Prof. P. Cheridito 11 August 2018

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 and paper 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			10
#2			8
#3			10
#4			12
#5			10
Total			50

Question 1 (10 Pts)

a) Define the notion of a coherent risk measure, and give a financial interpretation of each axiom of coherence. (4 Pts)

b) Let $X \sim \text{Par}(\theta, \kappa)$ with cdf

$$F(x) = 1 - \left(\frac{\kappa}{\kappa + x}\right)^{\theta}, \quad x \geq 0,$$

for parameters $\kappa > 0$ and $\theta > 1$. Calculate $\text{VaR}_\alpha(X)$ and $\text{AVaR}_\alpha(X)$. (3 Pts)

c) Let L be a d-dimensional random vector whose components $L_1, \ldots, L_d$ are normally distributed with means $\mu_1, \ldots, \mu_d \in \mathbb{R}$ and variances $\sigma_1^2, \ldots, \sigma_d^2 > 0$. Fix a level $\alpha \in (1/2, 1)$. Is $\text{VaR}_\alpha(L_1 + \cdots + L_d)$ larger if the copula of the random vector L is the independence copula or the comonotonicity copula? (3 Pts)

Solution 1

a) A risk measure $\rho : \mathcal{L} \to \mathbb{R}$ is called coherent if it satisfies the following set of axioms:

- *monotonicity*: $\rho(L_1) \leq \rho(L_2)$ for $L_1 \overset{\text{a.s.}}{\leq} L_2$;
- *translation invariance*: $\rho(L + m) = \rho(L) + m$ for $m \in \mathbb{R}$;
- *subadditivity*: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$;
- *positive homogeneity*: $\rho(\lambda L) = \lambda \rho(L)$ for $\lambda \in \mathbb{R}_+$.

They admit the following interpretation:

- *monotonicity*: Portfolios that have higher losses in every possible scenario are regarded as more risky.
- *translation invariance*: By adding capital worth m, the total risk decreases/increases by the by the same amount. Note that if $m = -\rho(L)$, then we have that

  $$\rho(L + m) = \rho(L - \rho(L)) = \rho(L) - \rho(L) = 0$$

  (L is a loss random variable so $-\rho(L)$ would constitute a capital injection).
- *subadditivity*: Diversification does not lead to an increase in risk.
- *positive homogeneity*: Scaling a position (leveraging) up or down increases or decreases the risk by the same factor. Note that if $\lambda \in \mathbb{N}$ then we have

  $$\rho(\lambda L) = \rho\left(\sum_{i=1}^{\lambda} L\right).$$

  Subadditivity gives us that this should be less or equal to $\lambda \rho(L)$, but since there is no diversification, we require equality.

b) In order to compute AVaR, we first compute VaR. This can be done by simple inversion of the cdf. We obtain

$$\text{VaR}_\alpha(X) = \frac{\kappa}{(1-\alpha)^{\frac{1}{\theta}}} - \kappa = \kappa\left((1-\alpha)^{-\frac{1}{\theta}} - 1\right).$$

Then we have that

$$\begin{aligned}\mathrm{AVaR}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)\,du = \frac{1}{1-\alpha}\int_\alpha^1 \frac{\kappa}{(1-u)^{\frac{1}{\theta}}}du - \kappa \\ &= \kappa\frac{1}{1-\alpha}\frac{\theta}{\theta-1}(1-\alpha)^{1-\frac{1}{\theta}} - \kappa = \kappa\left(\frac{\theta}{\theta-1}(1-\alpha)^{-\frac{1}{\theta}} - 1\right).\end{aligned}$$

c) When the copula of L is the independence copula, then L is jointly normal with mean vector $\mu = (\mu_1, \ldots, \mu_d)^\top$ and a diagonal covariance matrix Σ whose i-th diagonal entry is equal to σ_i^2. Therefore,

$$\sum_{i=1}^d L_i \sim N\left(\sum_{i=1}^d \mu_i, \sum_{i=1}^d \sigma_i^2\right).$$

