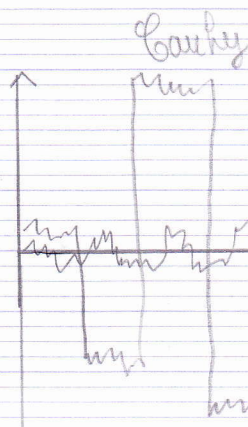


(∞ ∞) $x = L(m) m^{\frac{1}{\alpha}} x_m$ avec $L(m)$ variation lente donc en gal on le renvoie dans x_m .

des maxima: $\mathbb{P}(M_m \geq x)$.

A - Convergence des 3 types de lois.

Rappel: On s'intéresse aux lois telles que $\{M_m > x\}$ déterminent $\{S_m > x\}$. On a introduit la classe des lois sous-exponentielles S qui vérifient cette propriété. 2 comportements distincts sont possibles pour le seuil x :



$\xrightarrow{(\infty \infty)}$ $x = \sqrt{m} x_m$ (à la limite, ça se comporte comme $\mathcal{P}(0, \sqrt{m})$)
 lorsque FES de variance finie.
 $\xrightarrow{(\infty \infty)}$ $x = m^{\frac{1}{\alpha}} x_m$ lorsque FE RV(α) $\alpha \leq 2$ dit de variance infinie (seul plus élevé) (ex: Cauchy)

En résumé, pour FES, connaissant α ($\alpha = +\infty$ dans le cas log-normal, Weibull): degré du 1^{er} moment = $+\infty$

$$\alpha \wedge 2 = \min(\alpha, 2) \quad \mathbb{P}(S_m > \underbrace{m^{\frac{1}{\alpha \wedge 2}} x_m}_{> \sqrt{m}}) \underset{m \rightarrow +\infty}{\sim} \mathbb{P}(M_m > m^{\frac{1}{\alpha \wedge 2}} x_m)$$

toujours

Pour alors étudier la probabilité d'occurrence de ces événements avec $\mathbb{P}(M_m > m^{\frac{1}{\alpha \wedge 2}} x_m)$.

$\mathbb{P}(X \leq ax+b)$
 $= F(ax+b)$
 $= G(x)$
 $= \mathbb{P}(Y \leq x)$

def: On dit que X et Y sont du même type de loi si $\exists a > 0$ et $b \in \mathbb{R}$ tq $Y \sim aX+b$:
 si $X \sim F$ et $Y \sim G$ alors $F(ax+b) = G(x)$ Y est une transformation affine de X
 -ex: toutes les lois normales sont du même type

(imp) th: (Fischer - Tippett): si $\exists (a_m)$ et $(b_m) > 0$, une fonction de répartition G , une w.a. Z telles que:

$$b_m^{-1} (M_m - a_m) \xrightarrow[m \rightarrow +\infty]{\mathcal{L}} Z$$

alors $Z \sim G$ et du même type que 3 lois:

→ Fréchet: $\Phi_\alpha(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ e^{-x^\alpha} & \text{si } x > 0 \end{cases}, \alpha > 0$

→ Weibull négative: (\neq Weibull classique mais c'est sa symétrisée)

$$\Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha} & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}, \alpha > 0$$

→ Gumbel: $\Lambda(x) = e^{-e^{-x}}, x \in \mathbb{R}$

Les lois de type Φ_α, Ψ_α et Λ sont les lois max-stables.

exo 13: • montrer que $X \sim \Lambda \Leftrightarrow e^{\frac{x}{\alpha}} \sim \Phi_\alpha$
 $\Leftrightarrow -e^{-\frac{x}{\alpha}} \sim \Psi_\alpha$

(imp) • montrer que: si $X \sim$ Fréchet alors
 $M_n \sim n^{\frac{1}{\alpha}} X$ (idée: $\mathbb{P}(M_n > x) = 1 - \mathbb{P}(M_n \leq x)$
 et les $(X_i)_i$ sont iid).

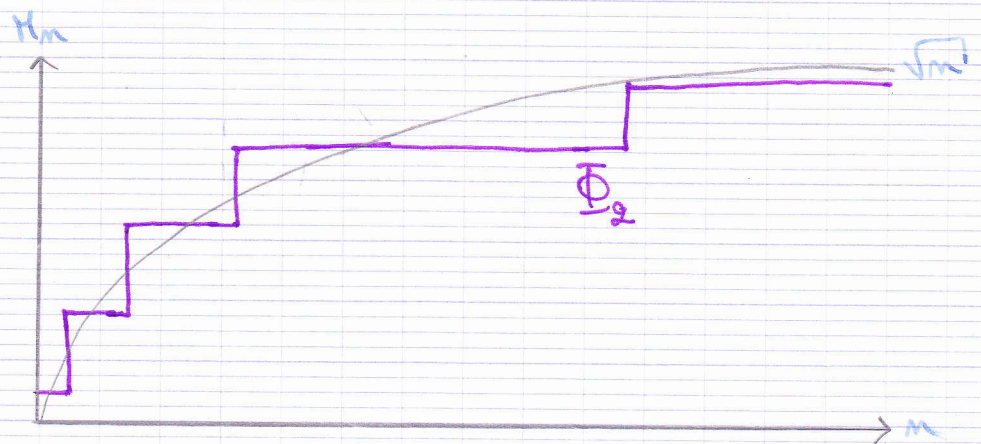
$M_n \sim n^{\frac{1}{\alpha} \leq 0} X$

→ si $X \sim$ Weibull alors

→ si $X \sim$ Gumbel alors

$M_n \sim X + \ln(n)$.

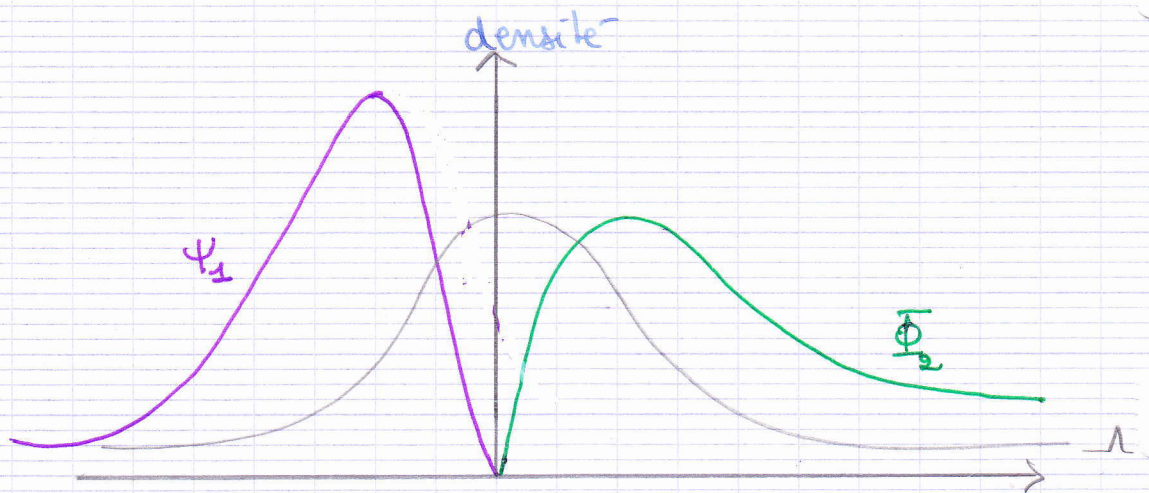
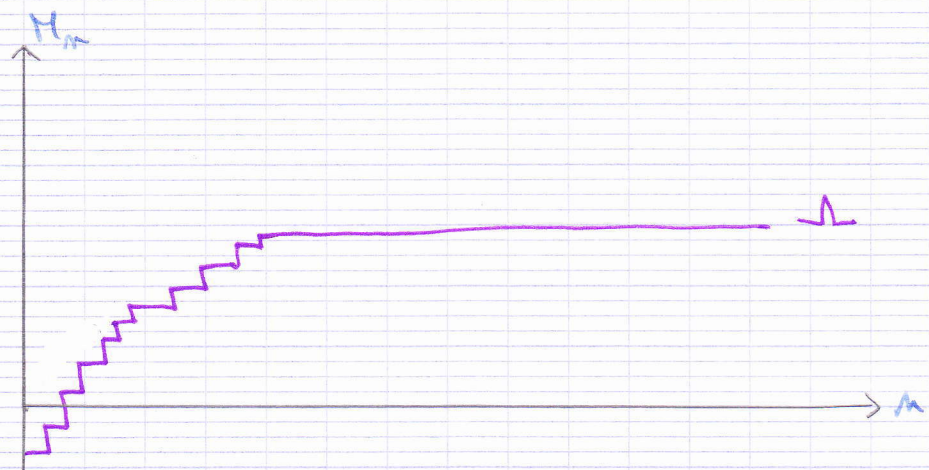
Graphes des records: (On trace M_n en fonction de n)



il n'y a pas de limite du maximum ici.



très rapidement, on est proche de la limite (la limite du maximum est 0)



vitesse de convergence = b_m (croissante en général) (8)

constante de renormalisation = $\frac{1}{b_m}$

exco 14: Montrez que $m M_m \xrightarrow{\mathcal{L}} \Psi_1$ pour $X \sim \mathcal{U}([-1, 0])$ (ici la vitesse est $1/m$).

• Montrez que $M_m - \ln(n) \xrightarrow{\mathcal{L}} \Lambda$ pour $X \sim \text{Exp}(1)$ il n'y a pas d'ordre de grandeur (ici constante de renormalisation est 1)

• Montrez que $\frac{\pi}{n} M_m \xrightarrow{\mathcal{L}} \Phi_1$ pour $(\infty \otimes)$

$X \sim \text{Cauchy} = G_1$ (ici, on renormalise par une puissance négative de n) ($f(x) = \frac{1}{\pi(1+x^2)}$).

