

**La Théorie des Extrêmes
et
La Gestion des Risques de Marché**

Cours ENSAI de 3^{ème} année

Thierry RONCALLI

**Groupe de Recherche Opérationnelle
Crédit Lyonnais
Bercy-Expo – Immeuble Bercy Sud – 4^{ème} étage
90, Quai de Bercy — 75613 Paris Cedex 12
France**

thierry.roncalli@creditlyonnais.fr

Janvier 2001

Table des matières

Introduction	1
1 Quelques éléments sur la théorie et les distributions unidimensionnelles des extrêmes	3
1.1 La loi du maximum (ou du minimum)	3
1.2 Le théorème de Fisher-Tippet	10
1.3 Complément bibliographique	11
2 La distribution généralisée des extrêmes	15
2.1 Définition de la distribution GEV	15
2.2 Quelques propriétés	16
2.3 Estimation par la méthode du maximum de vraisemblance	17
2.4 Applications	20
2.4.1 Le concept de temps de retour	20
2.4.2 Construction de la VaR GEV	21
2.4.3 Construction de scénarios de crise unidimensionnels	23
2.5 Complément bibliographique	25
3 Les distributions multidimensionnelles des extrêmes	27
3.1 Les copules	27
3.2 Distributions bidimensionnelles des extrêmes	33
3.2.1 Max-domaine d'attraction	33
3.2.2 Première caractérisation des copules à valeurs extrêmes	33
3.2.3 Seconde caractérisation des copules à valeurs extrêmes	33
3.2.4 Le coefficient de dépendance de queue	34
3.3 Estimation par la méthode du maximum de vraisemblance	36
3.4 Application au stress-testing	38
3.5 Complément bibliographique	42
Conclusion	45

Introduction

Ce cours s'intitule "La théorie des extrêmes et la gestion des risques de marché" ; c'est un cours qui se veut pratique. Du fait notamment du format du cours (5 heures), je n'aborderai pas les aspects probabilistes de la théorie des extrêmes. Je consacrerai les cinq heures à l'application de la théorie des extrêmes à la gestion des risques de marché. Il est donc impératif que vous ayez suivi le module "Théorie des valeurs extrêmes" de Marian Hristache.

Cependant, afin que ce cours soit relativement indépendant, j'indique ici une bibliographie sélective. Dans un souci de simplicité, il est préférable dans un premier temps de consulter les ouvrages, et de lire les articles dans un second temps. Les deux livres d'Emile Gumbel [Gumbel 54] [Gumbel 58] sont très accessibles, et contiennent de nombreuses applications. Même s'ils sont relativement vieux, ils sont plaisants à lire, et surtout évitent une certaine mathématisation inutile que nous retrouvons dans certains ouvrages d'aujourd'hui. Je vous conseille aussi EMBRECHTS, KLÜPPELBERG et MIKOSCH [1997] et surtout KOTZ et NADARAJAH [2000]. Ce dernier ouvrage est très court — moins de 200 pages — et présente l'essentiel de la théorie, ainsi que de nombreuses références historiques. Je mentionne enfin deux autres livres qui sont d'un niveau mathématique plus élevé : celui [Galambos] de Janos Galambos et celui [Resnick] de Sidney Resnick.

Ces deux derniers ouvrages contiennent un chapitre consacré à la théorie multivariée des extrêmes. Néanmoins, ils sont difficiles d'accès, et il est sûrement préférable de consulter le troisième chapitre de KOTZ et NADARAJAH [2000]. Le problème majeur avec ces 3 livres est qu'ils abordent la théorie des extrêmes dans un cadre multidimensionnel sans faire référence aux copules, ce qui complique la formalisation sans pour autant donner des théorèmes plus généraux. Pour avoir un exposé rigoureux basé sur les copules, vous pouvez consulter le chapitre 6 de JOE [1997]. Le chapitre 53 de [KBJ] peut être considéré comme une "vulgarisation" et une simplification du chapitre de JOE [1997].

Enfin, il me serait difficile de ne pas citer trois autres travaux qui vont servir de fil conducteur pour ce cours. Le premier concerne le Groupe de Travail ENSAE 1999/2000 portant le même nom que ce cours. Le document [BN] d'Antoine Bezat et d'Ashkan Nikeghbali est de mon point de vue le plus complet, et d'une grande qualité pédagogique pour comprendre comment appliquer la théorie des extrêmes à la gestion du risque. Il comporte de plus un chapitre tout à fait abordable sur le cas multidimensionnel. Arnaud Costinot a effectué un stage au GRO sur la construction de programmes quantitatifs de scénarios de crise (stress-testing). Vous trouverez un résumé de son

travail dans [CRR]. Enfin, le paragraphe 4.3.2 de [BDNRR] contient des éléments d'application de la théorie des extrêmes dans le cas multidimensionnel. Certains des exercices qui suivent sont directement inspirés de ce paragraphe, ainsi que du document [BN].