Let us denote

$$m = \sum_{i=1}^d \mu_i \quad \text{and} \quad s^2 = \sum_{i=1}^d \sigma_i^2$$

Using the formula for VaR of a normally distributed random variable derived in the class, it follows that

$$\mathrm{VaR}_\alpha\left(\sum_{i=1}^d L_i\right) = m + s\,\Phi^{-1}(\alpha),$$

where $\Phi^{-1}(\alpha)$ denotes the α-quantile of $N(0,1)$.

In case the copula of L is the comonotonicity copula, we can use the fact that VaR is a comonotone additive risk measure, which gives us that

$$\mathrm{VaR}_\alpha\left(\sum_{i=1}^d L_i\right) = \sum_{i=1}^d \mathrm{VaR}_\alpha(L_i) = \sum_{i=1}^d \left(\mu_i + \sigma_i\,\Phi^{-1}(\alpha)\right) = m + \left(\sum_{i=1}^d \sigma_i\right)\Phi^{-1}(\alpha).$$

Since

$$s^2 = \sum_{i=1}^d \sigma_i^2 < \sum_{i=1}^d \sigma_i^2 + 2\sum_{i=1}^d \sum_{j=i+1}^d \sigma_i\sigma_j = \left(\sum_{i=1}^d \sigma_i\right)^2$$

implies that

$$s = \sqrt{\sum_{i=1}^d \sigma_i^2} < \sum_{i=1}^d \sigma_i,$$

it is clear that $\mathrm{VaR}_\alpha(L_1 + \cdots + L_d)$ is greater under comonotonicity than it is under independence for all $d \geq 2$ and $\alpha \in (1/2, 1)$ (because in this case we have that $\Phi^{-1}(\alpha) > 0$).

Question 2 (8 Pts)

a) Is every normal variance mixture distribution elliptical? Explain your answer. (4 Pts)

b) Let d financial returns be modeled by the components $X_1, \ldots, X_d$ of a d-dimensional random vector X. Assume X has an elliptical distribution such that $\mathbb{E}[X_i^2] < \infty$ for all $i = 1, \ldots, d$, and $\mathbb{E}[X_1] = \cdots = \mathbb{E}[X_d]$. We want to show that the minimum variance portfolio also minimizes Value-at-Risk. More precisely, denote

$$\Delta := \left\{w \in \mathbb{R}^d : \sum_{i=1}^d w_i = 1\right\}$$

and fix a level $\alpha \in (1/2, 1)$. Then show that the two minimization problems

$$\min_{w\in\Delta} \operatorname{Var}\left(\sum_{i=1}^{d} w_i X_i\right) \quad \text{and} \quad \min_{w\in\Delta} \operatorname{VaR}_\alpha\left(-\sum_{i=1}^{d} w_i X_i\right)$$

have the same minimizer $w^* \in \Delta$. (4 Pts)

Solution 2

a) Let's assume that a random vector $X \in \mathbb{R}^d$ is a normal variance mixture. Then, by definition, X admits a stochastic representation as

$$X \stackrel{(d)}{=} \mu + \sqrt{W} A Z$$

for a random vector $Z \sim N_k(0, I_k)$, a random variable $W \geq 0$, $\mu \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times k}$. By definition, X is also elliptically distributed if we have that

$$X \stackrel{(d)}{=} m + BY$$

for some $m \in \mathbb{R}^d$, $B \in \mathbb{R}^{d\times l}$ and $Y \sim S_l(\psi)$. We can obviously set $l = k$, $m = \mu$ and $B = A$. If we can prove that $\sqrt{W}Z$ is spherical, we can also set $Y = \sqrt{W}Z$ and we are done.

