Statistique des risques extrêmes

Support de cours II

Nicolas Jeannelle

Direction des Risques – Risques Financiers

Confédération Nationale du Crédit Mutuel

Sommaire

Introduction

I. Value-at-Risk : Fondements et modélisation

II. Value-at-Risk par modélisation des extrêmes

- La Théorie des Valeurs Extrêmes
- Modélisation : Méthode *Bloc Maxima*
- Modélisation : Méthode *Peak Over Threshold*

I. Value-at-Risk dynamique

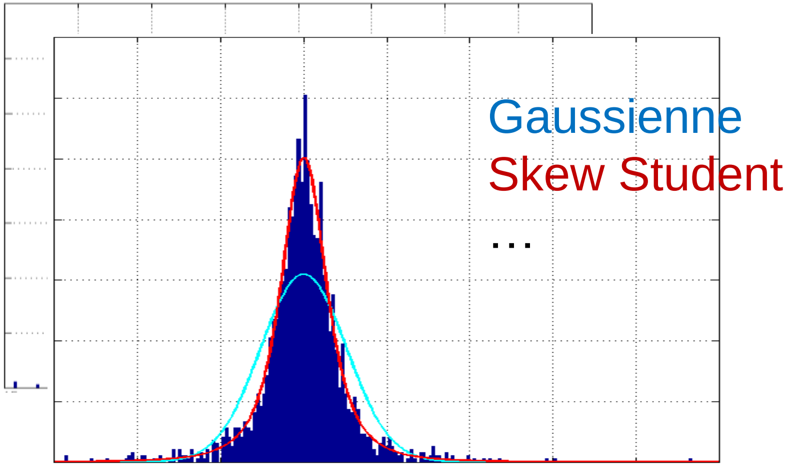
Conclusion

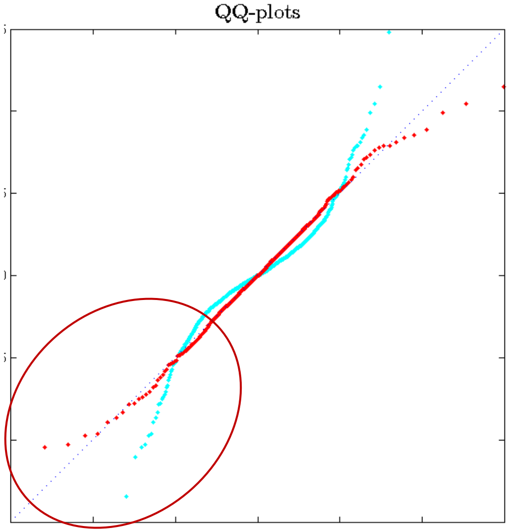
Sommaire

1. La Théorie des Valeurs Extrêmes

Introduction

- ***Rappel*** **: La Value-at-Risk se définit par 3 paramètres : le niveau de confiance α, l'horizon h et la distribution des pertes F.**
  - Seul F est réellement à la main de l'analyste
  - Différentes distributions peuvent être testées
  - Leur ajustement est évalué par des tests d'adéquation et des considérations graphiques
  - Et la mesure de risque est obtenue comme quantile d'ordre α





- ***Mais*** **:**
  - Est-il logique de modéliser **l'intégralité du profil de risque** pour n'en extraire qu'un **quantile extrême** ?
  - Qui plus est par une méthode accordant un **poids identique** à chaque observation …

➢ Se focaliser sur les extrêmes via la **Théorie des Valeurs Extrêmes**.

1. La Théorie des Valeurs Extrêmes

Préliminaires

- ***Motivation historique*** : **Utilisée pour la première fois pour quantifier des hauteurs de digues pouvant contenir une montée des eaux survenant tous les 10 000 ans.**
  - Utilisée aujourd'hui pour la prévision de catastrophes naturelles (inondations, vents violents…) ou autres (chute de valeurs de marché…)

- ***Principe*** : **Quantifier la loi des « extrêmes » afin de pouvoir estimer de manière précise des quantiles d'ordre très élevé, même non observables dans l'échantillon.**

- ***Approche*** : **On s'intéresse à la loi du maximum d'un échantillon de variables *iid***

$$X_1, \dots, X_n \qquad X_i \sim F \ \text{inconnue}$$

Soit M_n le maximum de cet échantillon

$$M_n = \max(X_1, \dots, X_n)$$

alors

$$\mathbb{P}[M_n \leq x] = \mathbb{P}[\{X_1 \leq x\} \cap \dots \cap \{X_n \leq x\}] = \prod_i \mathbb{P}[X_i \leq x] = F^n(x)$$