A la fin de chaque chapitre, vous trouverez une bibliographie complémentaire. Celle-ci est exclusivement composée d'articles de référence. Si vous désirez approfondir le sujet, elle complète la bibliographie de base ci-dessous.

Remarque 1 *Ce document est un support du cours contenant un certain nombre d'exercices et l'ensemble du code GAUSS. La partie explicative des programmes ainsi que les rappels théoriques seront présentés lors du cours. A cette occasion, n'hésitez pas à intervenir lorsque la solution d'un exercice ne vous paraît pas évidente ou lorsque des points théoriques restent obscurs.*

- [BN] BEZAT, A. et A. NIKEGHBALI [2000], La théorie des extrêmes et la gestion des risques de marché, GT ENSAE (document disponible sur le site web du GRO)
- [BDNRR] BOUYÉ, E., DURRLEMAN, V., A. NIKEGHBALI, G. RIBOULET and T. RONCALLI [2000], Copulas for finance : a reading guide and some applications, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (document disponible sur le site web du GRO)
- [CRR] COSTINOT, A., G. RIBOULET et T. RONCALLI [2000], Stress-testing et théorie des valeurs extrêmes : une vision quantifiée du risque extrême, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper* (document disponible sur le site web du GRO)
- [EKM] EMBRECHTS, P., C. KLÜPPELBERG et T. MIKOSCH [1997], Modelling Extremal Events for Insurance and Finance, Springer-Verlag, Berlin
- [Galambos] GALAMBOS, J. [1987], The Asymptotic Theory of Extreme Order Statistics, seconde édition, Krieger (première édition : John Wiley & Sons, 1987)
- [Gumbel 54] GUMBEL, E.J. [1954], Statistical Theory of Extreme Values and Some Practical Applications, *Applied Mathematics Series*, **33**, National Bureau of Standards, Washington
- [Gumbel 58] GUMBEL, E.J. [1958], Statistics of Extremes, Columbia University Press, Columbia
- [Joe] JOE, H. [1997], Multivariate Models and Dependence Concepts, Chapman & Hall, London
- [KBJ] KOTZ, S., N. BALAKRISHNAN et N.L. JOHNSON [2000], Continuous Multivariate Distributions — Volume 1 : Models and Applications, John Wiley & Sons, New York
- [KN] KOTZ, S. et S. NADARAJAH [2000], Extreme Value Distributions, Imperial College Press, London
- [Resnick] RESNICK, S.I. [1987], Extreme Values, Point Processes and Regular Variation, Springer-Verlag, New York

1

Quelques éléments sur la théorie et les distributions unidimensionnelles des extrêmes

1.1 La loi du maximum (ou du minimum)

Nous cherchons à caractériser la fonction de distribution \mathbf{G}_N de la statistique d'ordre χ_N^+ définie par

$$\chi_N^+ = \bigvee_{n=1}^N X_n = \max(X_1, \dots, X_n, \dots, X_N) \quad (1.1)$$

Nous supposons que les variables aléatoires sont indépendantes et de même loi de distribution \mathbf{F} .

Exercice 1 *Montrer que \mathbf{G}_N a pour expression*

$$\mathbf{G}_N(x) = [\mathbf{F}(x)]^N \quad (1.2)$$

En déduire la fonction de densité g_N associée, ainsi que l'expression du quantile.

Exercice 2 *Montrer que la loi du minimum χ_N^- est*

$$\mathbf{G}_N(x) = 1 - [1 - \mathbf{F}(x)]^N \quad (1.3)$$

Exercice 3 *Généraliser les résultats précédents lorsque les variables aléatoires X_n présentent des distributions différentes \mathbf{F}_n .*

Exercice 4 *Généraliser les résultats précédents lorsque les variables aléatoires X_n ne sont pas indépendantes.*

Exercice 5 *Généraliser les résultats précédents lorsque les variables aléatoires X_n présentent des distributions différentes \mathbf{F}_n et ne sont pas indépendantes.*

Remarque 2 *Les deux derniers exercices peuvent être facilement résolus si on utilise une représentation copule. Dans ce cas, les lois du maximum et du minimum correspondent à la section diagonale de la copule et du dual de la copule.*