(M_n croît de manière positive \Rightarrow on divise par n)

B - Max-domaine d'attraction.

def: Le max-domaine d'attraction, $\text{MDA}(G)$, pour $G = \Psi_x, \Psi_x, \Lambda$, est l'ensemble des lois F tel que si $X_1, \dots, X_n \sim F$ iid, $\exists (a_n), (b_n) > 0$ tels que $b_n^{-1} (M_n - a_n) \xrightarrow{\mathcal{L}} G$.

exco: $\rightarrow \mathcal{U}([-1, 0])$ est le max-domaine d'attraction de la Fréchet Ψ_1 .

$\rightarrow \text{Exp}(1) \in \text{MDA}(\Lambda)$

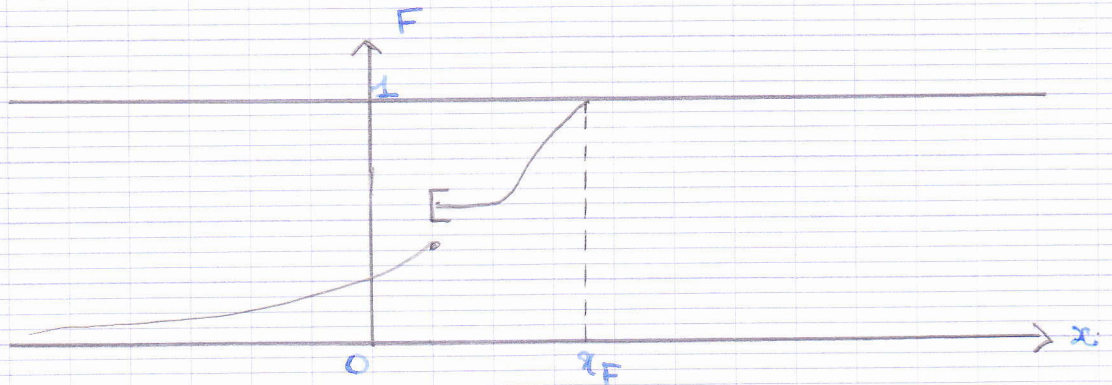
$\rightarrow \text{Cauchy} \in \text{MDA}(\Phi_1)$

Rq: Comme Ψ_x (Weibull) est une loi dont le support est $]-\infty, 0]$, bornée par 0, si $F \in \text{MDA}(\Psi_x)$ alors $\exists x_F > 0$ tel que $x_F = F^{-1}(1)$ (le support est inclus dans $]-\infty; x_F]$)
 F croissante par forcément continue

notation: F^{-1} est l'inverse généralisée de F , ie $F^{-1}(t) = \inf \{ x \in \mathbb{R}, F(x) \geq t \}$

($\infty \otimes$) $M_n \xrightarrow{p.s.} -\infty$ dans le cas d'une Cauchy donc il faut diviser par n pour avoir Φ_1 . Lorsque $X \sim \mathcal{U}([0, 1])$ (∞ X v.a. bornée) donc pour avoir $n \rightarrow +\infty$, il faut multiplier par n .

1 borne supérieure sur la loi \Leftrightarrow 1 borne supérieure sur des extrêmes.



si on est dans le MDA (Fréchet), il n'y a aucune raison que les maxima (ie record) décroissent,

th: MDA (Φ_x) = RV(α), $\forall \alpha \geq 0$

ex: Pareto, Burr, G_x ($0 < \alpha < 2$), loggamma
 $(f(x) = \frac{x^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{\alpha-1})$,

exo 15: $\Phi_x \in$ MDA (Φ_x) = RV(α).

prop: si $F \in$ MDA (Φ_x) alors :

les records ont un ordre de grandeur qui explose avec le temps

$$b_m^{-1} M_n \xrightarrow[n \rightarrow +\infty]{} \Phi_x$$

ou $b_m = F^{-1}\left(1 - \frac{1}{m}\right)$.

exo 16: • Montrer qu'il existe $K > 0$ tel que $b_m = (Km)^{\frac{1}{\alpha}}$ lorsque F est une Pareto.

• Montrer que $\exists K > 0$ tel que $b_m = (K[\ln(m)]^{\beta-1} m)^{\frac{1}{\alpha}}$ lorsque F est une loggamma.

\Rightarrow ordre de grandeur des maxima ($\approx m^{\frac{1}{\alpha}}$) augmente (\Rightarrow les records augmentent)

• Vérifier que $\frac{M_n}{F^{-1}(1 - \frac{1}{n})} \xrightarrow[n \rightarrow +\infty]{\varphi} \Phi_x$

pour $F = \Phi_x$ et $F = G_{\pm}$. ($G_{\pm} \in \text{MDA}(\Phi_{\pm})$ ce qui est logique car Cauchy = 1-stable)

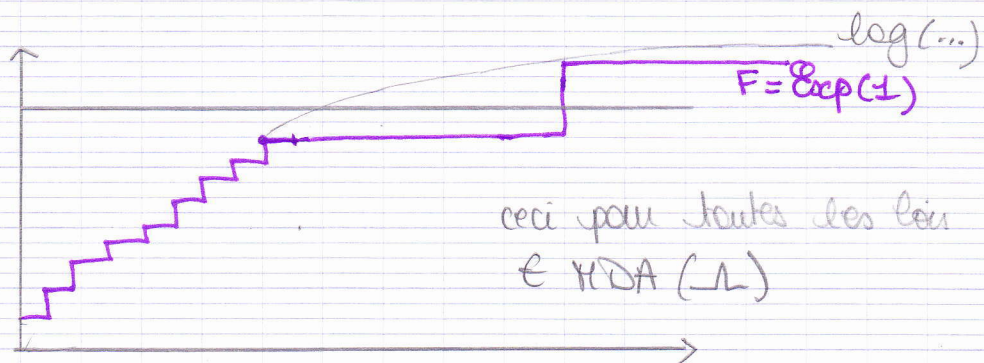
On ne peut pas caractériser $\text{MDA}(\Lambda)$. Cependant:

- si $F \in \text{MDA}(\Phi_x)$ alors $E[|X|^{\delta}] < +\infty, \forall \delta < x$ et $E[|X|^x] = +\infty$.
- si $F \in \text{MDA}(\Lambda)$ alors $E[|X|^{\delta}] < +\infty \forall \delta$
- (imp) • si $F \in \text{MDA}(\Phi_x)$ alors $Y = \ln^{(\max)}(1 \vee X) \in \text{MDA}(\Lambda)$ où $X \sim F$ est telle que $Y \sim G$.

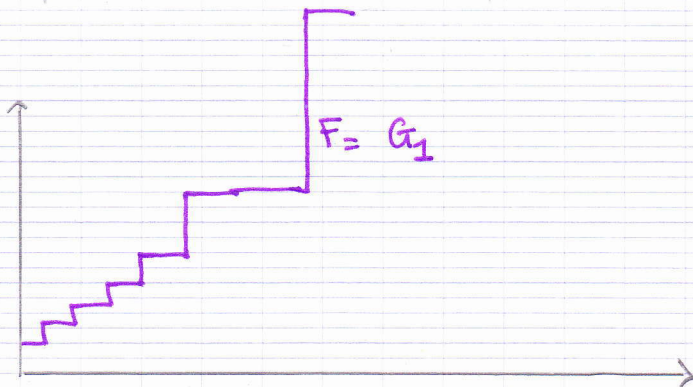
$$\begin{aligned} \ln(1 \vee M_n) &= \text{umax}(\ln(1 \vee X_n)) \\ &= \text{umax } Y_i \quad (\text{continue}) \end{aligned}$$

$$\rightsquigarrow \text{umax } Y_i = \ln(1 \vee b_n) \xrightarrow{\varphi} \Lambda$$

Les records sont du même ordre de grandeur lorsque $F \in \text{MDA}(\Lambda)$.



le prochain record sera du même ordre de grandeur des records passés. (record $\rightarrow +\infty$ très lentement)



les événements extrêmes dirigent la globalité des observations et l'ordre de grandeur des futurs événements extrêmes est différent.

• La loi log-normale et la loi Weibull (MDKA).

On a ici: $b_m^{-1} (M_m - a_m) \xrightarrow{m \rightarrow +\infty} \Lambda$
avec $b_m \sim (\log(m))^\beta$ pour $\beta > 0$.

Plus α indice est grand ($\alpha > 2$), plus la probabilité d'occurrence de $\{M_m \geq \sqrt{m} x_m\}$ est faible ($x_m \rightarrow +\infty$).

niveau fixé dans le TLC

$R_0: \sqrt{m} = \text{constante}$ de renormalisation $\alpha > 2, x_m \nearrow +\infty, \frac{\mathbb{P}(S_m \geq \sqrt{m} x_m)}{\mathbb{P}(M_m \geq \sqrt{m} x_m)} \xrightarrow{m \rightarrow +\infty} 1$ (*)

Or $\sqrt{m} \gg m^{\frac{1}{\alpha}}$ donc $\mathbb{P}(M_m \geq \sqrt{m} x_m) \rightarrow 0$

$\mathbb{P}(M_m \geq \sqrt{m} x_m)$ d'autant plus vite que α est grand!
 $= \mathbb{P}(M_m \geq m^{\frac{1}{\alpha}} \frac{\sqrt{m} x_m}{m^{\frac{1}{\alpha}}})$ $\{S_m \geq \sqrt{m} x_m\}$ et $\{M_m \geq \sqrt{m} x_m\}$ sont des événements rares de même importance.

$\frac{\sqrt{m}}{m^{\frac{1}{\alpha}}} \xrightarrow{m \rightarrow +\infty} +\infty$ $\alpha > 2$ est un indice de rareté de ces événements d'autant plus vite que α est grand.

(*) ici, on s'intéresse aux événements rares qui caractérisent les événements globaux.

Dans ce cadre ($\alpha > 2$), α est un indice de la rareté des événements extrêmes de type $\{M_m \geq \sqrt{m} x_m\}$ événements qui caractérisent les extrêmes de S_m . ($\alpha < 2$ joue sur l'importance des événements rares)

IV Inférence statistique.

