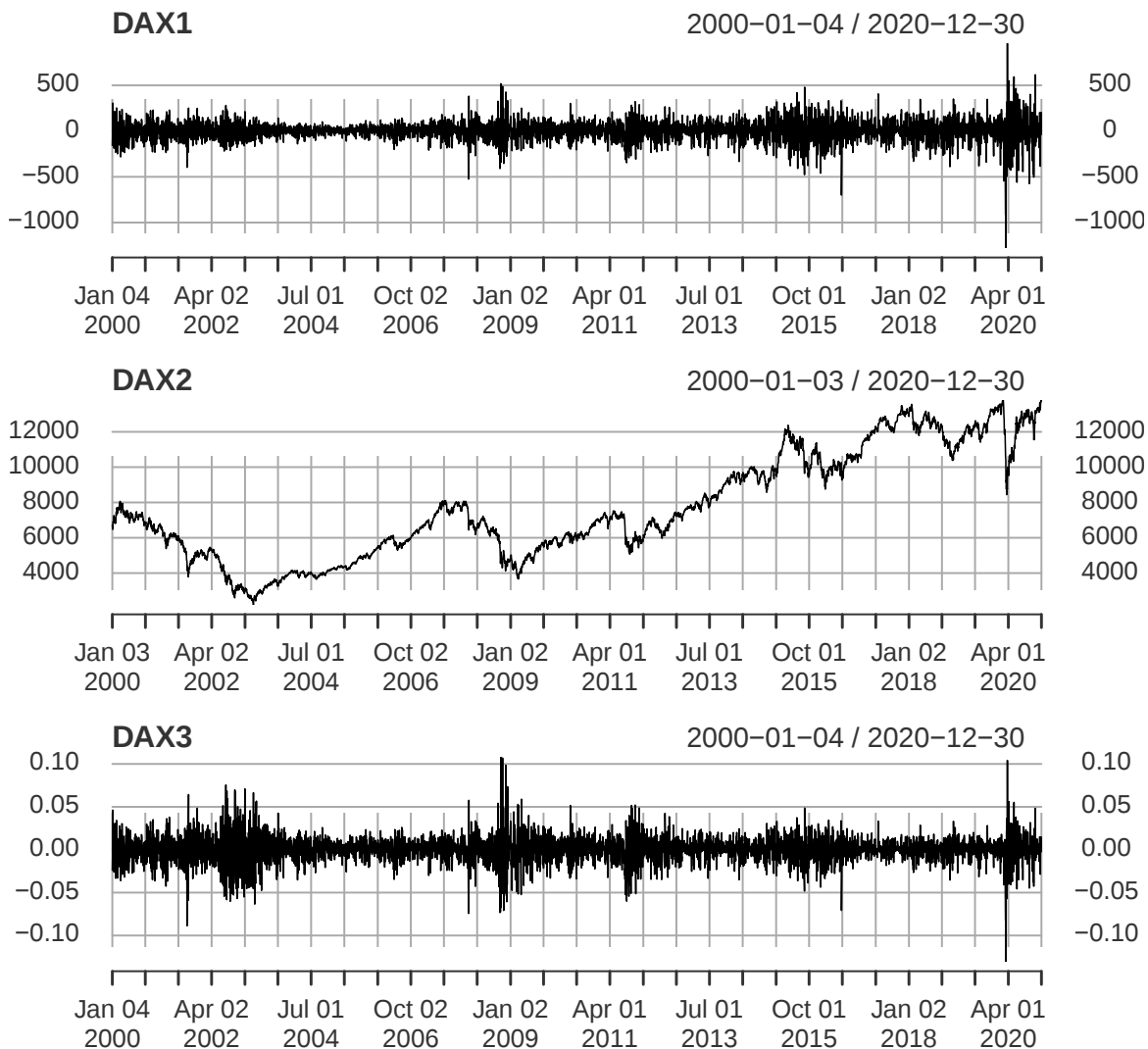
Solution:

From top to down: Absolute returns, DAX itself, log-returns.