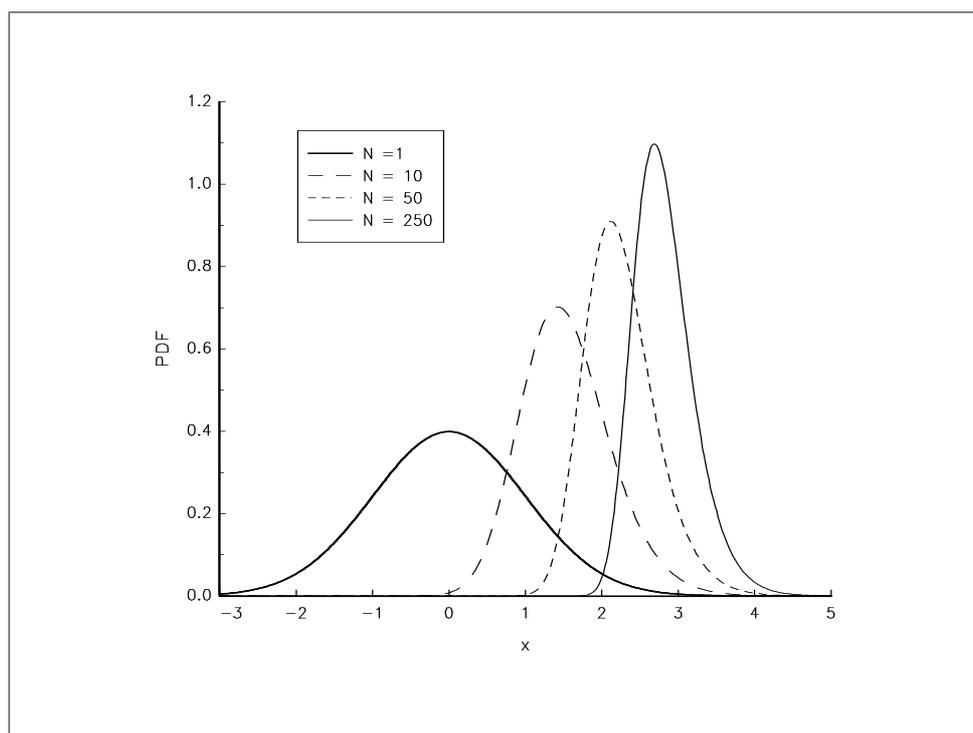
Remarque 3 *En remarquant que $|\mathbf{G}_1^{-1}(\alpha)|$ est la Valeur en Risque à un jour d'un portefeuille dont la variation suit la distribution \mathbf{F} , quelle est l'interprétation du quantile α de \mathbf{G}_N ?*

Voyons un exemple de programme GAUSS qui calcule la densité de la loi du maximum lorsque \mathbf{F} est la distribution gaussienne standardisée :

```
new;
library pgraph;

N = 1~10~50~250;
x = seqa(-10,20/1000,1000);
pdf = N .* cdfn(x)^(N-1) .* pdfn(x);

graphset;
_pdate = '''; _pnum = 2; _pframe = 0;
_pltype = 6|1|3|6; _plwidth = 5|0|0|0;
xtics(-3,5,1,0);
ytics(0,1.2,.2,2);
xlabel('x');
ylabel('PDF');
_plegstr = ''N =1\000N = 10\000N = 50\000N = 250'';
_plegctl = {2 5 2 5};
graphprt(''-c=1 -cf=g1.ps'');
xy(x,pdf);
```



Graphique 1.1. PDF de G_N pour différentes valeurs de N

Dans le programme qui suit, nous estimons les paramètres de la distribution \mathbf{F} par maximum de vraisemblance à partir de la loi du maximum.

```
new;
library pgraph;
```

```

rndseed 123;
N = 50;
M = 250;
data = 0.5*randn(N,M) + 2; /* random numbers */
x = maxc(data);          /* componentwise maximum */

proc ml(theta);
  local mu,sigma,xc,pdf;

  mu = theta[1];
  sigma = sqrt(theta[2]^2);
  xc = (x - mu)./sigma;
  pdf = N .* cdfn(xc)^(N-1) .* pdfn(xc) ./ sigma;

  retp( ln(pdf) );
endp;

proc negml(theta);
  retp( - sumc(ml(theta)) );
endp;

let sv = 1 1;
_qn_PrintIters = 1;

output file = g2.out reset;

{theta,f,g,retcode} = Qnewton(&negml,sv); /* Optimization */

H = hessp(&negml,theta); /* - Hessian */
cov = invpd(H);          /* Covariance matrix - form 1 */
stderr = sqrt(diag(cov)); /* standard errors of the ML estimates */
tstudent = theta ./ miss(stderr,0); /* test t */
df = rows(x) - rows(theta); /* degrees of freedom */
pvalue = 2*cdfn(abs(tstudent),df); /* p-value */

let parnames = 'mu' 'sigma';
mask=0~1~1~1~1;
let fmt[5,3]= ' *.*s' 8 8 ' *.*lf' 16 6 ' *.*lf' 16 6
              ' *.*lf' 17 6 ' *.*lf' 16 6;
omat = parnames~theta~stderr~tstudent~pvalue;
print; print;
print 'Parameters      estimates      std.err.    '\
      ' t-statistic      p-value    ';
print '-----'\
      '-----';
call printfm(omat,mask,fmt);

output off;

```

Nous obtenons les résultats suivants :

```

=====
QNewton Version 3.2.39                      1/09/2001 10:03 am
=====

return code =      0
normal convergence

Value of objective function      -25.531119

Parameters      Estimates      Gradient
-----
P01              2.0531          0.0001
P02              0.4806          0.0001

Number of iterations      10