Soit (X_1, \dots, X_m) une suite d'observations iid. Que peut-on dire sur les événements

raies de type extrême?

On se restreint au problème suivant: On se fixe $p \in]0, 1[$, on veut estimer la VaR x_p pour une marginale telle que:

$$P(X > x_p) = p \quad (x_p = F^{-1}(1-p)).$$

A-GEV et GPD.

def: On représente les 3 types de lois univariées stables sous une forme unique appelée Generalized Extreme Value distribution:

$$H_\xi(x) = \begin{cases} 1 - \exp(- (1 + \xi x)^{1/\xi}) & \text{si } \xi \neq 0 \\ \exp(- \exp(-x)) & \text{si } \xi = 0 \end{cases}$$

pour $1 + \xi x \geq 0, \xi \in \mathbb{R}$.
à x fixé, $(1 + \xi x)^{1/\xi} \xrightarrow{\xi \rightarrow 0} x$

- Rq:
- si $\xi = 0$, H_0 est du même type que Λ
 - si $\xi > 0$, H_ξ est du même type que Φ_x avec $x = \frac{1}{\xi}$. (intéressant pour la théorie des valeurs extrêmes)
 - si $\xi < 0$, H_ξ est du même type que Ψ_x avec $x = -\frac{1}{\xi}$. (on est dans le cas bornée)

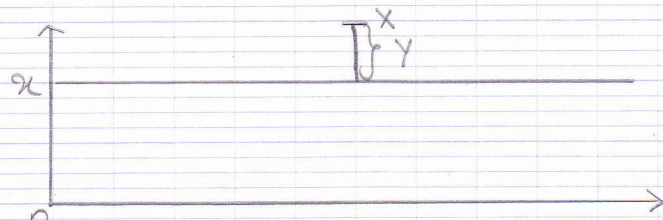
Dans la suite, $F \in \text{MDA}(H_\xi)$ pour $\xi \in \mathbb{R}$ (Poisson(p) n'appartient à aucun $\text{MDA}(H_\xi)$) et on souhaite estimer ξ inconnu. On n'est pas ici dans un cadre paramétrique.

def: La fonction de répartition d'une Pareto Généralisée (GPD) est définie par:

fonction de répartition $1 - G_{\xi, \tau} = \bar{G}_{\xi, \tau}(x) = \frac{1}{(1 - \tau x)^{1/\xi}}, \xi \in \mathbb{R}, \frac{\xi}{\tau} < 0$
 $\forall x \geq 0$ si $\xi \geq 0$ et $0 \leq x \leq \frac{1}{\tau}$ si $\xi < 0$.

Rq: GEV et GPD sont étroitement liées.

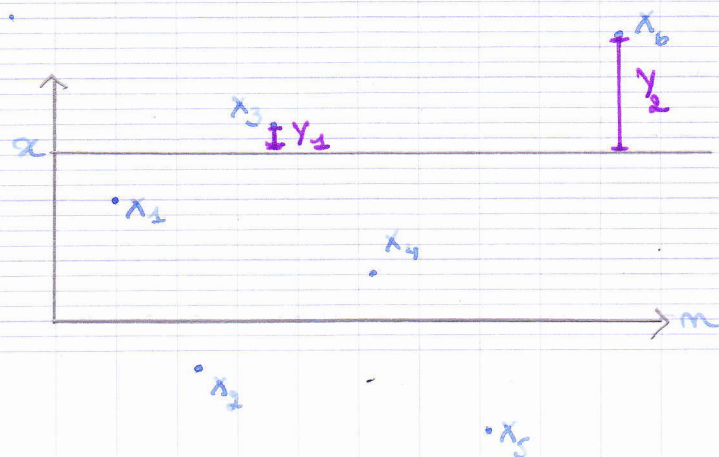
th: On note $x_F = F^{-1}(1) =$ borne supérieure du support $\in \mathbb{R} \cup \{+\infty\}$. Alors, $\forall \xi \in \mathbb{R}, F \in \text{MDA}(H_\xi) \Leftrightarrow \lim_{x \rightarrow x_F} \sup_{0 < y < x_F - x} \frac{F_x(y) - G_{\xi, \tau}(x, y)}{1 - G_{\xi, \tau}(x, y)} = 1 = 0$
 où $F_x(y) = \mathbb{P}(X - x \leq y \mid X > x)$ est appelée fonction de répartition des excès.



On fixe le seuil x . On regarde les valeurs au-dessus de ce seuil, i.e. les excès et on fait de l'inférence statistique.

B - Méthode POT (Peaks Over Threshold, "Pic au-dessus du seuil"),

Soient $X_1, \dots, X_n \sim F \in \text{MDA}(H_\xi)$ et on fixe un niveau x .



excès = de combien on dépasse = $X_i - x$

On note les excès $Y_1, Y_2, \dots, Y_{N(x)}$ où
 $N(x) = \# \{i, i=1, \dots, m, X_i > x\}$.

plus le seuil x est grand, moins d'excès on aura.

• Le processus des temps d'excès: temps de rebou T_j
 $[0, m] \rightarrow [0, \frac{1}{2}]$
 $T_j = \min \{ \frac{j}{m} > T_{j-1}, X_j > x \}$

renormalisé

$\frac{j}{m}$ = temps de rebou

(Le fait que $(X_i)_i$ sont iid est donné par $(T_j)_j$ Poisson (si les X_i sont iid), ($m \rightarrow +\infty$),

$(T_j - T_{j-1})_j$ sont iid avec $T_j - T_{j-1} \sim \text{Exp}(\lambda)$.

où $\lambda > 0$ ne dépend pas de j (homogène)

Le temps de rebou moyen ne varie pas dans le temps

• Les (Y_i) sont iid et suivent une loi $G_{\xi, \tau}(x)$. On applique la méthode du maximum de vraisemblance (exo 17):

$$\begin{cases} \hat{\xi} = \frac{1}{N(x)} \sum_{i=1}^{N(x)} \ln(1 - \hat{\tau} Y_i) \\ \frac{1}{\hat{\tau}} + \frac{1}{N(x)} \left(\frac{1}{\hat{\xi}} + 1 \right) \sum_{i=1}^{N(x)} \frac{Y_i}{1 - \hat{\tau} Y_i} = 0 \quad (\dagger) \end{cases}$$

(pas de forme explicite des paramètres)

sous R, pour résoudre (\dagger) , il faut télécharger le package EVIR

⚠ Le support dépend de τ . (on n'a pas la normalité asymptotique alors mais on a bien $\hat{\xi} \rightarrow \xi$ et $\hat{\tau} \rightarrow \tau$). \rightarrow on ne pourra pas faire les tests classiques, il existe aussi une méthode type "moment", Bayésienne.

⚠ On est retombé dans un cadre paramétrique.

On a alors des approximations:

$$\rightarrow \mathbb{P}(X > x+y) = \frac{N(x)}{n} \left(1 - \frac{1}{\hat{\xi}}(y)\right)^{-\frac{1}{\hat{\xi}}}$$

se comporte comme une queue de distribution d'une Pareto généralisée

$$\rightarrow \hat{x}_\alpha = x - \hat{\xi} \left(\left(\frac{n}{N(x)} \alpha \right)^{\hat{\xi}} - 1 \right) \text{ estimateur de la VaR,}$$

Rq: si $\hat{\xi} < 0$ alors $\hat{\xi} > 0$ donc $\hat{x}_\alpha < x = \text{seuil}$ qu'on avait fixé \Rightarrow on est dans le cadre des extrêmes tronqués

C- Validation de la méthode POT.

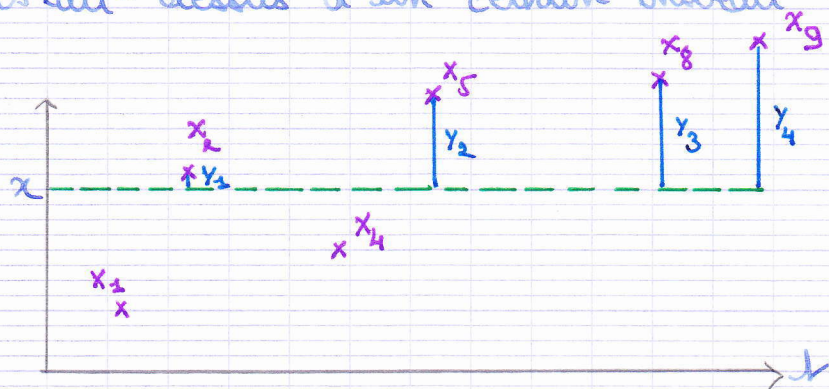
• La méthode ne fonctionne que si $\xi > -\frac{1}{2}$.

• Choix de x ?

\rightarrow Il faut que $1 \ll N(x) \ll n \dots$

par la méthode des valeurs extrêmes, il suffit de garder les valeurs extrêmes. Le nombre dépend de x .

Rappel: La méthode POT consiste à inférer des événements de type extrême $\{X > x\}$ on ne considère que les excès au-dessus d'un certain niveau x qu'on choisit



si μ est un peu alors il n'est pas nécessaire de faire appel aux théorèmes des valeurs extrêmes

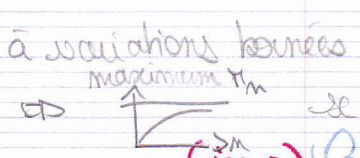
$\alpha =$ seuil

Soient $(Y_1, \dots, Y_{n(\alpha)})$ les excès, alors si (X_1, \dots, X_m) sont iid les lois \in MDA(H_ξ), on a:

$\rightarrow (Y_1, \dots, Y_{n(\alpha)})$ sont iid de loi de Pareto généralisée: $G_{\xi, \bar{\sigma}}(y) = \frac{1}{(1 - \bar{\sigma} y)^\xi}$, $\xi \in \mathbb{R}, \bar{\sigma} < 0$
si $\xi < 0, \forall y \in [0, \xi/\bar{\sigma}]$

Rq: Quand $\xi \rightarrow 0$ alors $\bar{\sigma} \rightarrow 0$ et $\lim_{\xi \rightarrow 0} \frac{\bar{\sigma}}{\xi} < 0$

donc $G_{\xi, \bar{\sigma}}(y) = \exp(-\frac{1}{\xi} \ln(1 - \bar{\sigma} y)) \xrightarrow{\xi \rightarrow 0} \exp(\frac{\bar{\sigma}}{\xi} y)$



à variations bornées • si $\xi < 0$ alors $G_{\xi, \bar{\sigma}}(x) \xrightarrow{x \rightarrow +\infty} +\infty$ donc il faut se restreindre à un intervalle.