- (ii) [3 Points] Briefly discuss for each of the three time series if it seems to be stationary or not.

Solution:

DAX1 / Absolute returns: Not stationary. Variance increases with time.
DAX2 / raw data: Not stationary. Increasing trend (increasing mean function).
DAX3 / log-returns: Stationary. Mean close to zero and constant. Variance also stable in the long run (though volatility clusters visible, but they are in line with stationarity).



- (c) [1 Point] Briefly discuss the relevance of stationarity in the statistical analysis of time series.

Solution:

Stationarity is a crucial precondition for the past to be representative for the presence and future. Hence, the assumption of stationarity lies at the heart of statistical analysis of time series.

- (d) [1 Point] Provide the definition of a GARCH(1,1) process.

Solution:

A GARCH(1,1) process is a time series $(X_t)_{t \in \mathbb{Z}}$ with the specification

$$X_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where $(Z_t)_{t \in \mathbb{Z}}$ is a strict white noise process with variance 1 and where $\alpha_0 > 0$, $\alpha_1, \beta \geq 0$ and $\alpha_1 + \beta_1 < 1$.

- (e) [3 Points] Mention three stylized facts of financial time series consisting of daily log-returns, which a GARCH(1,1) process can replicate well, and *briefly* explain how / why.

Solution:

(U3) Conditional expected returns are close to 0. Indeed. If $\mathcal{F}_{t-1}$ is generated by $X_{t-1}, X_{t-2}, \dots$, then

$$\mathbb{E}[X_t \mid \mathcal{F}_{t-1}] = \sigma_t \mathbb{E}[Z_t] = 0.$$

(U1) Return series are not iid although they show little serial correlation. Indeed, a GARCH(1,1) process has vanishing autocorrelation. Indeed, for $s < t$ it holds that

$$\mathbb{E}[X_s X_t] = \mathbb{E}[\mathbb{E}[X_s X_t \mid \mathcal{F}_{t-1}]] = \mathbb{E}\left[X_s \underbrace{\mathbb{E}[X_t \mid \mathcal{F}_{t-1}]}_{=0}\right] = 0.$$

(U4) Volatility (conditional standard deviation) appears to vary over time. Indeed, we can show that σ_t is the volatility of the GARCH process and this quantity varies over time.

Instead of (U4), one can also mention and elaborate on (U2) (Series of absolute or squared returns show profound serial correlation.) or (U5) (Extreme returns appear in clusters – due to the fact that σ_t^2 is an increasing function of σ_{t-1}^2).

Question 3

Let X be a random random variable with cdf

$$F_X(x) = \begin{cases} 0, & x < 2 \\ \frac{x^2-1}{x^2}, & x \geq 2. \end{cases}$$

- (a) [1 Point] Does X have a density? If yes, derive it. If no, provide an argument why not.

Solution:

No, since F_X has a jump at $x = 2$. Indeed, $F_X(x) = 0$ for $x < 2$ and $F_X(2) = \frac{3}{4}$.

- (b) [2 Points] Show that F_X belongs to an $\text{MDA}(H_\xi)$ for a standard GEV distribution H_ξ and determine the corresponding ξ .

Solution:

We may use a characterisation result from the lecture for $\xi > 0$: $F \in \text{MDA}(H_\xi)$ if and only if $1 - F(x) = x^{-1/\xi} L(x)$ where L is a slowly varying function.

It holds that

$$1 - F(x) = x^{-2}.$$

Since any constant is a slowly varying function, it holds that $F \in \text{MDA}(H_\xi)$ for $\xi = \frac{1}{2}$.

Alternatively, it is also possible to show that $F \in \text{MDA}(H_{1/2})$ by correctly applying the definition of the MDA with the normalising sequences $c_n = \sqrt{n}/2$ and $d_n = \sqrt{n}$. Indeed, for any $x > 0$ there is an $N \in \mathbb{N}$ such that for all $n > N$ it holds that

$$F^n(c_n x + d_n) = \left(1 - \frac{1}{(c_n x + d_n)^2}\right)^n = \left(1 - \frac{(1 + \frac{1}{2}x)^{-2}}{n}\right)^n \rightarrow \exp(-(1 + \frac{1}{2}x)^{-2}) = H_{1/2}(x).$$

- (c) [1 Point] Calculate the excess distribution function $F_u(x) = \mathbb{P}[X - u \leq x \mid X > u]$, where $x \geq 0, u \geq 2$.