One way to conclude quickly is to realize that $\sqrt{W}Z \sim M_k(0, I_k, \hat{F}_W)$ and to recall that we have seen in the lectures that this distribution is indeed spherical. More precisely, let $U \in \mathbb{R}^{k\times k}$ be an arbitrary orthogonal matrix. Then

$$U\sqrt{W}Z \stackrel{(d)}{=} \sqrt{W}UZ \stackrel{(d)}{=} \sqrt{W}Z$$

because $UZ \sim N_k(0, UU^\top) \iff UZ \sim N_k(0, I_k)$, which shows that $\sqrt{W}Z$ is spherical.

b) Let $X \sim E_d(m, \Sigma, \psi)$ for an $m = (\mu, \dots, \mu)^\top \in \mathbb{R}^d$ with $\mu = \mathbb{E}[X_1]$, a matrix $\Sigma \in \mathbb{R}^{d\times d}$ and a characteristic generator ψ. Note that since elliptical distributions are closed under affine transformations and since we have that $\sum_{i=1}^d w_i = 1$,

$$w^\top X \sim E_1(\mu, w^\top \Sigma w, \psi).$$

By the definition of an elliptical distribution, $w^\top X$ admits a stochastic representation as

$$w^\top X \stackrel{(d)}{=} \mu + \sqrt{w^\top \Sigma w} Y,$$

for some $Y \sim S_1(\psi)$. We therefore obtain by the properties of variance and the positive homogeneity and translation invariance of VaR that

$$\begin{aligned}
\operatorname{Var}\left(w^\top X\right) &= \operatorname{Var}\left(\mu + \sqrt{w^\top \Sigma w} Y\right) = \left(w^\top \Sigma w\right) \operatorname{Var}(Y), \\
\operatorname{VaR}_\alpha\left(-w^\top X\right) &= \operatorname{VaR}_\alpha\left(-\mu - \sqrt{w^\top \Sigma w} Y\right) = -\mu + \sqrt{w^\top \Sigma w}\, \operatorname{VaR}_\alpha(-Y).
\end{aligned}$$

Note that since $-Y \stackrel{(d)}{=} Y$ (because -1 can be interpreted as an orthogonal matrix in $\mathbb{R}^{1\times 1}$), Y is symmetric and with zero mean, so

$$\operatorname{VaR}_\alpha(-Y) = \operatorname{VaR}_\alpha(Y) > 0 \quad \text{for } \alpha \in (1/2, 1).$$

The factors $\operatorname{Var}(Y)$ and $\operatorname{VaR}_\alpha(Y)$ are therefore of the same sign and are independent of w, thus do not affect the value of the minimizer $w^* \in \Delta$. This means that for all $\alpha \in (1/2, 1)$

$$\operatorname*{argmin}_{w\in\Delta} \operatorname{Var}\left(w^\top X\right) = \operatorname*{argmin}_{w\in\Delta} w^\top \Sigma w \quad \text{and} \quad \operatorname*{argmin}_{w\in\Delta} \operatorname{VaR}_\alpha\left(-w^\top X\right) = \operatorname*{argmin}_{w\in\Delta} \sqrt{w^\top \Sigma w}.$$

Since $x \mapsto \sqrt{x}$ is an increasing function, it also does not affect the value of the minimizer w^* and we can drop the square root in the last expression. Since we have reduced both optimization problems to the same one, we are done.

Question 3 (10 Pts)

a) Let (X, Y) be a two-dimensional random vector with Exp(1)-marginals and copula

$$C(u, v) = uv + (1 - u)(1 - v)uv.$$

Does (X, Y) have a density? If yes, can you compute it? (3 Pts)

b) Calculate Spearman's rank correlation between X and Y given in a). (2 Pts)

c) Calculate the coefficient of upper tail dependence λ_u between X and Y given in a). (2 Pts)

d) Let (X, Y) be a two dimensional random vector with cdf

$$\frac{1 - e^{-x} - e^{-y} + e^{-x-y}}{1 - e^{-x-y}}$$

on $\mathbb{R}^2_+$. What are the marginal distributions and copula of (X, Y)? (3 Pts)