(imp) Ceci n'est vrai que pour un niveau α bien choisi!

donc il suffit de regarder q_d n grand donc on restreint d'intervalle

Rq: Si α est trop petit, $(Y_1, \dots, Y_{n(\alpha)}) = (X_1, \dots, X_m)$ pour $X > 0$ (loi des excès = loi de départ) (ex: $\alpha = 0$).
 $(X \sim \text{Gamma}(1, 2))$. La loi Gamma ne peut pas être représentée comme une Pareto généralisée car $X > 0$. donc $(Y_1, \dots, Y_{n(\alpha)}) \sim \text{Gamma}(1, 2)$ qui n'est pas une Pareto généralisée $\forall \xi \in \mathbb{R}, \bar{\sigma} < 0$.

On dit qu'on est dans un cadre de statistique semi-paramétrique (après transformation, on se ramène au cadre paramétrique, c'est-à-dire la Pareto généralisée si ξ et $\bar{\sigma}$ sont à estimer): les observations suivent une loi inconnue ("cadre non-paramétrique") mais la théorie nous permet de se ramener au cadre paramétrique par la transformation $(X_1, \dots, X_m) \rightarrow (Y_1, \dots, Y_{n(\alpha)})$ pour α bien choisi.

\rightarrow Les temps d'excès (normalisés):

$$T_j = \inf \left\{ \frac{i}{n} > T_{j-1}, i=1, \dots, n, X_i > \alpha \right\}$$

$$(T_0 = 0).$$

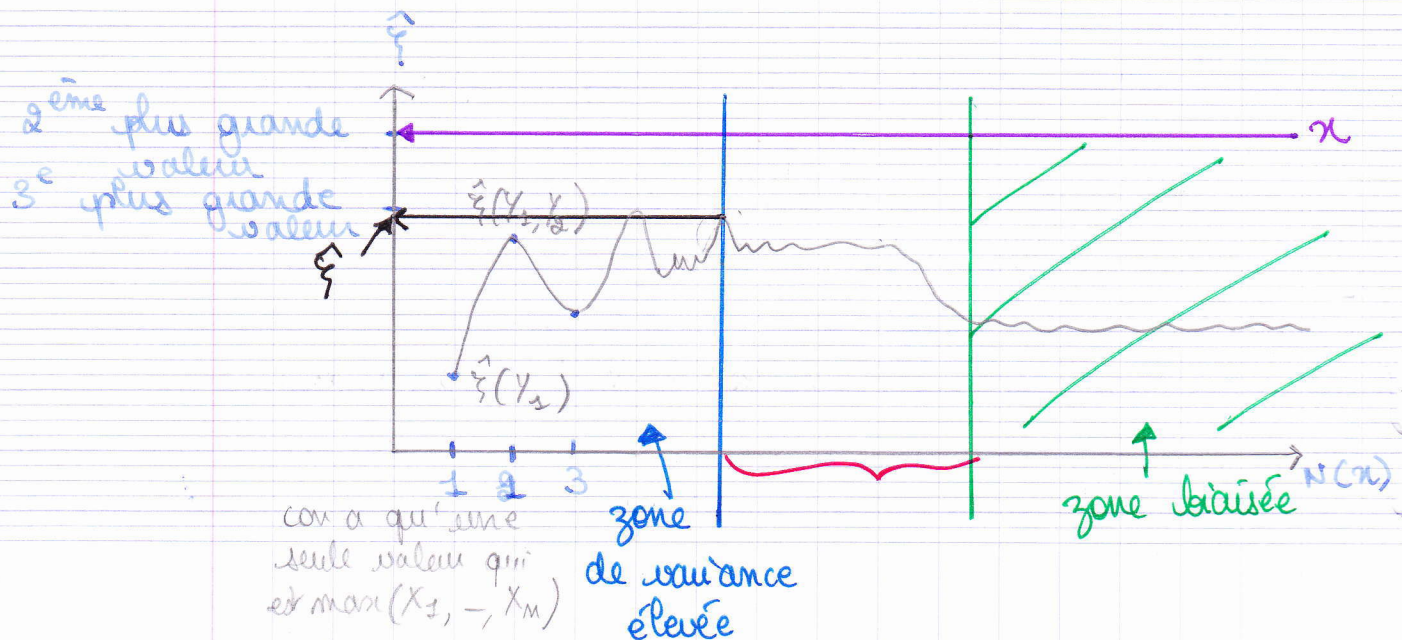
exc: Sur le graphique, $n=9$, on a: $T_1 = \frac{2}{9}$, $T_2 = \frac{5}{9}$, $T_3 = \frac{8}{9}$ et $T_4 = 1$.
On définit le processus des temps d'excès $(T_j)_{j \in \mathbb{N}}$ qui est, approximativement quand $n \rightarrow +\infty$, tel que $(T_j - T_{j-1})$ sont iid $\sim \text{Exp}(\lambda)$, $\lambda > 0$

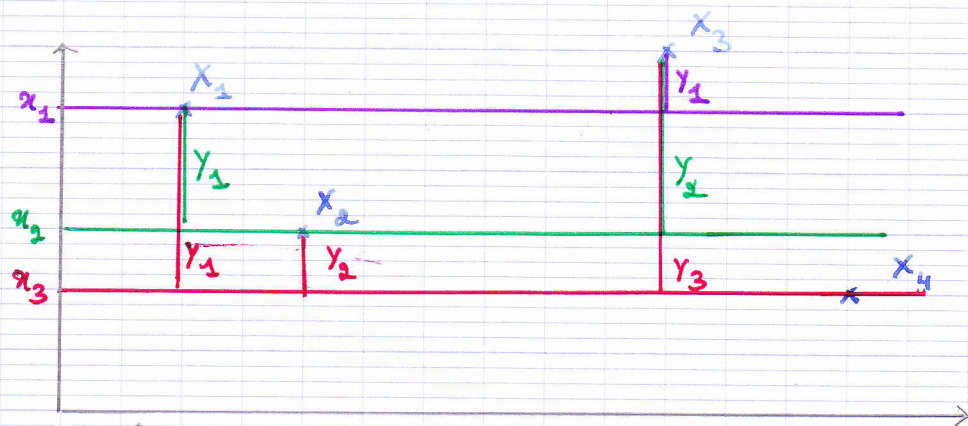
La méthode POT repose sur de nombreuses hypothèses (seuil α bien choisi, FEMDA (H_2), iid).
en g^{al}, toutes les lois classiques \in MDA (\Rightarrow jamais remis en cause)
 \uparrow est principalement le prob

prob: Il n'existe pas de méthodes de tests statistiques rigoureux dans ce cadre semi-paramétrique.

1 - Choix du seuil. (méthode non rigoureuse)

En pratique, on trace $\hat{\xi}(\alpha)$ où $\hat{\xi}$ est le maximum de vraisemblance obtenu à partir de $(Y_1, \dots, Y_{[n\alpha]})$ pour différents $\alpha > 0$.





A chaque fois qu'on fixe un new seuil, on recalcule le maximum de vraisemblance.

A chaque fois, on abaisse le seuil α_m en prenant la valeur suivante de la $(m+1)^{\text{ième}}$ plus grande valeur.

Au bout d'un moment, on a une certaine stabilité.

En abaissant le seuil et qu'une à un certain moment, baisser le seuil ne changera rien (on a gardé un certain nombre de points).

Plus $N(x)$ grand, plus on a une meilleure estimation.

zone de variance élevée \Rightarrow seuil α trop élevé, on ne garde pas.

zone biaisée \Rightarrow on garde $N(x)$ suffisant de valeurs mais on en garde trop donc les excès ne suivent plus une loi généralisée.

à la limite, on sait qu'on garde toutes les observations $(X_i)_{i \in \mathcal{M}}$ donc dans cette zone qui varie peu, on sait que ce n'est pas bon. \Rightarrow pas intéressant.

Dans la zone des seuils α restreints, on choisit le seuil α le plus grand; c'est un compromis biais-variance. ($\hat{\xi}$ est plus imp que $\hat{\sigma}$).

2 - modélisation des queues de distribution.

Le choix de x (de $N(x)$) fournit des estimateurs $(\hat{\xi}, \hat{\zeta})$. on vérifie que x est compatible avec l'approche non-paramétrique.

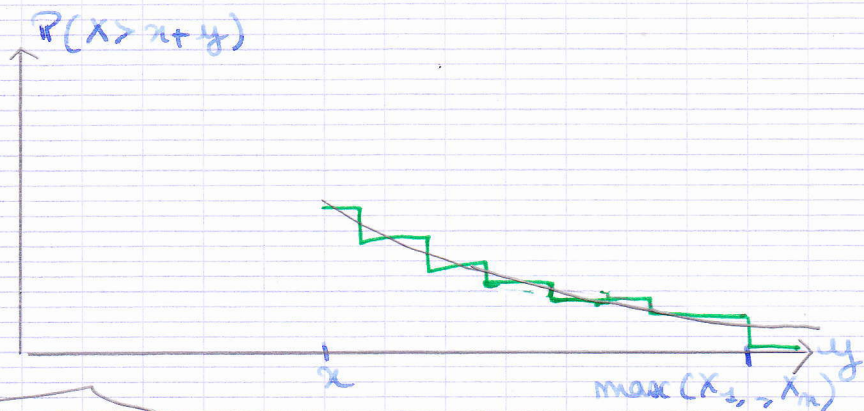
$\hat{\xi}$ et $\hat{\zeta}$ permettent de calculer une VaR.