Solution:

We have for $x \geq 0$ and $u \geq 2$,

$$F_u(x) = \frac{F_X(x+u) - F_X(u)}{1 - F_X(u)} = \frac{-(x+u)^{-2} + u^{-2}}{u^{-2}} = 1 - \frac{u^2}{(x+u)^2}.$$

- (d) [2 Points] Does there exist a parameter $\xi \in \mathbb{R}$ and a function β such that

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

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

Solution:

Pickands–Balkema–de Haan Theorem implies that there exists a measurable function $\beta: \mathbb{R} \rightarrow (0, \infty)$ such that

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

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

We have shown in (b) that $F \in \text{MDA}(H_{1/2})$, thus (1) holds for $\xi = 1/2$ and for some function $\beta(u)$.

We have that for $x \geq 0$

$$G_{1/2, \beta}(x) = 1 - \frac{1}{(1 + \frac{x}{2\beta})^2}$$

Furthermore, it can be seen from (c) that (for $u \geq 2$)

$$F_u(x) = G_{1/2, \beta(u)}(x) \quad \text{for } \beta(u) = u/2.$$

So, the convergence holds for $\xi = 1/2$ and $\beta(u) = u/2$.

- (e) [**1 Point**] Let $X_1, X_2, \dots$ be iid random variables with cdf F_X given above. Let $M_n = \max\{X_1, \dots, X_n\}$ be the maximum. Where does the maximum converge to as $n \rightarrow \infty$?

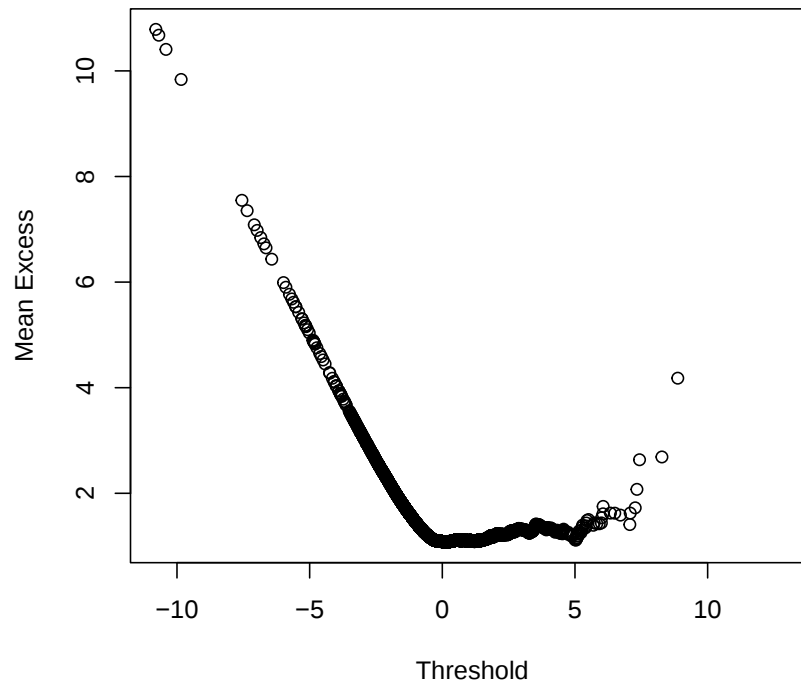
Solution:

Since $x_F = \infty$, the maximum converges to ∞ almost surely.

E: 0.5 if they write $F_{M_n}(x) = \mathbb{P}(M_n \leq x) = (F_X(x))^n \xrightarrow[n \rightarrow \infty]{} 0 \quad \forall x$

- (f) [**2 Points**] Below, you may see a sample mean excess function $e_n(v)$ from an iid sample $Y_1, Y_2, \dots$ with cdf G . Suppose that $G \in \text{MDA}(H_{\xi'})$.