Finalement $M_n \sim F^n$

1. La Théorie des Valeurs Extrêmes

Théorème de Fisher - Tippet

- ***Enoncé fondamental sur la limite du maximum*** **:**

On suppose qu'il existe

$a_n > 0, b_n$	**des suites de constantes**
H	**une distribution non dégénérée**

telles que

$$\lim_{n\to\infty} \frac{M_n - b_n}{a_n} = H \quad \text{ou encore} \quad \lim_{n\to\infty} \mathbb{P}\left[\frac{M_n - b_n}{a_n} \leq x\right] = \lim_{n\to\infty} F^n(a_n x + b_n) = H(x)$$

alors *H* ne peut être que de 3 types :

- Fréchet : $H_{1,\gamma}(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\gamma}) & x > 0, \gamma > 0 \end{cases}$
- Weibull : $H_{2,\gamma}(x) = \begin{cases} \exp(-(-x)^{\gamma}) & x < 0, \gamma < 0 \\ 1 & x \geq 0 \end{cases}$
- Gumbel : $H_{3,0}(x) = \exp\big(-\exp(-x)\big) \qquad \gamma = 0$

➢ On dit alors que F est dans le **domaine d'attraction** de Gumbel, Fréchet ou Weibull.

1. La Théorie des Valeurs Extrêmes

Théorème de Fisher - Tippet

- ***Théorème Central Limite*** :
  - **Indépendamment de F…**
  - … sous certaines hypothèses …
  - … la **loi de la moyenne** pour de grands échantillons *iid* converge vers une **distribution limite**

$$\frac{\overline{X_n} - \mu}{\sigma / \sqrt{n}} \rightarrow \mathcal{N}(0, 1)$$

- ***Théorème de Fisher-Tippet*** :
  - **Indépendamment de F…**
  - … sous certaines hypothèses …
  - … la **loi du maximum** pour de grands échantillons *iid* converge vers une **distribution limite**

$$\frac{M_n - b_n}{a_n} \rightarrow H$$

- ***On va voir que*** :
  - On peut déterminer le domaine d'attraction par la connaissance de F
  - Il existe une distribution capable d'unifier les 3 distributions limites

1. La Théorie des Valeurs Extrêmes

Caractérisation des domaines d'attraction

- ***Principe*** **: Une étude du comportement de la fonction de survie $\overline{F}$ en les queues de distribution permet de caractériser entièrement le domaine d'attraction.**

- ***Définitions*** **:**
  - Le **point terminal** de F est défini par $\omega = \inf\{x; F(x) = 1\}$
  - U est à **variations régulières d'indice δ** $\in \mathbb{R}$ si $\lim_{x\to\infty} \frac{U(tx)}{U(x)} = t^{\delta}$. On écrit $U \in \mathcal{V}_{\delta}$
  - Si $\delta = 0$, U est à **variations lentes**. On écrit $U \in \mathcal{V}_0$
  - Si $U \in \mathcal{V}_{\delta}$ alors $U(x) = x^{\delta}L(x)$ avec $L \in \mathcal{V}_0$
  - Si $L \in \mathcal{V}_0$ alors $L(x) = c(x)\exp\left\{\int_1^x \frac{\Delta(u)}{u}du\right\}$ avec $c(x) \to c > 0, \Delta(x) \to 0$

- ***Caractérisation des domaines d'attraction*** **:**
  - F appartient au domaine d'attraction de **Fréchet** d'indice $\gamma > 0$ ssi $\overline{F} \in \mathcal{V}_{-1/\gamma}$
  - F appartient au domaine d'attraction de **Weibull** d'indice $\gamma < 0$ ssi

    ω est fini

    $\overline{F_*} \in \mathcal{V}_{-1/\gamma}$ avec $F_* = F\left(\omega - \frac{1}{x}\right)$
  - F appartient au domaine d'attraction de **Gumbel** ssi $\exists z < \omega \leq \infty$ tel que

$$\overline{F}(x) = c(x)exp\left\{-\int_z^x \frac{1}{a(t)}dt\right\}$$

1. La Théorie des Valeurs Extrêmes

Caractérisation des domaines d'attraction

- ***Convergence ?***
  - **En théorie** : la convergence n'est pas assurée, sauf si ***F*** **est continue** et si la queue de distribution est **« suffisamment régulière »**.
  - **En pratique** : il est souvent justifié de supposer que la loi du maximum d'un phénomène observé converge.
- ***Domaines d'attraction de lois usuelles*** :