Solution 3

a) Using Sklar's theorem, the cdf $F_{X,Y}$ of (X, Y) is given by

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)),$$

where F_X and F_Y are the ' of X and Y, respectively. Using the fact that the margins are Exp(1)-distributed, the above gives

$$\begin{aligned} F_{X,Y}(x, y) &= (1 - e^{-x})(1 - e^{-y}) + e^{-x-y}(1 - e^{-x})(1 - e^{-y}) = (1 - e^{-x})(1 - e^{-y})(1 + e^{-x-y}) \\ &= 1 - e^{-y} - e^{-x} + 2e^{-x-y} - e^{-x-2y} - e^{-2x-y} + e^{-2x-2y} \end{aligned}$$

for $(x, y) \in \mathbb{R}^2_+$. We have that $F_{X,Y} \in C^\infty(\mathbb{R}^2_+)$ so the density $f_{X,Y}$ does exist and is given by

$$f_{X,Y}(x, y) = \frac{\partial^2 F}{\partial x \partial y}(x, y) = 2e^{-x-y} - 2e^{-x-2y} - 2e^{-2x-y} + 4e^{-2x-2y}$$

for $(x, y) \in \mathbb{R}^2_+$.

b) Spearman's rank correlation ρ_S between two random variables X and Y with continuous marginal distributions given by cdfs F_X and F_Y is defined as

$$\rho_S(X, Y) = \rho(F_X(X), F_Y(Y)),$$

where ρ denotes the standard Pearson's linear correlation coefficient. We know from the lecture that ρ_S is independent of the marginal distributions and can be computed as

$$\begin{aligned} \rho_S(X, Y) &= 12 \iint_{[0,1]^2} \big(C(u, v) - uv\big) du dv = 12 \iint_{[0,1]^2} uv(1 - u)(1 - v) du dv \\ &= 12 \left(\int_0^1 u(1 - u) du \right)^2 = 12 \left(\frac{1}{2} - \frac{1}{3} \right)^2 = \frac{1}{3}, \end{aligned}$$

where C denotes the copula of (X, Y).

c) The coefficient of upper tail dependence is defined as

$$\lambda_u = \lim_{\alpha \uparrow 1} \mathbb{P}\left[X > q_X(\alpha) \,|\, Y > q_Y(\alpha)\right].$$

We have seen in the lecture that it is independent of the marginal distributions and can be computed as

$$\lambda_u = \lim_{\alpha \uparrow 1} \frac{1 - 2\alpha + C(\alpha, \alpha)}{1 - \alpha}.$$

Since

$$C(\alpha, \alpha) = \alpha^2 + (1-\alpha)^2 \alpha^2 = 2\alpha^2 - 2\alpha^3 + \alpha^4$$

is, as a function of one variable α, differentiable in α, l'Hospital's rule gives

$$\lambda_u = 2 - \lim_{\alpha \uparrow 1} \frac{d}{d\alpha} C(\alpha, \alpha) = 2 - \lim_{\alpha \uparrow 1} 4\alpha - 6\alpha^2 + 4\alpha^3 = 0.$$

d) The marginal distributions are easily computed as

$$F_X(x) = F(x, \infty) = 1 - e^{-x} \quad \text{and} \quad F_Y(y) = F(\infty, y) = 1 - e^{-y}.$$

We now want to use Sklar's theorem, which states that we can compute the copula of X as $C(u, v) = F\left(q_X(u), q_Y(v)\right)$. We thus need to compute the quantile function of X (the margins are identical). By inverting F_X, we get that $q_X(u) = q_Y(u) = -\log(1-u)$. The copula is therefore given by

$$C(u, v) = \frac{uv}{1 - (1-u)(1-v)} = \frac{uv}{v + u - uv} = \frac{1}{u^{-1} + v^{-1} - 1},$$

which is the Clayton copula with $\theta = 1$.