Describe how you can use the sample mean excess function $e_n(v)$ to determine the sign of ξ' in general. Then determine the sign of ξ' for this situation.



Solution:

We know theoretically that beyond a certain threshold u , the sample mean excess function, $e_n(v)$, $v \geq u$, behaves like a linear function.

The sign of the slope indicates the sign of ξ' , where $G \in \text{MDA}(H_{\xi'})$.

We can see that beyond the threshold of about $u = 1$, the slope of $e_n(v)$ is positive. This provides evidence that $\xi' > 0$.

Question 4

(a) Suppose you have iid observations $(X_1, Y_1), \dots, (X_n, Y_n)$ of a two-dimensional random vector (X, Y) .

(i) [1 Point] Provide a formula for the sample version of Kendall's tau, $r_\tau(n)$.

Solution:

$$\begin{aligned} r_\tau(n) &= \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \text{sign}((X_i - X_j)(Y_i - Y_j)) \\ &= \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \text{sign}((X_i - X_j)(Y_i - Y_j)) \end{aligned}$$

(ii) [1 Point] Suppose that $n = 5$ and your random sample is

$$\{(-2, -3), (-1, 1), (0, -1), (1, 5), (3, 3)\}.$$

Compute $r_\tau(n)$ on this sample.

Solution:

$$\begin{aligned} r_\tau(n) &= \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \text{sign}((X_i - X_j)(Y_i - Y_j)) \\ &= \frac{1}{10}(4 + 1 + 2 - 1) = 0.6 \end{aligned}$$

(iii) [1 Point] Describe how your result from (ii) can help you to determine $r_\tau(n)$ on the sample

$$\{(-2, (-3)^5), (-1, (1)^5), (0, (-1)^5), (1, (5)^5), (3, (3)^5)\}.$$

Hint: You can solve this exercise even if you don't have a specific result for (ii).

Solution:

Kendall's tau is invariant with respect to strictly increasing transformations of one of its components. Since $\mathbb{R} \ni x \mapsto x^5$ is strictly increasing, $r_\tau(n)$ remains the same as above.

(b) Let (X, Y) be a random vector with joint cdf F and continuous marginal cdfs F_X and F_Y , which are unknown.

Suppose you know that X has mean 0 and standard deviation 1, while Y has mean 5 and standard deviation 2.

(i) [2 Points] Suppose you know that X and Y are uncorrelated and you know F_X . Is it possible to determine F_Y ? Justify your answer either by a counterexample or specifically calculate F_Y .

Solution:

No, it is not possible.

F_Y could be any distribution with mean 5 and standard deviation 2.

- (ii) [2 Points] Suppose you know that X and Y are uncorrelated and you know both F_X and F_Y . Is it possible to determine F ? Justify your answer either by a counterexample or specifically calculate F .

Solution:

No, it is not possible.

There are multiple possible copulas which lead to a correlation of 0, not only the independence copula.

- (iii) [2 Points] Suppose you know that X and Y have a correlation of 1 and you know F_X . Is it possible to determine F_Y ? Justify your answer either by a counterexample or specifically calculate F_Y .

Solution:

Yes, it is possible.

We know that if X and Y have correlation 1, then $Y = aX + b$ almost surely for some $a > 0$. Since we know the means and standard deviations of X and Y , we know that $Y = 2X + 5$ almost surely. This determines F_Y as $F_Y(t) = F_X((t - 5)/2)$.

- (iv) [2 Points] Suppose you that X and Y have a correlation of 1 and you know both F_X and F_Y . Is it possible to determine F ? Justify your answer either by a counterexample or specifically calculate F .

Solution:

Yes, it is possible.

A correlation of 1 implies that X and Y are comonotonic. Since the marginal distributions are continuous, the copula of the joint distribution must be unique and therefore must be the comonotonicity copula.