On compare cette approche semi-paramétrique avec l'approche non-paramétrique.

cadre semi-paramétrique: (x fixe pour la méthode graphique):

$$\forall y > 0, \quad \mathbb{P}(X > x+y) = \frac{N(x)}{n} \left[\frac{1}{(1 - \hat{\zeta} y)^{\frac{1}{\hat{\xi}}}} \right]$$

On s'est fixé le seul x qui est censé être une bonne estimation.



$$\mathbb{P}(X > x+y | X > x) = \text{Pareto généralisée.}$$

cadre non-paramétrique: (queue de distribution empirique):

$$\mathbb{P}(X > x+y) = \frac{\#\{i, i \in \{1, \dots, n\}, \text{ avec } X_i > x+y\}}{n}$$

↳ on ne peut pas calculer la VaR avec cette méthode car VaR dépend de ξ et ζ .

↳ graphique vert: saut de taille $\frac{1}{n}$.

Il faut que les 2 graphiques soient proches (imp).

$\max(X_1, \dots, X_n) = \text{Var}$ de risque $\frac{1}{n}$ (empiriquement)

Rq: L'approche non-paramétrique ne permet pas d'estimer des Var car: $\widehat{\text{Var}}_\alpha = F_n^{-1}(1-\alpha) = \max(X_1, \dots, X_n)$
 $\forall \alpha \leq \frac{1}{n}$

(où $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq x\}}$)

\Rightarrow queue de distribution $= X_{(n)}$ (approche non-paramétrique) alors que de ce cadre semi-paramétrique, la Var peut être calculée $\forall \alpha$ car on a une queue de distribution continue.

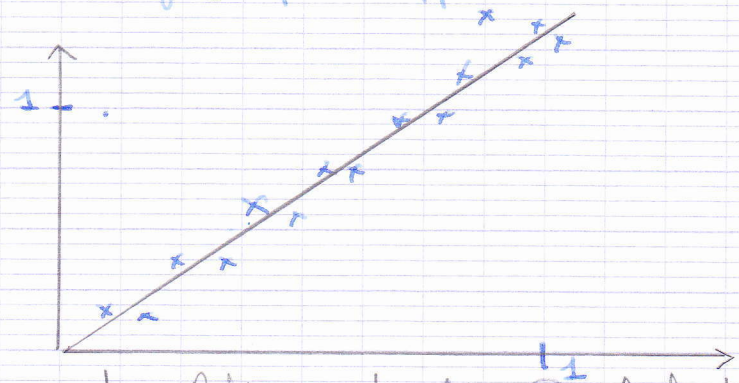
Dans l'approche semi-paramétrique, on peut estimer la Var quelque soit $\alpha \in]0, 1[$.

non paramétrique: $\widehat{\text{Var}}_\alpha = \text{constante} \Rightarrow$ on sous-estime le risque

semi-paramétrique: il peut y avoir de plus grande valeurs que $X_{(n)}$.

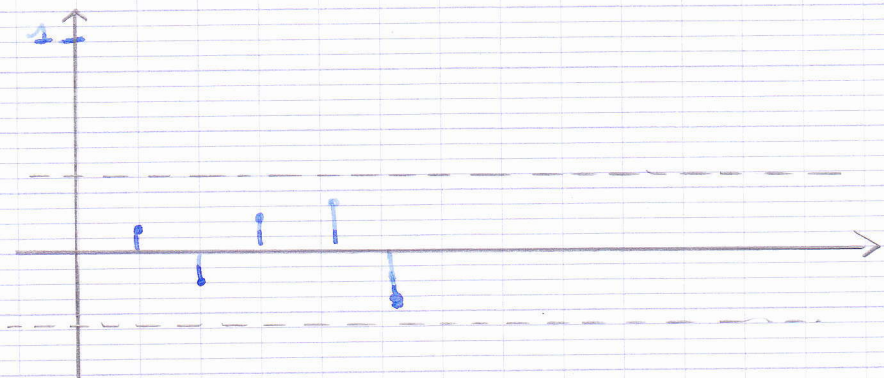
3 - Discussion de l'hypothèse i.i.d.

On trace: \rightarrow QQ-plot des incréments des dates d'excès $T_j - T_{j-1}$ par rapport à la loi $\text{Exp}(1)$.



les points sont relativement alignés, il faut que ce soit linéaire car pente = λ

→ l'autocorrélogramme de $(T_j - T_{j-1})$



points correspondent à l'hypothèse H_0 : il n'y a pas de corrélation.

- ① Vérifier que les points sont proches d'une droite affine $y = \lambda x$ avec $0 < \lambda < +\infty$ (QQ-plot)
- ② Vérifier qu'on reste dans la bande de confiance de H_0 : non-corrélation (autocorrélogramme)

$$\text{corrélation} = \frac{\frac{1}{n-h} \sum_{i=1}^{n-h} (T_i - T_{i-1})(T_{i+h} - T_{i+h-1}) - \left[\frac{1}{n} \sum_{i=1}^n (T_i - T_{i-1}) \right]^2}{\frac{1}{n} \sum_{i=1}^n (T_i - T_{i-1} - \frac{1}{n} \sum_{i=1}^n (T_i - T_{i-1}))^2}$$

indépendamment distribuées \Rightarrow impossible à identifier

- ① tracer les données
- ② identifier un emp^t aberrant: si \pm justificat^o ex-ternéure aux données (élément perturbateur) \Rightarrow on enlève des données.

Rq: S_m est croissante en n donc il est inutile de fixer le seuil à une constante d'où $P(S_m \geq \alpha_m)$

Chapitre III :

Théorie des valeurs extrêmes bivariées à l'aide des copules.

ref. → Gestion des risques multiples, cours ENSAI, de Roncalli.
→ Introduction to Copulas, Deaton.

I Les Copules.

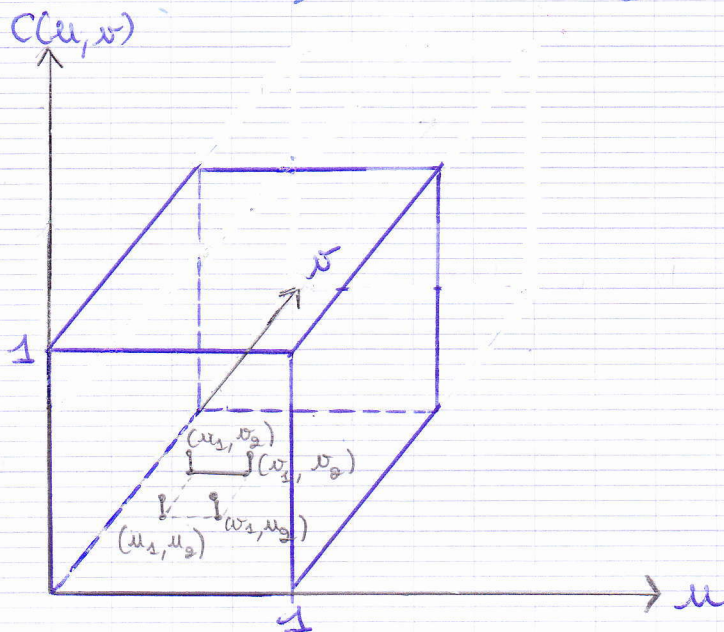
A - Théorème de Sklar (1959),

def. Une copule C est une fonction de $[0, 1]^2$ dans $[0, 1]$ telle que :

- $C(0, u) = C(u, 0) = 0, \quad \forall 0 \leq u \leq 1$
- $C(u, 1) = C(1, u) = u, \quad \forall 0 \leq u \leq 1.$
- C est 2-croissante, c'est-à-dire :

$$C(w_1, w_2) - C(w_1, u_2) - C(u_1, w_2) + C(u_1, u_2) \geq 0$$

$$\forall 0 \leq u_1 \leq w_1 \leq 1, \quad \forall 0 \leq u_2 \leq w_2 \leq 1.$$



La copule

prop: C est la distribution d'un couple (U_1, U_2) de lois marginales uniformes $U_1 \sim \mathcal{U}([0, 1])$ et $U_2 \sim \mathcal{U}([0, 1])$.

En effet, on a: $C(u, v) = \mathbb{P}(U_1 \leq u, U_2 \leq v) \quad \forall u, v \in [0, 1]$

- * $\mathbb{P}(U_1 \leq 0, U_2 \leq u) = \mathbb{P}(U_1 \leq u, U_2 \leq 0) = 0$ ($= \mathbb{P}(U_1 \leq 0)$)
- * $\mathbb{P}(U_1 \leq u, U_2 \leq 1) = \mathbb{P}(U_1 \leq 1, U_2 \leq u) = u$ ($= \mathbb{P}(U_1 \leq u)$)
- * $\mathbb{P}(u_1 \leq U_1 \leq w_1, u_2 \leq U_2 \leq w_2)$
 $= \mathbb{P}(U_1 \leq w_1, U_2 \leq w_2) - \mathbb{P}(U_1 \leq w_1, U_2 \leq u_2)$
 $- \mathbb{P}(U_1 \leq u_1, U_2 \leq w_2) + \mathbb{P}(U_1 \leq u_1, U_2 \leq u_2)$
 ≥ 0 (grâce à la 2-coissance).

exco 1: • Montre que $C^+(u, v) = uv$ est une copule (appelée copule produit). ($\Leftrightarrow U_1 \perp U_2$)

• Soient $X_1 \sim F_1$ et $X_2 \sim F_2$. Alors

$F_1(X_1) \sim \mathcal{U}([0, 1])$ et $F_2(X_2) \sim \mathcal{U}([0, 1])$ donc $C(F_1(x_1), F_2(x_2))$ définit une distribution de probabilité bidimensionnelle.