Domaine d'attraction	Gumbel $\gamma = 0$	Fréchet $\gamma > 0$	Weibull $\gamma < 0$
Loi	*Normale* *Exponentielle* *Lognormale* *Gamma* *Weibull*	*Cauchy* *Pareto* *Burr*	*Uniforme* *Bêta*

  - Pour un même domaine d'attraction, les vitesses de convergence de 2 lois peuvent être très différentes.
  - Grossièrement :
    - Les lois suffisamment régulières à queue fine appartiennent au domaine d'attraction de Gumbel
    - Les lois suffisamment régulières à queue épaisse appartiennent au domaine d'attraction de Fréchet
    - Les lois suffisamment régulières à support fini appartiennent au domaine d'attraction de Weibull

1. La Théorie des Valeurs Extrêmes

La distribution GEV

- ***Principe*** **: Les 3 profils limites de distribution du théorème de Fisher-Tippet peuvent être définis par une seule distribution paramétrique commune, la distribution GEV.**
- ***Définition*** **: La *Generalized Extreme Value Distribution* est une loi paramétrique définie par sa fonction de répartition**

$$G_{\mu,\sigma,\xi}(x) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \qquad \text{avec } 1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0$$

- μ est le paramètre de **localisation**
  σ est le paramètre **d'échelle**
  ξ est le paramètre de **forme**

- ξ **caractérise entièrement** le profil de la distribution limite:
  - $\xi > 0$ donne la distribution de Fréchet, support borné par valeurs inférieures
  - $\xi < 0$ donne la distribution de Weibull, support borné par valeurs supérieures
  - $\xi = 0$ donne la distribution de Gumbel, support non borné

- **La GEV est la seule distribution limite non dégénérée pour un échantillon de maxima.**

- **La GEV est la seule distribution max-stable…**

Sommaire

2. Modélisation par maxima par blocs

Méthodologie

- ***Principe*** **: Estimer une distribution GEV afin de modéliser la loi du maximum de l'échantillon.**

- ***Mise en œuvre*** **:**
  - **i.** Construire un **échantillon de maxima**
  - **ii.** Estimer les paramètres de la loi GEV
  - **iii.** Calculer la mesure de risque associée
  - **iv.** Evaluer l'adéquation aux données

- ***Remarques*** **:**
  - Bien que l'on possède des résultats concernant les domaines d'attraction selon F, **on ne s'intéresse pas à la distribution F directement**.
  - En effet, on ne connait jamais F, on pose une hypothèse qui apparait cohérente.
  - Par exemple, si les tests globaux d'adéquation semblent acceptés l'hypothèse selon laquelle F est normal, il se peut que la loi de Fréchet soit plus adaptée à la modélisation des extrêmes.

2. Modélisation par maxima par blocs

Echantillon des maxima et loi adaptée

- ***Construction de l'échantillon des maxima*** :
  - On dispose d'un échantillon $X_1, \dots, X_n$ de variables *iid*
  - On découpe cet échantillon en k **blocs disjoints de même taille** s, sur lesquels on calcule le maximum :

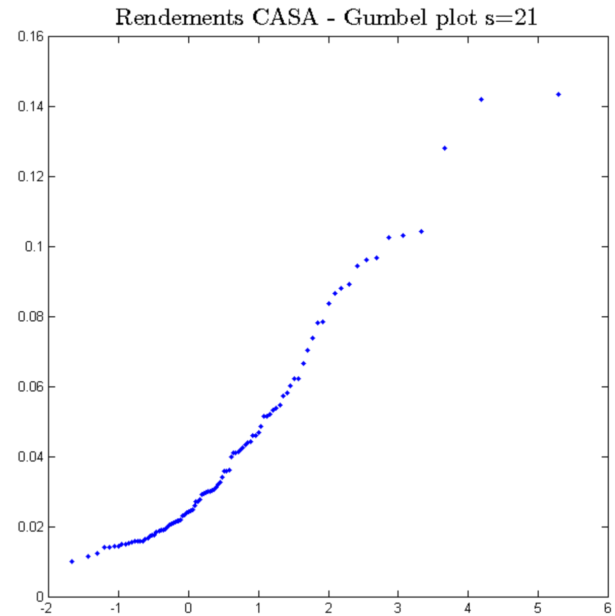
$$Y_j = \max\left(X_{s*(j-1)+1}, \dots, X_{s*(j-1)+s}\right), \quad j \in [\![1, k]\!]$$ que l'on suppose aussi *iid*