```

Minutes to convergence 0.00283

Parameters	estimates	std.err.	t-statistic	p-value
mu	2.053096	0.047049	43.637009	0.000000
sigma	0.480573	0.022499	21.360149	0.000000

Remarque 4 *Comment pourrions-nous utiliser le programme précédent pour faire un test d'adéquation de distribution des extrêmes ?*

Exercice 6 *Nous utilisons la base de données cac-dj. Nous voulons modéliser le rendement de l'indice CAC40 de la façon suivante :*

$$\frac{r_t - \mu}{\sigma} \sim \mathbf{F} \quad (1.4)$$

avec \mathbf{F} une distribution unidimensionnelle de support \mathbb{R} . Estimer les paramètres (μ, σ) lorsque \mathbf{F} est respectivement la distribution gaussienne $\mathcal{N}(0, 1)$ et la distribution de Student à 3 degrés de liberté t_3 . Quelle est l'incidence du choix de la distribution sur la loi du maximum ? En utilisant les éléments de réponse de la remarque précédente, construire un test de Kolmogorov-Smirnov de la distribution du maximum. Discuter alors l'influence du caractère leptokurtique de la distribution des rendements du CAC40.

Voici le programme GAUSS correspondant à la première partie de l'exercice.

```
new;
library pgraph;

data = load('cac-dj');
CAC40 = data[.,2]; /* CAC 40 */
y = ln(CAC40);
r = packr(y - lag1(y)); /* (log)returns of the CAC40 index */

muGaussian = meanc(r); /* estimated parameters for the Gaussian distribution */
sigmaGaussian = stdc(r);

nu = 3; /* df of the Student distribution */
muStudent = meanc(r); /* estimated parameters for the Student distribution */
sigmaStudent = stdc(r)/sqrt(nu/(nu-2));

N = 1^22^262^262*10;

x = seqa(-0.20,0.50/1000,1001);
x1 = (x - muGaussian) ./ sigmaGaussian;
g1 = N .* cdfn(x1)^(N-1) .* pdfn(x1) ./ sigmaGaussian;

x2 = (x - muStudent) ./ sigmaStudent;
g2 = N .* cdfn(x2,nu)^(N-1) .* pdfn(x2,nu) ./ sigmaStudent;

graphset;
begwind;
window(2,2,0);

_pframe = 0; _pnum = 2; _pltype = 6|3;
_pnumht = 0.20; _ptitlht = 0.25; _paxht = 0.20;
xlabel('return of the CAC40 index');
ylabel('pdf of G|N[']');

setwind(1);
title('One trading day');
xtics(-0.05,0.10,0.05,2);
ytics(0,50,10,2);
xy(x,g1[.,1]^g2[.,1]);
```

```

setwind(2);
  title('One month of trading');
  xtics(0,0.15,0.05,2);
  ytics(0,60,20,2);
  xy(x,g1[.,2]^g2[.,2]);

setwind(3);
  title('One year of trading');
  xtics(0,0.20,0.05,2);
  ytics(0,80,20,2);
  xy(x,g1[.,3]^g2[.,3]);

setwind(4);
  title('Ten years of trading');
  xtics(0,0.25,0.05,2);
  ytics(0,90,30,2);
  _plegstr = 'Gaussian\000Student';
  _plegctl = {2 6 5 3};
  xy(x,g1[.,4]^g2[.,4]);

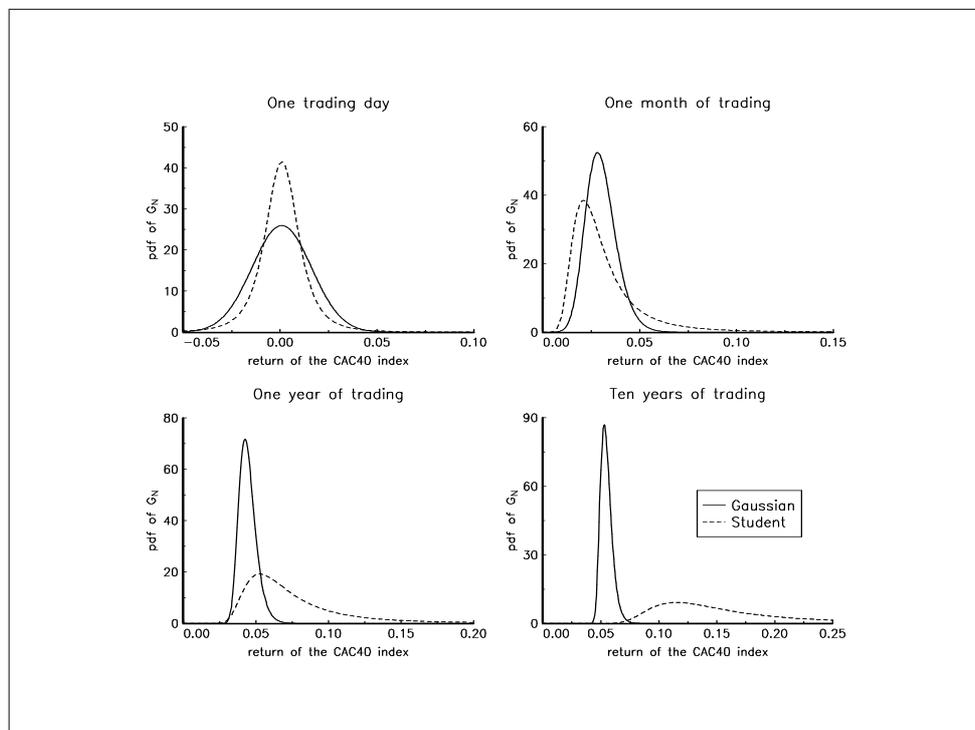
  graphprt('-c=1 -cf=cac1.ps');

endwind;

proc cdft(x,nu);
  retp( 1 - cdftc(x,nu) );
endp;

proc pdft(x,nu);
  retp( gamma((nu+1)/2) ./ gamma(nu./2) ./ sqrt(nu.*pi) .*
    (1+x^2./nu)^(-(nu+1)/2) );
endp;