D'autre part, $F(x_1, x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2)$ est aussi une distribution de probabilité bidimensionnelle.

th: (Sklar): Soit $(X_1, X_2) \sim F$ telle que $X_1 \sim F_1$ et $X_2 \sim F_2$ alors il existe une copule C telle que:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (\text{pas unique})$$

De plus, la copule $C = (X_1, X_2)$ est unique dès que F_1 et F_2 sont continues.

notations: La copule du couple (X_1, X_2) vérifie:

$$C(X_1, X_2)(u, w) = F(F_1^{-1}(u), F_2^{-1}(w)), \forall (u, w) \in [0, 1]^2$$

def: La copule normale est définie par:

$$C(u, w; \rho) = \Phi(\Phi^{-1}(u), \Phi^{-1}(w))$$

- où: $\rightarrow \Phi$ la fonction de répartition de $\mathcal{N}(0, 1)$.
- $\rightarrow \Phi$ est la distribution de $\mathcal{N}_2(0_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$

avec $\rho =$ coefficient de corrélation, $-1 \leq \rho \leq 1$.

$C(u, w; \rho)$ est, à ρ -près, unique.

$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ est définie positive d'où $-1 \leq \rho \leq 1$
 ($\Leftrightarrow \Delta = 1 - \rho^2 \geq 0 \Leftrightarrow \rho^2 \leq 1$)
 = matrice de variance-covariance

Rappel: $\rho = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}$ et d'après l'inégalité

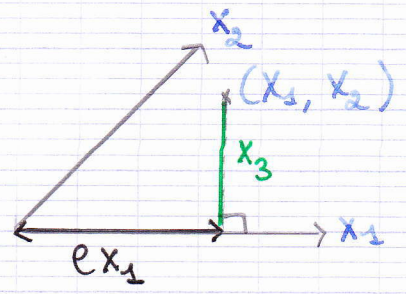
de Cauchy-Schwarz, on a bien $|\rho| \leq 1$

th: La copule normale vaut:

$$C(u, w; \rho) = \int_0^u \Phi\left(\frac{\Phi^{-1}(w) - \rho \Phi^{-1}(u)}{\sqrt{1 - \rho^2}}\right) du$$

preuve: Soient $(X_1, X_2) \sim \mathcal{N}_2(0_2, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})$, alors

$E[X_2 | X_1] = \rho X_1$ (spécifique au cas gaussien). De plus, théorème de Cochran, nous assure qu'il y a une projection orthogonale dans L^2 .



existe $X_3 \perp X_1$ tel que $X_2 = \rho X_1 + \sqrt{1 - \rho^2} X_3$ avec $X_3 \sim \mathcal{N}(0, 1)$.

B - Densité bidimensionnelle.

Si F est ^{distribution du couple} absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^2 ($\lambda \times \lambda$), elle admet une densité qui vérifie: $f(x_1, x_2) = c(F_1(x_1), F_2(x_2)) f_1(x_1) f_2(x_2)$ où c est la densité de la copule, c'est-à-dire:

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}$$

prop.: montrer que $c \geq 0 \Leftrightarrow C$ est α -croissante.

Rq.: La copule produit $C^+(u, v) = uv$ admet une densité $c^+ = 1 \geq 0$. On en déduit que $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, C^+ est la loi de (U_1, U_2) avec $U_1 \sim \mathcal{U}([0, 1])$, $U_2 \sim \mathcal{U}([0, 1])$ et $U_1 \perp U_2$.

C - Représentation graphique d'une copule.

def.: Une courbe de niveau $\alpha \in [0, 1]$ de la copule C est définie par:

$$L(C)(\alpha) = \{(u, v) \in [0, 1]^2, C(u, v) = \alpha\}$$

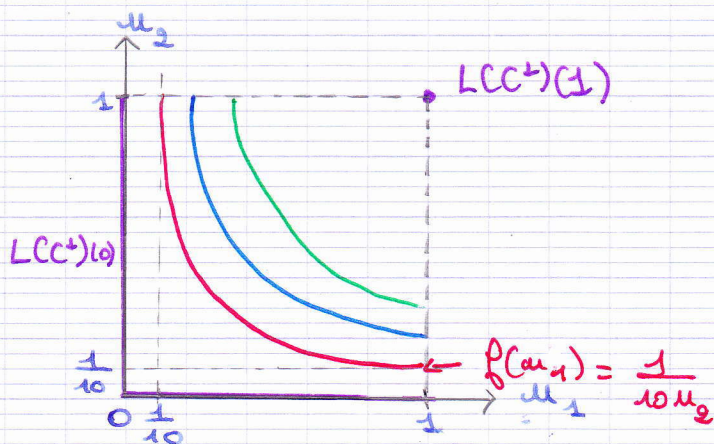
Pour représenter une copule (fonction de $[0, 1]^2 \rightarrow [0, 1]$), on trace différentes lignes de niveau, par exemple $L(C)(\alpha)$ pour $\alpha = 0, \frac{4}{10}, \dots, \frac{9}{10}, 1$.

Représentation graphique de la copule produit.

$$L(C^+)(0) = [0, 1] \times \{0\} \cup \{0\} \times [0, 1]$$

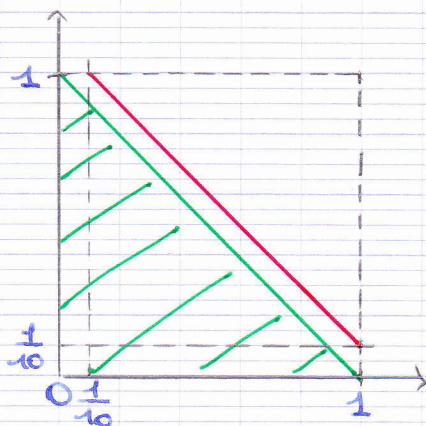
$$L(C^+)\left(\frac{1}{10}\right) = \left\{ (u_1, u_2), u_2 = \frac{1}{10u_1} \right\}$$

... $L(C^+)(1) = \{1\} \times \{1\}$



def: Soient C_1 et C_2 2 copules. On dit que C_1 est plus petite que C_2 ($C_1 < C_2$ ordre partiel) ssi:
 $C_1(u_1, u_2) \leq C_2(u_1, u_2)$, $\forall (u_1, u_2) \in [0, 1]^2$

th: (Bornes de Fréchet) Toutes copules C vérifient
 $C^- < C < C^+$ où: $\rightarrow C^-(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$
 copule négative.
 $\rightarrow C^+(u_1, u_2) = \min(u_1, u_2)$
 copule positive



$L(C^-)(0)$
 $= \{u_2 \leq 1 - u_1, (u_1, u_2) \in [0, 1]^2\}$

$L(C^-)(1/10)$
 $= \{(u_1, u_2) \in [0, 1]^2, u_2 = \frac{11}{10} - u_1\}$

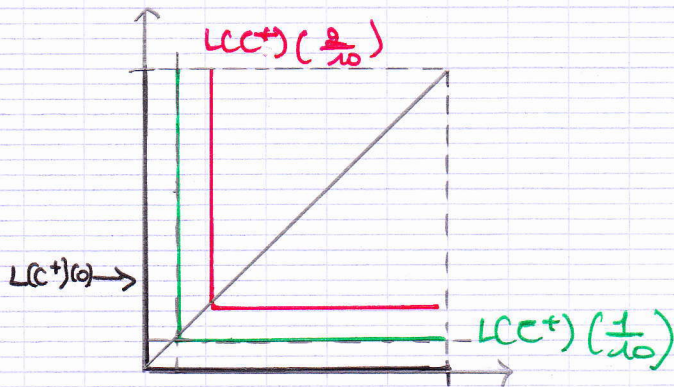
- $L(C^-)(0)$
- $L(C^-)(1/10)$

$L(C^-)(\alpha)$ pour $\alpha = \frac{2}{10}, \frac{9}{10}, 1$ sont des droites.

Δ on connaît toutes les extrémités des courbes de niveaux

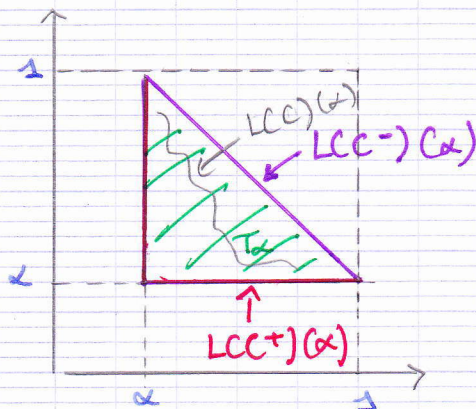
$$L(C^+)(0) = [0, 1] \times \{0\} \cup \{0\} \times [0, 1]$$

$$L(C^+)(\alpha) = [\alpha, 1] \times \{\alpha\} \cup \{\alpha\} \times [\alpha, 1]$$



Rq: Comme $C^- \prec C \prec C^+$, nous avons pour tout $\alpha \in [0, 1]$, $L(C)(\alpha)$ appartient au triangle T_α .

\Rightarrow Les courbes de niveau ne se croisent pas



les points sur les côtés sont fixés

Rq: Une copule est plus grande lorsque ces courbes de niveau sont plus proches de 0 (pour cela, il faut regarder par rapport à la 1^{ère} diagonale pour atteindre le niveau).

exo 8: Soit la copule de Frank de paramètre $\theta \in \mathbb{R}$:

$$C^F(u_1, u_2; \theta) = -\frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right) - 1$$

1] Vérifier que c'est une copule (densité $c \geq 0$)

2] Etudier les cas limites $\theta \rightarrow 0$, $\theta \rightarrow +\infty$, $\theta \rightarrow -\infty$.