  - avec s…
    - devant être suffisamment grand pour satisfaire les conditions asymptotiques du théorème
    - mais suffisamment faible pour obtenir un échantillon de maxima de taille convenable

- ***Choix de la distribution GEV*** :
  - Intuitivement, une optimisation des 3 paramètres ne donnera pas la valeur $\xi = 0$.
  - Utilisation du **Gumbel plot** :

$$\left\{\left(-log\left(-log\left(\frac{i-1/2}{k}\right)\right); Y_{(i)}\right),\ i \in [\![1, k]\!]\right\}$$

  - Si la distribution adaptée est celle de Gumbel alors le Gumbel plot est **linéaire**
  - Si la courbe obtenue présente une **courbure**, alors la distribution adaptée est plutôt Fréchet ou Weibull



2. Modélisation par maxima par blocs

Estimation des paramètres

- ***Estimation par maximum de vraisemblance*** :
  - La **densité** de la loi GEV s'écrit :

$$g_{\mu,\sigma,\xi}(y) = \frac{1}{\sigma}\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-(1+\xi)/\xi} \exp\left\{-\left[1+\xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

  - L'échantillon des maxima par blocs $(y_1, \dots, y_k)$ étant supposé *iid*, la **log vraisemblance** est donnée par

$$\ell_{\mu,\sigma,\xi}(y_1, \dots, y_k) = \sum_{i=1}^{k} \log\left(g_{\mu,\sigma,\xi}(y_i)\right)$$

  - On obtient le **paramétrage optimal $\hat{\theta}$** en maximisant la log vraisemblance par rapport au triplet (μ, σ, ξ)

$$\hat{\theta} = \underset{\mu,\sigma,\xi}{\arg\max}\ \ell_{\mu,\sigma,\xi}(y_1, \dots, y_k) = \underset{\mu,\sigma,\xi}{\arg\min}\ -\ell_{\mu,\sigma,\xi}(y_1, \dots, y_k)$$

- ***Remarques*** :
  - Si $\xi = 0$ alors $g_{\mu,\sigma,0}(y) = \frac{1}{\sigma}\exp\left\{-\left(\frac{y-\mu}{\sigma}\right)\right\}\exp\left\{-\exp\left\{-\left(\frac{y-\mu}{\sigma}\right)\right\}\right\}$
  - Si $\xi > 0$ alors attention aux valeurs de ξ pouvant conduire à une moyenne ou une variance infinie.

- ***Dans la pratique*** : la loi GEV est une distribution usuelle généralement **disponible dans les packages** avec différentes fonctions déjà existantes (*.fit*, *.pdf*, *.cdf*, etc.).

2. Modélisation par maxima par blocs

Mise en œuvre de l'évaluation de la VaR par BM

1. *Préliminaire* :

- Si l'on souhaite modéliser des **minima** (comme c'est le cas pour les rendements), <u>on modélise alors le maximum de l'opposé de la série</u>
- On pourrait également procéder à un changement de variables dans les écritures…

2. *Choix de s* :

- On rappelle que s doit correspondre à un **arbitrage** entre le respect des conditions asymptotiques et nombre de points dans l'échantillon
- Il peut correspondre à 1 mois ($s = 21$ jours), 1 an ($s = 251$ jours), etc…

3. *Estimation des paramètres* :

- Par Maximum de vraisemblance
- Par Méthode des Moments…

4. *Validation ex-ante* :

- Graphe des densités,
- QQ-plot,
- Tests d'adéquation

2. Modélisation par maxima par blocs

Mise en œuvre de l'évaluation de la VaR par BM

5. ***Redressement du niveau de quantile*** :
Comme on ne se concentre que sur l'échantillon des maximas, l'estimation de la VaR à partir de la GEV ne doit pas se faire de niveau α mais de niveau α_{BM} défini par

$$\alpha_{BM} = \alpha^s$$

- Provient du fait que le quantile redressé est tel que $F^s(\alpha_{BM}) = F(\alpha)$
- Formule pouvant être approchée par DL à l'ordre 1 par $1 - \alpha_{BM} = s \times (1 - \alpha)$