```



Graphique 1.2. Modélisation de la loi du maximum du rendement de l'indice CAC40

Remarque 5 Nous voyons clairement sur le graphique l'influence du choix de la distribution \mathbf{F} sur la loi du maximum des rendements du CAC40. A titre d'illustration, modifier la valeur des degrés de liberté ν de la distribution de Student. En particulier, vérifier que nous obtenons le cas de la distribution gaussienne lorsque ν tend vers $+\infty$.

Exercice 7 A partir des estimations précédentes, montrer que la loi gaussienne est incapable d'expliquer les rendements extrêmes du CAC40. Pour cela, comparer la fonction quantile \mathbf{G}_N^{-1} avec les rendements les plus extrêmes¹. Calculer la probabilité que χ_N^+ prenne des valeurs supérieures aux statistiques d'ordre observées $r_{(m)}$ pour $m = 1, 2, 3, 4, 5, 10, 15, 20$. Conclure.

```

new;
library pgraph;

data = loadadd('cac-dj');

CAC40 = data[:,2];      /* CAC 40 */

y = ln(CAC40);
r = packr(y - lag1(y)); /* (log)returns of the CAC40 index */

muGaussian = meanc(r); /* estimated parameters for the Gaussian distribution */
sigmaGaussian = stdc(r);

nu = 3; /* df of the Student distribution */

muStudent = meanc(r); /* estimated parameters for the Student distribution */
sigmaStudent = stdc(r)/sqrt(nu/(nu-2));

N = rows(r); /* Number of observations of the database */

alpha = seqa(1/100,1/100,90);
x = cdfni(alpha^(1./N));
x1 = sigmaGaussian * x + muGaussian;
x = cdfpci(1-alpha^(1./N),nu);
x2 = sigmaStudent * x + muStudent;

r = rev(sortc(r,1));
Nmax = 10;
rmax = r[1:Nmax];

graphset;
  begwind;
  window(1,2,0);
  _pnum = 2; _pframe = 0;
  _paxht = 0.25; _pnumht = 0.20; _ptitlht = 0.25;
  fonts('simplex simgrma');
  ylabel('return (in percent)');
  xlabel('\202a\201');
  xtics(0,1,0.2,2);

  setwind(1);
  title('Gaussian distribution');
  _pline = ones(Nmax,1).*(1~3~0)~100*rmax~ones(Nmax,1)~100*rmax~ones(Nmax,1).*(1~15~0);
  ytics(4,7,0.5,0);
  xy(alpha,100*x1);

  setwind(2);
  title('Student distribution');
  ytics(5,20,2.5,0);
  xy(alpha,100*x2);