Question 4 (12 Pts)

Let X be a non-negative random random variable with cdf

$$F_X(x) = \frac{x}{x+1}, \quad x \geq 0.$$

a) Does X have a density? If yes, can you derive it? (2 Pts)

b) Find all $k \in \mathbb{N}$ such that $\mathbb{E}[|X|^k] < \infty$. (2 Pts)

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

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

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

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

for a generalized Pareto distribution $G_{\xi,\beta}$? If yes, for which ξ and β does this hold? (3 Pts)

Solution 4

a) The density f_X of X exists and is given for all $x \geq 0$ by

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{x + 1 - x}{(x+1)^2} = \frac{1}{(x+1)^2}.$$

b) We can conclude that $\mathbb{E}\left[|X|^k\right] = \infty$ for all $k \in \mathbb{N}$ from (c), since there we show that $F_X \in \text{MDA}(H_1)$ and we know that if $F \in \text{MDA}(H_\xi)$ with $\xi > 0$, then $\mathbb{E}\left[|X|^k\right] = \infty$ for $k \geq 1/\xi$.

We can also show this directly. We have that

$$\mathbb{E}\left[|X|^k\right] = \mathbb{E}\left[X^k\right] = \int_0^\infty \frac{x^k}{(x+1)^2} dx.$$

But at ∞, $\frac{x^k}{(x+1)^2} \sim x^{k-2}$ and $\int_a^\infty x^{k-2} dx, a > 0$ is finite if and only if $k < 1$. This implies that the entire integral cannot converge for any $k \in \mathbb{N}$.

c) It will be helpful to rewrite the given cdf as

$$F_X(x) = \frac{x}{x+1} = \frac{x}{x+1} - 1 + 1 = 1 - \frac{1}{x+1}.$$

Since we know from (b) that $\mathbb{E}\left[|X|^k\right] = \infty$ for all $k \in \mathbb{N}$ or simply observing that f_X exhibits a power decay, we would expect that $F_X \in \text{MDA}(H_\xi)$ for $\xi > 0$. This observation helps with constructing our normalizing sequences for instance as $c_n = n$ and $d_n = n - 1$. We then have that

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

as $n \to \infty$ for all $x \geq 0$, which is the GEV distribution with $\xi = 1$. That is $F_X \in \text{MDA}(H_1)$.

d) We can easily derive of simply use the formula for F_u from the class:

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

This gives

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

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

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

if and only if $F_X \in \text{MDA}(H_\xi)$. We have shown in (c) that $F_X \in \text{MDA}(H_1)$, thus (1) holds for $\xi = 1$ and for some function $\beta(u)$ yet to be determined.

We have that

$$\begin{aligned}\left|F_u(x) - G_{1,\beta(u)}(x)\right| &= \left|1 - \frac{u+1}{x+u+1} - 1 + \left(1 + \frac{x}{\beta(u)}\right)^{-1}\right| = \left|\left(\frac{x+\beta(u)}{\beta(u)}\right)^{-1} - \frac{u+1}{x+u+1}\right| \\ &= \left|\frac{\beta(u)}{x+\beta(u)} - \frac{u+1}{x+u+1}\right|,\end{aligned}$$

which is equal to 0 for $\beta(u) = u + 1$. This choice of beta will also render the limit in (1) equal to 0, and we are done.