6. ***Calcul de la VaR*** :
La VaR de niveau de confiance α à horizon d'investissement h modélisée par la loi GEV est donnée par

$$VaR_h(\alpha) = -G_{\hat{\mu},\hat{\sigma},\hat{\xi}}{}^{-1}(\alpha_{BM})$$

Remarque : $G_{\hat{\mu},\hat{\sigma},\hat{\xi}}{}^{-1}(\alpha_{BM})$ *si on se place en convention risque*

7. ***Validation ex-post*** :
Calcul du nombre d'exceptions

Sommaire

3. Modélisation par distribution des excès

Introduction

- ***Idée*** : Plutôt que d'estimer la loi du maximum de l'échantillon, on va s'intéresser au comportement des **pertes au-delà d'un certain seuil**.
- ***Principe*** :
  - Soit un échantillon de variables *iid* $X_1, \dots, X_n$ $\quad X_i \sim F$ inconnue
  - Etant donné un **seuil u élevé**, on dénombre N_u **pertes extrêmes** positives dépassant ce seuil
  - On va s'intéresser à la **loi des excès** de ces pertes

$$Z_{u,i} = \widetilde{X}_i - u\,, \qquad i \in [\![1, N_u]\!]$$
$$\text{tel que} \quad \exists j \quad \widetilde{X}_i = X_j \text{ et } X_j > u$$

  - On suppose également l'échantillon $Z_{u,1}, \dots, Z_{u,N_u}$ *iid*
  - Cela revient à s'intéresser à la loi des excès des statistiques d'ordre élevées
- ***Enjeux*** :
  - Existe-t-il une distribution adaptée à cette problématique ?
  - Comment choisir u de manière optimale ?
- ***Formulation*** : On s'intéresse à la **fonction de répartition des excès F_u**

$$F_u(y) = \mathbb{P}[Z \le y | X > u] = \frac{\mathbb{P}[u < X \le y + u]}{\mathbb{P}[X > u]} = \frac{F(y+u) - F(u)}{1 - F(u)}$$

3. Modélisation par distribution des excès

Comportement Asymptotique

- ***Théorème de Balkema - de Haan - Pickands*** :

Soit un échantillon de variables aléatoires *iid* $(X_1, \dots, X_n)$

$M_n = \max(X_1, \dots, X_n)$

Z_u **les excès de l'échantillon au-delà d'un seuil** u

Si

$$M_n \underset{n\to\infty}{\sim} GEV(\mu, \sigma, \xi)$$

Alors

$$Z_u \underset{u\to\infty}{\sim} GPD(\tilde{\sigma}, \xi)$$

- La distribution limite est la **Generalized Pareto Distribution** définie par

$$\lim_{u\to\infty} F_u(y) = H_{\tilde{\sigma},\xi}(y) = \begin{cases} 1 - \left(1 + \xi \dfrac{y}{\tilde{\sigma}}\right)^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp\left(-\dfrac{y}{\tilde{\sigma}}\right), & \xi = 0 \end{cases}$$

- $\tilde{\sigma} = \sigma + \xi(u - \mu) > 0$
- Si $\xi \geq 0$ alors le support est $\mathbb{R}_+$, et $[0, -1/\xi]$ sinon
- La $GPD_{1,\xi}$ regroupe 3 distributions selon ξ : Pareto ($\xi > 0$), Pareto II ($\xi < 0$), Exponentielle ($\xi = 0$)
- Si la **distribution des excès est** $H_{\tilde{\sigma},\xi}(y)$ alors **celle des extrêmes est** $u + H_{\tilde{\sigma},\xi}(y)$

3. Modélisation par distribution des excès

Mise en œuvre

- ***Dans la pratique*** :
  - On suppose que les excès suivent exactement une GPD pour un seuil « suffisamment élevé »
  - On définit un seuil u à l'aide d'une méthode adaptée
  - On estime les paramètres de loi sur l'échantillon des excès $(z_1, \dots, z_{N_u})$
  - On vérifie *a posteriori* que l'hypothèse de distribution des excès est justifiée

- ***Détermination du seuil u*** : **Elle s'effectue généralement à l'aide d'un outil graphique : le Mean-Excess plot.**