```

¹ N est égal à la taille de la base de données.

```

graphprt(''-c=1 -cf=cac2.ps'');

endwind;

let indx = 1 2 3 4 5 10 15 20;
rmax = r[indx];
x1 = (rmax - muGaussian) ./ sigmaGaussian;
Pr1 = 1-cdfn(x1)^N;

x2 = (rmax - muStudent) ./ sigmaStudent;
Pr2 = 1-(1-cdfstc(x2,nu))^N;

output file = cac2.out reset;

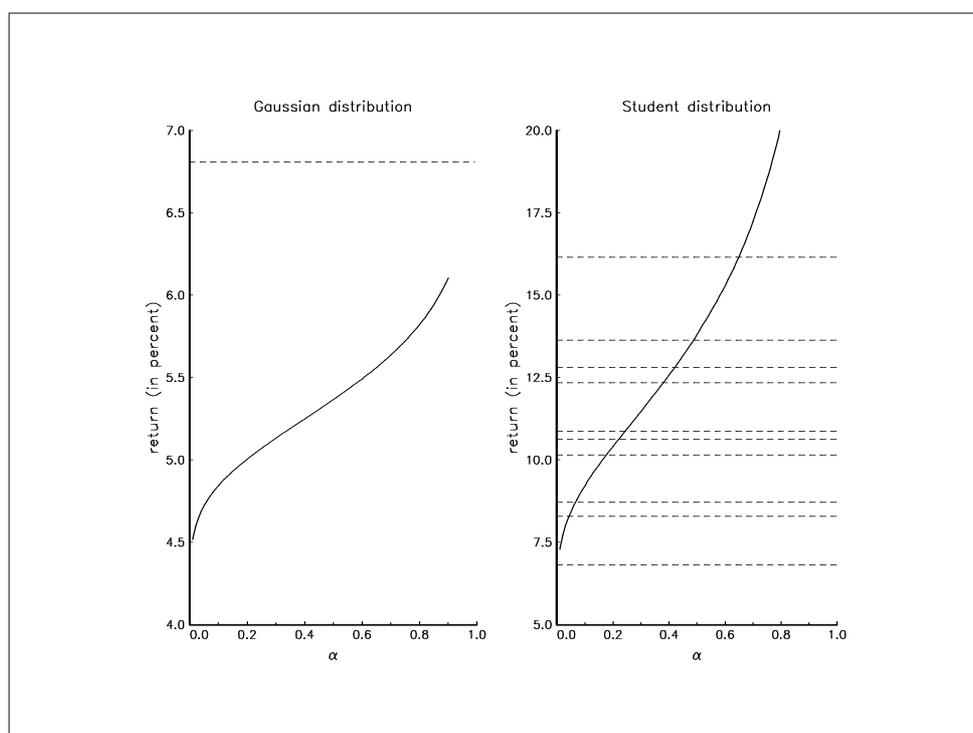
call printfm(100*rmax~pr1~pr2,1~1~1,'%lf''^8~3);

output off;

```

Les probabilités $\Pr \{ \chi_N^+ > r_{(m)} \}$ sont les suivantes

m	$r_{(m)}$	$\mathcal{N}(0,1)$	t_3
1	16.15%		35.2%
2	13.63%		51.3%
3	12.81%		58.0%
4	12.34%		62.1%
5	10.87%		75.7%
10	6.81%	1.4%	99.6%
15	5.12%	70.8%	100%
20	3.86%	100%	100%



Graphique 1.3. Quantile $G_N^{-1}(\alpha)$

1.2 Le théorème de Fisher-Tippett

Théorème 1 (Fisher-Tippett) *Supposons N variables aléatoires $X_1, \dots, X_n, \dots, X_N$ indépendantes et de même loi de distribution \mathbf{F} . S'il existe des constantes a_N et b_N et une distribution limite non dégénérée \mathbf{G} telles que*

$$\lim_{N \rightarrow \infty} \Pr \left(\frac{\chi_N^+ - b_N}{a_N} \leq x \right) = \mathbf{G}(x) \quad \forall x \in \mathbb{R} \quad (1.5)$$

alors \mathbf{G} appartient à l'un des trois types suivants de distribution :

(Gumbel)	$\mathbf{G}(x) = \Lambda(x) = \exp(-e^{-x})$	$x \in \mathbb{R}$	$g(x) = \exp(-x - e^{-x})$
(Fréchet)	$\mathbf{G}(x) = \Upsilon_\alpha(x) = \begin{cases} 0 & x \leq 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases}$	$x \leq 0$ $x > 0$	$g(x) = \begin{cases} 0 & \\ \alpha x^{-(1+\alpha)} \exp(-x^{-\alpha}) & \end{cases}$
(Weibull)	$\mathbf{G}(x) = \Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & x \leq 0 \\ 1 & x > 0 \end{cases}$	$x \leq 0$ $x > 0$	$g(x) = \begin{cases} \alpha (-x)^{\alpha-1} \exp(-(-x)^\alpha) & \\ 0 & \end{cases}$

Exercice 8 *Considérer la loi exponentielle \mathcal{E}_1 — $\mathbf{F}(x) = 1 - e^{-x}$. En posant $a_N = 1$ et $b_N = \ln N$, montrer **analytiquement** que $(\chi_N^+ - b_N)/a_N$ tend asymptotiquement vers la distribution de Gumbel.*

Exercice 9 *Reprendre l'exercice précédent, et illustrer graphiquement la convergence pour $N = 2, 5, 10, 15$. Étudier la vitesse $\sup \left| \Lambda(x) - [\mathbf{F}(x + \ln N)]^N \right|$ de la convergence en loi. Conclure.*

```

new;
library pgraph;

proc cdfExp(x);
  local cdf;
  cdf = 1 - exp(-x);
  retp( missex(cdf, cdf .< 0) );
endp;

proc (1) = cdfMax(CDF,N,x);
  local CDF:proc;
  local y;

  y = CDF(x)^N;

  retp(y);
endp;

proc cdfGumbel(x);
  local cdf;

  cdf = exp(-exp(-x));

  retp(cdf);
endp;

x = seqa(-5,0.01,1001);
N = 2^5~10~15;

F = cdfExp(x);
F_n = cdfMax(&cdfEXP,N,x);