That means

$$F(x_1, x_2) = M(F_1(x_1), F_2(x_2)) = \min(F_1(x_1), F_2(x_2))$$

Question 5

- (a) You backtest two different models for $\text{VaR}_{0.95}$ -forecasts of losses. Model M1 uses a static approach based on historical simulation, and M2 uses a GARCH(1,1) specification where the distribution of the innovations has also been fitted with historical simulation.

The evaluation window consists of 1000 daily losses. You compute the number of exceedances (how often the time series exceeds the specific forecasts). Moreover, you compute the average score $\bar{S}$ of the strictly consistent score $S(x, y) = (1_{\{y \leq x\}} - 0.95)(x - y)$. You may find the results in the following table.

	M1	M2
Exceedances	48	112
$\bar{S}$	0.15	0.06

- (i) [2 Points] What do these results indicate in terms of calibration and accuracy of the two models?

Solution:

For $\text{VaR}_{0.95}$ -forecasts, we expect an exceedance with a probability of 5%. That means, on the evaluation window of 1000 observations, we expect about 50 exceedances if the forecasts are calibrated.

The exceedances of M1 are close to this number, so it seems to be well calibrated. The exceedances of M2 are far too high, it is not well calibrated.

Since S is a strictly consistent score for $\text{VaR}_{0.95}$, a smaller average score indicates a better accuracy of a model. Hence, M2 has a better accuracy than M1.

- (ii) [2 Points] Explain this assessment explicitly referring to the decomposition of strictly consistent scores and the structure of the two models.

Solution:

Due to the score decomposition, strictly consistent scores simultaneously assess (conditional) calibration and the relevant information. Since M2 has a smaller average score than M1, but is less well calibrated, it must be based on more relevant information. And indeed, that's the case due to the autoregressive structure of the GARCH model in comparison to the static model (no relevant information used).

- (iii) [1 Point] How conclusive are the backtesting results presented in the table? Mention briefly how you could improve their validity.

Solution:

The backtesting results are just point estimates without any uncertainty quantification. To improve their validity, we should perform statistical tests.

E.g., we could test the null hypothesis of calibration or we could perform Diebold–Mariano tests for comparative backtests.

(b) [6 Points] On the space of square integrable random variables L^2 we can define the risk measure

$$\rho(X) = \text{sd}(X) + \mathbb{E}[X], \quad X \in L^2.$$

That is, ρ is the sum of the standard deviation and the mean.

Examine if ρ is a coherent risk measure on L^2 . To this end, check all four axioms explicitly. (You may use properties of the mean and standard deviation shown in the lectures and problem sets without a proof.)

Hint: For the monotonicity, it may be helpful to consider a random variable with a Bernoulli distribution.

Solution:

In the lecture / on the problem sets it was shown that the mean is a coherent risk measure. Moreover, sd fails to be monotone, is translation *invariant* (not equivariant), is positively homogeneous and subadditive.

Translation equivariance For $X \in L^2$ and $c \in \mathbb{R}$ we have

$$\rho(X + c) = \text{sd}(X + c) + \mathbb{E}[X + c] = \text{sd}(X) + \mathbb{E}[X] + c = \rho(X) + c.$$

So translation equivariance holds.

Positive Homogeneity For $X \in L^2$ and $c > 0$ we have

$$\rho(cX) = \text{sd}(cX) + \mathbb{E}[cX] = c\text{sd}(X) + c\mathbb{E}[X] = c\rho(X).$$

So positive homogeneity holds.

Subadditivity Since sd is subadditive and $\mathbb{E}$ is additive, subadditivity holds.

Monotonicity For a random variable $X \in L^2$ with $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$ it holds that $\text{sd}(X) + \mathbb{E}[X] = \sqrt{p(1-p)} + p$.

For $p = 0.8$, this is 1.2. On the other hand, for a random variable Y which is almost surely 1.1, we have that $Y > X$ almost surely and $\text{sd}(Y) + \mathbb{E}[Y] = 1.1 < 1.2$. So monotonicity fails.

In total, ρ is not a coherent risk measure on L^2 .