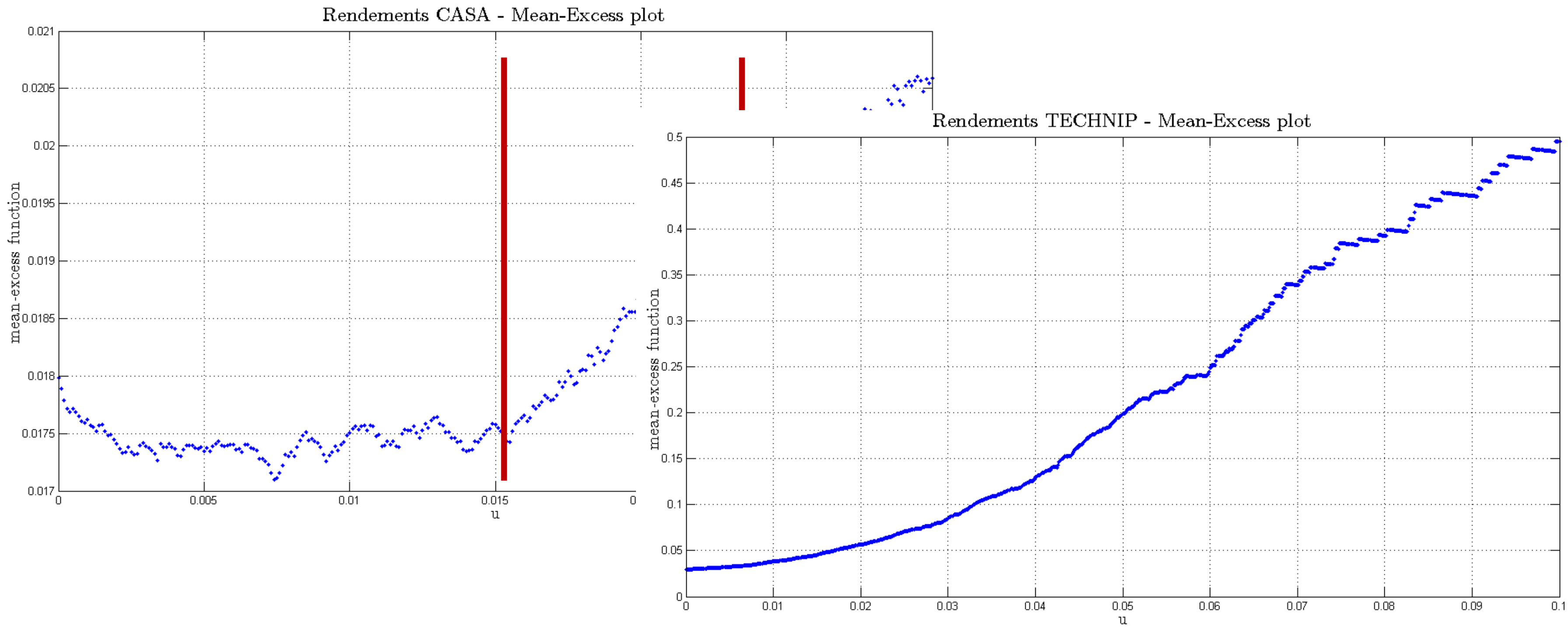
  Il se définit par

$$\left\{\left(u\,;e_n(u)\right), \qquad u \in \left[x_{(1)}, x_{(n)}\right]\right\}$$

  - $e_n(u) = \frac{1}{N_u}\sum_{i=1}^{n}(x_i - u)_+$ est l'estimateur empirique de $e(u) = \mathbb{E}[X - u|X > u]$, la moyenne des excès (**mean excess function**)
  - **Propriété** : pour la loi GPD, $e(u) = \frac{\sigma+\xi u}{1-\xi}$
  - Si les excès au-delà de u suivent une loi GPD, alors le **mean-excess plot a un comportement linéaire**

3. Modélisation par distribution des excès

Mean-Excess Plot



- ***Remarques*** **:**
  - **u doit être suffisamment élevé pour respecter les conditions asymptotiques…**
  - **…mais doit également être calibré de façon à observer assez de données dans l'échantillon.**
  - **Le *mean-excess plot* doit être tracé sur des abscisses positives pour être lisible.**
  - **Il n'est pas toujours facile de conclure quant à la valeur de u**