3] Représenter graphiquement C^F pour différentes valeurs de θ . Montrer que :

$$C^F(\theta_1) < C^F(\theta_2) \text{ pour } \theta_1 \leq \theta_2$$

Rq: Plus θ est grand, plus la copule est grande.
 \Rightarrow au bout d'un moment, $L(C^F)$ va se confondre avec $L(C^{F-})$ (ie pour certaines valeurs de θ) et de même pour $L(C^{F+})$.

- prop: Soit $(X_1, X_2) \sim \mathcal{N}_2(0, I_2)$ alors: ($e \rightarrow 1$ ou $e \rightarrow -1$)
- $\rightarrow C^-$ correspond à $C(X_1, -X_1)$.
 - $\rightarrow C^+$ correspond à $C(X_1, X_1)$.
 - $\rightarrow C^+$ correspond à $C(X_1, X_2)$. (*)

prop 4: Montrer que si X_1, X_2 sont iid \sim log-normale alors: $\rightarrow C^- = C(X_1, \frac{1}{X_1})$

$$\rightarrow C^+ = C(X_1, X_1)$$

- Montrer que si U_1, U_2 sont iid $\sim U([0, 1])$ alors: $\rightarrow C^- = C(U_1, 1 - U_1)$.
- $\rightarrow C^+ = C(U_1, U_1)$.

Rq: • copule positive: une marginale est égale à l'autre marginale (\Rightarrow dépendance positive)

• copule négative \Rightarrow dépendance négative

$$\Rightarrow \begin{cases} X_2 = -X_1 & \text{pour la normale} \\ X_2 = -\frac{1}{X_1} & \text{pour la lognormale} \\ X_2 = 1 - X_1 & \text{pour l'uniforme} \end{cases}$$

th: • Si (X_1, X_2) est tel que $C(X_1, X_2) = C^-$ alors f_1 croissante, f_2 décroissante et une v.a. X telles

que: $X_1 = f_1(X)$ et $X_2 = f_2(X)$.

On dit que (X_1, X_2) est contre-monotone.

(quand une marginale croît, l'autre décroît)

• Si $C(X_1, X_2) = C^*$ alors $\exists f_1$ croissante, f_2 croissante et une w.a.r. X telles que:

$$X_1 = f_1(X) \quad \text{et} \quad X_2 = f_2(X)$$

On dit que (X_1, X_2) est co-monotone (quand une marginale croît, l'autre croît).

Rq: Il y a une relation déterministe entre les deux marginales.

prop: Si h_1 croissante et h_2 croissante ^{h_1 et h_2} strictement _↓ alors: $C(h_1(X_1), h_2(X_2)) = C(X_1, X_2)$

Rq: la copule caractérise la dépendance entre les 2 marginales.

On dit que la copule est invariante par transformations strictement croissantes des marginales.

(la transformation ici conserve la relation d'ordre)

Ici, l'ordre de grandeur des marginales n'influence pas la copule (mais influence le coefficient de corrélation).

D - Les coefficients de dépendance de queues λ_u et λ_l .

Rq: On peut trouver h_1 croissante et h_2 croissante telles que $\text{cor}(h_1(X_1), h_2(X_2)) \neq \text{cor}(X_1, X_2)$.
La corrélation est sensible à des transformations sur des marginales.

corrélation = constante $\in [-1, 1]$, copule = fonction

Pour éviter ce phénomène, nous allons construire des coefficients de dépendance à partir des queues.

def: soit $\bar{C}(u, u) = \mathbb{P}(U_1 \geq u, U_2 \geq u)$
 $= 1 - 2u + C(u, u)$ (à vérifier)

est la queue de distribution de la copule.

On appelle coefficient de dépendance des queues de distribution: • à droite λ_u^{upper} défini par:

$$\lim_{u \rightarrow 1^-} \frac{\bar{C}(u, u)}{1-u} = \lambda_u \text{ lorsqu'elle existe.}$$

• à gauche λ_u^{lower} défini par:

$$\lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} = \lambda_L \text{ lorsqu'elle existe.}$$

Rq: • $\lambda_u = \lim_{u \rightarrow 1^-} \mathbb{P}(U_1 > u | U_2 > u)$

• $\lambda_L = \lim_{u \rightarrow 0^+} \mathbb{P}(U_2 \leq u | U_1 \leq u)$
 U_2 a une valeur extrêmement petite

exco 5: • montrer que: $\rightarrow \lambda_u(C^+) = \lambda_L(C^+) = 0$

$\rightarrow \lambda_u(C^+) = \lambda_L(C^+) = 1$

$\rightarrow \lambda_u(C^-) = \lambda_L(C^-) = 0$

c'est l'équivalent de la loi normale pour les lois

(imp) • montrer que: $\lambda_u = 2 - 2^{1/\theta}$ et $\lambda_L = 0$

pour la copule de Gumbel ($\theta \geq 1$) définie par:

$$C(u_1, u_2; \theta) = \exp\left[-\left\{(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\right\}^{1/\theta}\right]$$

• montrer que $\lambda_u = 0$ et $\lambda_L = 2^{-1/\theta}$ pour la

copule de Clayton ($\theta > 0$) définie par:

$$C(u_1, u_2; \theta) = \frac{1}{(u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}}$$

En pratique, on utilise la dépendance quantile -

quantile :

$$\text{Var}_\alpha(X) = F_X^{-1}(1-\alpha)$$

$$\lambda(\alpha) = \mathbb{P}(X_2 > F_2^{-1}(1-\alpha) \mid X_1 > F_1^{-1}(1-\alpha)) \\ = \mathbb{P}(X_2 > \text{Var}_\alpha(X_2) \mid X_1 > \text{Var}_\alpha(X_1)).$$

On remarque que : $\lambda_0 = \lim_{\alpha \rightarrow 0^+} \lambda(\alpha)$: le

coefficient de dépendance des queues à droite correspond à la probabilité d'un événement rare de type extrême ait lieu pour X_2 sachant qu'il a eu lieu pour X_1 .

Rq: Pour la copule normale, $\begin{cases} \lambda_0 = 0 & \text{si } \rho < 1 \\ \lambda_0 = 1 & \text{si } \rho = 1 \end{cases}$
($X_2 = X_1$) cas extrême

\Rightarrow on retombe sur la copule positive

\Rightarrow copule normale non intéressante pour modéliser les queues de distributions des copules.

Contrairement à la copule de Gumbel, la copule normale ne permet pas de modéliser la dépendance entre les événements rares de type extrêmes entre deux v.a. X_1 et X_2 (idem pour la corrélation ρ).

Pour la loi normale, avoir un coefficient de corrélation ρ revient à avoir :

$$X_2 = \rho X_1 + \sqrt{1-\rho^2} X_3$$

où X_3 indépendant (interfère dans les extrêmes). \Rightarrow on ne peut pas faire une régression linéaire mais on peut faire 1 régression quantile.

E. Les copules des valeurs extrêmes

def: Une copule des valeurs extrêmes est une

copule C^* qui vérifie la relation:

$$C^{*t}(u_1, u_2) = C^*(u_1^t, u_2^t), \quad \forall t \geq 0, \quad \forall (u_1, u_2) \in [0, 1]^2$$

Rq: En particulier $C(\max_{1 \leq i \leq n} (X_i), \max_{1 \leq i \leq n} (Y_i))$

$$= C(X_i, Y_i) \text{ pour } (X_i, Y_i) \text{ iid}$$

si $C(X_i, Y_i)$ est une copule des valeurs extrêmes copule stable par maximas.

exo 6: montrer que la copule de Gumbel est une copule des valeurs extrêmes.

th: Soient $(X_i, Y_i)_{i \geq 0}$ une suite iid de vecteurs aléatoires de \mathbb{R}^2 . Si $\exists (a_n, (b_n) > 0, (a'_n)$ et $(b'_n) > 0$ telles que:

$$\begin{aligned} & \text{si } (X_i) \in \text{MDA}(F_1^*) \left(b_n^{-1} (\text{Max}(X_i) - a_n) \right) \xrightarrow[n \rightarrow \infty]{} F^* \\ & \text{et } (Y_i) \in \text{MDA}(F_2^*) \left(b'_n^{-1} (\text{Max}(Y_i) - a'_n) \right) \end{aligned}$$

d'où F_1^* et F_2^*

alors $F^*(x_1, x_2) = C^*(F_1^*(x_1), F_2^*(x_2))$ où F_1^*, F_2^* (max-stables) sont des lois appartenant aux 3 types (Fréchet, Weibull, Gumbel) et C^* est une copule de valeurs extrêmes.