Question 5 (10 Pts)

a) Write down the specification of a GARCH(1,1) model. (2 Pts)

b) Which stylized facts of daily log-returns can a GARCH(1,1) model capture? (2 Pts)

c) Let the distribution of a d-dimensional random vector X be given by univariate marginal cdfs $F_1, \ldots, F_d$ and a Gaussian copula C_P^{Ga}. Describe an algorithm for simulating X. Justify your approach. (3 Pts)

d) Describe the Peaks-Over-Threshold method. (3 Pts)

Solution 5

a) As seen in the lecture, we say that the process $X = (X_t)_{t \in \mathbb{Z}}$ follows a GARCH(1,1) process if it is stationary and we have for all $t \in \mathbb{Z}$ that

$$X_t = \sigma_t Z_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

for some $\alpha_0, \alpha_1, \beta_1 \geq 0$, where $Z_t \sim \mathrm{SWN}(0, 1)$ for all $t \in \mathbb{Z}$.

b) GARCH(1,1) model can capture the following stylized facts of daily log-returns:

- Volatility varying over time;
- Large (extreme) movements appearing in clusters (volatility clustering);
- Leptokurtic or heavy-tailed log-returns;
- Low correlation of raw log-returns;
- Profound correlation of absolute or squared log-returns;
- Conditional expected log-returns close/equal to zero.

Of these, the ones related to volatility and heavy tails are the most important since some of the other ones can obviously captured also by a sequence of i.i.d. normally distributed random variables with zero mean.

c) Let Φ denote the cdf of $N(0, 1)$ distribution given by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du,$$

and q_i the marginal quantile functions corresponding to F_i and given by

$$q_i(\alpha) = \min\{x \in \mathbb{R} \mid F_i(x) \geq \alpha\}$$

for all $i \in \{1, \ldots, d\}$. Simulation of X can be carried out as follows:

(i) Simulate $Z_i \sim N(0, 1)$, $i \in \{1, \ldots, d\}$ and define $Z = (Z_1, \ldots, Z_d)^\top$;

(ii) Compute the Cholesky decomposition $P = AA^\top$ of P;

(iii) Assign $Y = AZ$;

(iv) Assign $U = (\Phi(Y_1), \ldots, \Phi(Y_d))^\top$, where Y_i is the i-th component of Y;

(v) Return $X = (q_1(U_1), \ldots, q_d(U_d))^\top$, where U_i is the i-th component of U.

The X simulated this way has a distribution with copula C_P^{Ga} and margins $F_1, \ldots, F_d$ by Sklar's theorem and by the definition of multivariate normal distribution.

d) Consider the losses $X_1, \dots, X_n \sim F \in \mathrm{MDA}(H_\xi)$ and let N_u denote the number of these losses that exceed a selected large threshold u. According to the Pickands-Balkema-de Haan theorem, the excesses $Y_i = X_i - u$, $i \in \{1, \dots, N_u\}$ are roughly distributed as $G_{\xi,\beta}$. The estimates $\hat{\xi}$ and $\hat{\beta}$ can (often) be computed using the MLE, that is by maximizing the log-likelihood function

$$\ell(\xi, \beta; Y_1, \dots, Y_{N_u})) = \begin{cases} -N_u \log \beta - (1 + 1/\xi) \sum_{i=1}^{N_u} \log(1 + \xi Y_i/\beta) & \xi \neq 0, \\ -N_u \log \beta - \sum_{i=1}^{N_u} Y_i/\beta & \xi = 0 \end{cases}$$

with respect to ξ and β such that $\beta > 0$ and $1 + \xi Y_i/\beta > 0$ for all $i = 1, \dots, N_u$.

The only remaining problem is the selection of the threshold u. We have seen in the lecture that provided that the mean excess function $e(v) = \mathbb{E}[X - v \,|\, X > v]$ of $G_{\xi,\beta}$ exits, it is an affine function of a threshold $v \geq u$. We can therefore select the threshold u by constructing a mean excess plot

$$(X_{(i)}, e_n(X_{(i)})), \quad i \in \{1, \dots, n\},$$

where $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ are the ordered data and

$$e_n(v) = \frac{\sum_{i=1}^{n} (X_i - v) 1_{\{X_i > v\}}}{\sum_{i=1}^{n} 1_{\{X_i > v\}}}$$

is the sample mean excess function, and finding a point u^* such that plot grows roughly linearly above u^*.