3. Modélisation par distribution des excès

Estimation des paramètres

- ***Estimation par maximum de vraisemblance*** **: à partir de la densité**

$$h_{\sigma,\xi}(y) = \frac{1}{\sigma}\left[1 + \xi\frac{y}{\sigma}\right]^{-(1+\xi)/\xi} \qquad \ldots$$

- ***Estimation par méthode des moments pondérés*** **: On la privilégie à la méthode des moments classiques car certains moments peuvent ne pas être définis.**

  - On s'intéresse au **moment d'ordre r** $\in \mathbb{R}_+$ suivant :

$$\omega_r(\sigma,\xi) = \int_0^1 y^r \cdot \bar{H}_{\sigma,\xi}{}^{-1}(y)\,dy = \int_0^1 y^r \cdot \frac{\sigma}{\xi}\left(y^{-\xi} - 1\right)dy = \frac{\sigma}{(r+1)(r+1-\xi)}$$

  - Qui permet d'obtenir à partir des **moments théoriques** pour $r = 0$ et $r = 1$ l'écriture des paramètres

$$\sigma = \frac{2\omega_0\omega_1}{\omega_0 - 2\omega_1} \quad \text{et} \quad \xi = 2 - \frac{\omega_0}{\omega_0 - 2\omega_1}$$

  - Qui sont estimés par les **moments empiriques** calculés sur l'échantillon des excès $z_1, \ldots, z_{N_u}$ de fonction de répartition empirique $\widehat{F}$

$$\widehat{\omega}_r(\sigma,\xi) = \frac{1}{N_u}\sum_{i=1}^{N_u} z_i \widehat{F}^{\,r}(z_i), \qquad r = \{0, 1\}$$

- ***Dans la pratique*** **: la loi GPD est une distribution usuelle généralement disponible dans les packages avec différentes fonctions déjà existantes (*.fit*, *.pdf*, *.cdf*, etc.).**

3. Modélisation par distribution des excès

Mise en œuvre de l'évaluation de la VaR par POT

1. ***Préliminaire*** : idem BM
   - Si l'on souhaite modéliser des **minima** (comme c'est le cas pour les rendements), <u>on modélise alors le maximum de l'opposé de la série</u>
   - On pourrait également procéder à un changement de variables dans les écritures…

2. ***Détermination de*** u :
   - On rappelle que u doit correspondre à un **arbitrage** entre le respect des conditions asymptotiques et nombre de points dans l'échantillon
   - Il se détermine par **déduction graphique subjective**

3. ***Estimation des paramètres*** :
   - Par Maximum de vraisemblance
   - Par Méthode des Moments pondérés

4. ***Validation ex-ante*** :
   - Graphe des densités,
   - QQ-plot,
   - Tests d'adéquation

3. Modélisation par distribution des excès

Mise en œuvre de l'évaluation de la VaR par POT

5. ***Redressement du niveau de quantile*** :
Comme on ne se concentre que sur l'échantillon des maximas, l'estimation de la VaR à partir de la GEV ne doit pas se faire de niveau α mais de niveau α_{BM} défini par

$$1-\alpha_{POT}=\frac{n}{N_u}\times(1-\alpha)$$

Remarque : $\alpha_{POT}<\alpha$ *car* $n/N_u>1$

6. ***Calcul de la VaR*** :
La VaR de niveau de confiance α à horizon d'investissement h modélisée par la loi GPD est donnée par

$$VaR_h(\alpha)=-H_{\hat{\sigma},\hat{\xi}}^{-1}(\alpha_{POT})-u$$

Remarque : $H_{\hat{\sigma},\hat{\xi}}^{-1}(\alpha_{POT})$ *si on se place en convention risque*

7. ***Validation ex-post*** :
Calcul du nombre d'exceptions

Bilan

- 2 méthodes très **semblables sur la forme** …
  - Forme des fonctions de répartition
  - Paramètres à estimer de manière graphique
  - Compromis dans le choix du sous échantillon
  - Lien avec le théorème de BdHP
- … mais très **différentes sur le fond** …
  - Loi du maximum
  - Loi des excès
- … qui permettent de mieux appréhender le comportement des extrêmes.
- Concernant la prise en compte des phénomènes de ***clustering***
  - La méthode POT est plus appropriée car elle extrait les extrema au-delà d'un seuil, et prend ainsi en compte toutes les rentabilités extrêmes
  - Alors que la méthode BM n'en extrait qu'un seul par période s