On définit le Max Domaine d'Attraction $\text{MDA}(C^*)$ comme dans le cas univarié.

c'est une caractérisation complète: à la

th: Une copule des valeurs extrêmes C^* vérifie la relation: $C^*(u_1, u_2) = \exp \left[\ln(u_1 u_2) A \left(\frac{\ln(u_1)}{\ln(u_1 u_2)} \right) \right]$

copule C^* , on peut associer une fonction A convexe telle que:

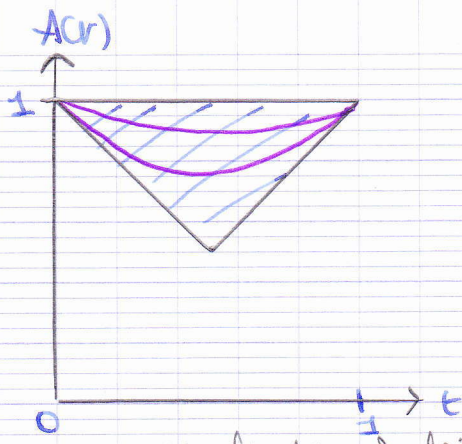
convexe A

$$\rightarrow A(0) = A(1)$$

$$\text{et } \max(t, 1-t) \leq A(t) \leq 1$$

et à la f^m convexe

A , on associe la copule C^*



$A(t)$ appartient au triangle.

copule de 2 lois marginales

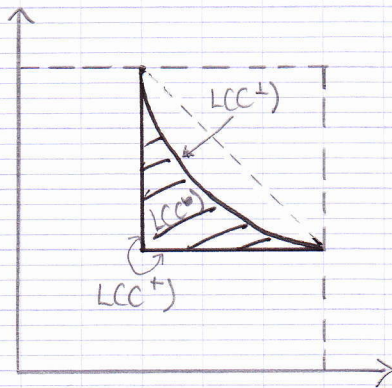
De plus, $F = C(F_1, F_2) \in \text{MDA}(F^* = C^*(F_1^*, F_2^*))$

(iii): ① $F_k \in \text{MDA}(F_k^*) \quad k=1,2.$

② $\lim_{u \rightarrow 0} 1 - C((1-u)^{1-t}, (1-u)^t) = A(t),$

(on regarde ce qui se passe proche de 1) $\forall t \in [0, 1].$

Rq: On a: $C^+ < C^* < C^- \Leftrightarrow \begin{cases} C^+ \Leftrightarrow A=1 \\ C^+ \Leftrightarrow A = \max(t, 1-t) \end{cases}$



• On a $\lambda_0 = 2(1 - A(\frac{1}{2})) > 0$ et $\lambda_L = 0$ sauf pour C^+ (car $A(\frac{1}{2}) = 1$ alors $\lambda_0 = 0$). Soient les

maximas et indep soient les maximas ont une dépendance de queues à droite forte

exco 7: etant que Gumbel $\Leftrightarrow A(t) = [t^\alpha + (1-t)^\alpha]^{\frac{1}{\alpha}}$

Rq: Pour les maximas, 2 comportements sont possibles: $\rightarrow \lambda_0 > 0 \Leftrightarrow$ dépendance de queue à droite.

ou $\lambda_0 = 0 \Rightarrow$ indépendance.

Le cas de la copule usnormale est exclu ($\lambda_0 = 0$ et dépendance $\rho < 1$).

VI Inférence statistique.

Soient n observations iid d'un couple $(X_i, Y_i)_{1 \leq i \leq n}$ tel que $(X_i, Y_i) \sim F$.

On souhaite estimer la copule :

$$C(X, Y)(x, y) = F(F_1^{-1}(x), F_2^{-1}(y))$$

où $X_1 \sim F_1$ et $X_2 \sim F_2$.

A - Approche non-paramétrique

def: On appelle statistique de rang r_i de l'observation x_i son rang dans l'échantillon ordonné

$$(X_{(1)}, X_{(2)}, \dots, X_{(n)}) = (X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}).$$

ex: $(X_1, X_2, X_3, X_4) = (4, 8, 2, 5)$

$\Rightarrow (r_1, r_2, r_3, r_4) = (2, 4, 1, 3)$.

4 est ^{2ème} petite valeur dans l'échantillon ordonné.

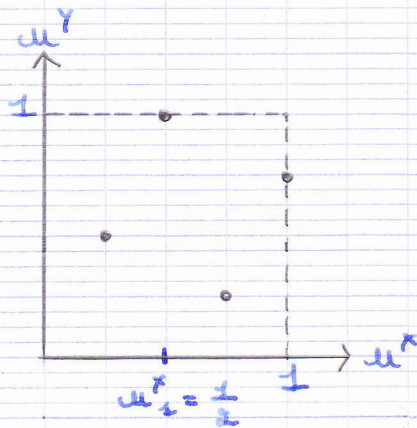
def: Le dépendogramme est le graphique des points $(u_i^x, u_i^y), \dots, (u_n^x, u_n^y)$ où $u_i^x = \frac{r_i^x}{n}$ et

$$u_i^y = \frac{r_i^y}{n} \text{ où } r_i^x \text{ est le rang de } x_i \text{ et } r_i^y \text{ est}$$

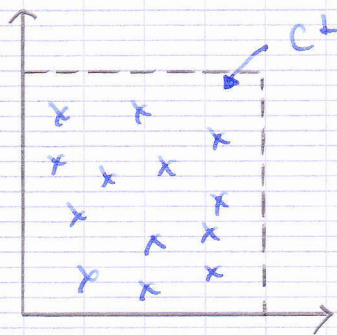
le rang de y_i .

ex: $(Y_1, Y_2, Y_3, Y_4) = (3, 2, 1, 0)$

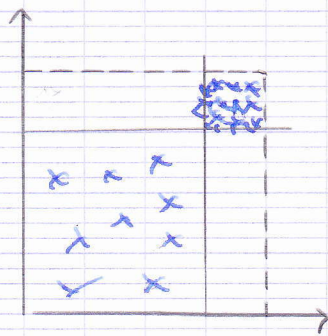
$$\Rightarrow (r_1^Y, r_2^Y, r_3^Y, r_4^Y) = (4, 3, 2, 1).$$



Dépendogrammes classiques:



(l'indépendance)
des queues
(répartition uniforme)



(dépendance de
queue à droite à droite)
quand on observe des
valeurs extrêmes pour X
alors on observe des va-
leurs extrêmes pour Y

Les valeurs extrêmes de Y sont observés que lorsqu'il y a des valeurs extrêmes pour X (et vice-versa)

Graphiquement, un amas de points en haut à droite correspond à une dépendance en queue à droite

On compare le dépendogramme des observations avec celui d'un couple $(U_i^1, U_i^2)_{1 \leq i \leq n}$ iid avec $(U_i^1, U_i^2) \sim$ Copule classique (normale, C^+ , C^- , C^+ Frank, Gumbel, Clayton), par simulation.

B - Simulation d'un dépendogramme.

On souhaite simuler une réalisation de $U = (U_1, U_2) \sim C$. On note :

$$C_{211}(u_1, u_2) = \mathbb{P}(U_2 \leq u_2 \mid U_1 = u_1)$$

la distribution conditionnelle de C

Prop 8: • montrer que : $C_{211}(u_1, u_2) = \frac{\partial C}{\partial u_1}(u_1, u_2)$

$$C_{211}(u_1, u_2) = \lim_{u \rightarrow 0^+} \mathbb{P}(U_2 \leq u_2 \mid u_1 \leq U_1 \leq u_1 + u)$$

• montrer que, conditionnellement à $\{U_1 = u_1\}$, $C_{211}(u_1, U_2) \sim \mathcal{U}([0, 1])$.
fixé

On procède en 3 étapes :

1] Simuler w_1 et w_2 deux réalisations $\mathcal{U}([0, 1])$.

2] Poser $u_1 = w_1$.

3] Prendre $U_2 = G^{-1}(w_2)$ où $G(u) = C_{211}(u_1, u)$

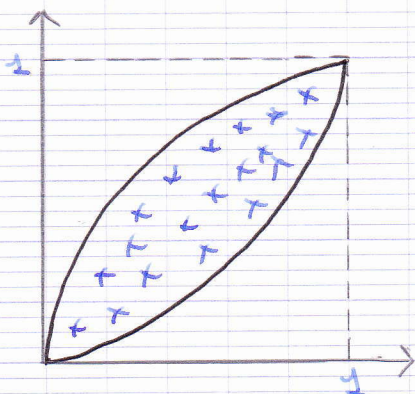
ex: Calculer C_{211} pour la Gumbel.

G^{-1} = inverse généralisée (existe toujours)

cas copules normales :

Pour $\epsilon > 0$:

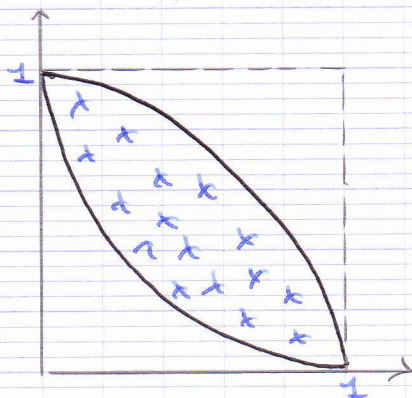
amande = région de confiance du dépendogramme



on s'approche de 1 selon 1 seule direction (car épaisseur de l'amande est restreinte) d'où $\lambda_0 = 0$

⇒ on n'a pas de phénomène d'amas

Pour $\rho < 0$:



C - La méthode du maximum de vraisemblance

On suppose que :

- F_1 est paramétrée par θ_1
- F_2 est paramétrée par θ_2
- C est paramétrée par θ_c

def: Le logarithme de la vraisemblance de $(X_i, Y_i)_{1 \leq i \leq n}$ vaut :

$$\begin{aligned} \ln(F(\theta)) &= \sum_{i=1}^M \ln(c(F_1(x_i; \theta_1), F_2(y_i; \theta_2))) \\ &+ \sum_{i=1}^M \ln(f_1(x_i; \theta_1)) \\ &+ \sum_{i=1}^M \ln(f_2(y_i; \theta_2)) \end{aligned}$$

avec $\theta = (\theta_1, \theta_2, \theta_c)$.

On a alors :

$$\hat{\theta}_{ML} = \underset{\theta \in \Theta}{\operatorname{argmax}} (\ell(\theta))$$

exo 9 : Pour la Gumbel avec $U_1, U_2 \sim U(0, 1]$

autre méthode : méthode IFM :

$$\hat{\theta}_{IFM} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_c)$$

où :

$$\rightarrow \hat{\theta}_1 = \underset{\theta_1}{\operatorname{argmax}} \sum_{i=1}^M \ln(f_1(x_i; \theta_1))$$

$$\rightarrow \hat{\theta}_2 = \underset{\theta_2}{\operatorname{argmax}} \sum_{i=1}^M \ln(f_2(y_i; \theta_2))$$

$$\rightarrow \hat{\theta}_c = \underset{\theta_c}{\operatorname{argmax}} \sum_{i=1}^M \ln(c(F_1(x_i; \hat{\theta}_1), F_2(y_i; \hat{\theta}_2); \theta_c))$$