```

```

F_n2 = cdfMax(&cdfEXP,N,x+ln(N));
F_n3 = cdfMax(&cdfEXP,50,x+ln(50));
G = cdfGumbel(x);

graphset;
begwind;
window(2,2,0);

_pdate = '''; _pnum = 2; _pltype = 6|3|1|4;
_pnumht = 0.20; _paxht = 0.25; _pframe = 0;
fonts('simplex simgrma');

ytics(0,1,0.2,2);
xtics(0,5,1,2);

setwind(1);
xlabel('x');
ylabel('F(x)');
xy(x,F);

setwind(2);
ylabel('2020[\201F(x)\2020]\201[N]');
_plegstr = 'N = 2\000N = 5\000N = 10\000N = 15';
_plegctl = {2 6 6 2};
xy(x,F_n);

xtics(-2,5,1,2);

setwind(3);
ylabel('2020[\201F(x+ln(N))\2020]\201[N]');
_plegctl = {2 6 5 2};
xy(x,F_n2);

setwind(4);
ylabel('CDF');
_plegstr = 'Gumbel\000N = 50';
_plwidth = 0|2.5;
xy(x,G~F_n3);

graphprt('-c=1 -cf=conv1.ps');

endwind;

N = seqa(1,1,100)';
F_ = cdfMax(&cdfEXP,N,x+ln(N));
eps = maxc(abs(F_-G));

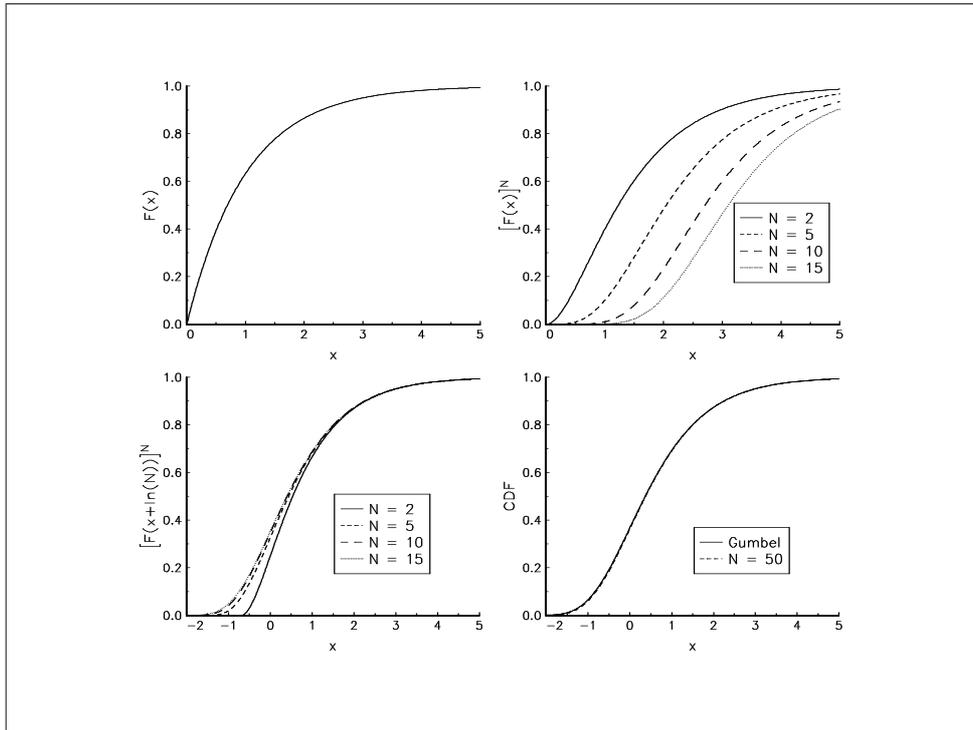
graphset;
_pdate = '''; _pnum = 2; _pframe = 0;
xlabel('N');
graphprt('-c=1 -cf=conv2.ps');
xy(N',ln(eps));

```

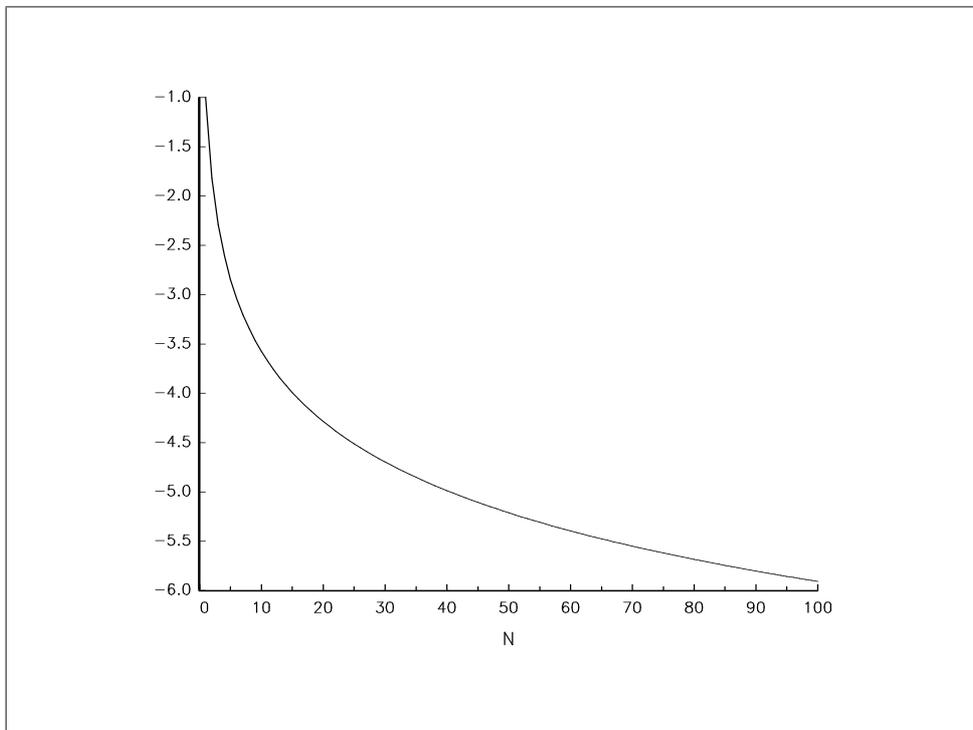
Exercice 10 Reprendre le code *GAUSS* précédent et illustrer la convergence graphiquement à partir des fonctions de densité.

1.3 Complément bibliographique

L'un des articles les plus aboutis sur la théorie unidimensionnelle des extrêmes est celui de GNEDENKO [1943]. Vous trouverez des informations sur les statistiques d'ordre dans le manuel de référence de DAVID [1981].



Graphique 1.4. Convergence de $[\mathbf{F}(x + \ln N)]^N$ vers la distribution de Gumbel



Graphique 1.5. Graphe de $\ln \left(\sup \left| \Lambda(x) - [\mathbf{F}(x + \ln N)]^N \right| \right)$

- [David] DAVID, H. A. [1981], Order Statistics, *Wiley Series in Probability and Mathematical Statistics*, seconde édition, John Wiley & Sons, New York (première édition : 1970)
- [Gnedenko] GNEDENKO, B. [1943], Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics*, **44**, 423-453

2

La distribution généralisée des extrêmes

2.1 Définition de la distribution GEV

La distribution GEV est définie de la façon suivante

$$\mathbf{G}(x) = \begin{cases} \exp\left(-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right) & \text{si } \xi \neq 0 \\ \exp\left(-\exp\left(\frac{x-\mu}{\sigma}\right)\right) & \text{si } \xi = 0 \end{cases} \quad (2.1)$$

Le support de \mathbf{G} est $\Delta = \{x : 1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0\}$. Cette fonction de distribution correspond à la loi de probabilité de Von Mises, mais elle est plus connue sous le nom de “**Generalized Extreme Value distribution**” (**GEV**). Nous avons alors les correspondances suivantes :

Gumbel	$\xi = 0$
Fréchet	$\xi > 0$
Weibull	$\xi < 0$

Exercice 11 *Ecrire une procédure qui calcule $\mathbf{G}(x)$.*

```
proc cdfGEV(x,mu,sigma,xi);
  local y,e,cdf;

  x = (x - mu) ./ sigma;
  y = 1 + xi .* x;
  e = y .> 0;
  y = y .* e + _macheps .* (1-e);
  cdf = exp(-(y^(-1./xi)));
  cdf = cdf .* e + (1-e) .* (xi .< 0);

  retp(cdf);
endp;
```

2.2 Quelques propriétés

La fonction de densité de la distribution GEV est

$$g(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1+\xi}{\xi}} \exp \left(- \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right) \quad (2.2)$$

Exercice 12 *Ecrire une procédure qui calcule $g(x)$.*

```
proc pdfGEV(x,mu,sigma,xi);
  local y,e,pdf;

  x = (x - mu) ./ sigma;
  y = (1 + xi .* x);
  e = y .> 0;
  y = y .* e + _macheps * (1-e);
  pdf = (y^(-(1+xi)./xi)) .* exp( -(y^(-1./xi)) ) ./ sigma;
  pdf = pdf .* e;

  retp(pdf);
endp;
```

Le quantile $\mathbf{G}^{-1}(\alpha)$ est donné par la formule suivante :

$$\mathbf{G}^{-1}(\alpha) = \mu - \frac{\sigma}{\xi} \left[1 - (-\ln \alpha)^{-\xi} \right] \quad (2.3)$$

Exercice 13 *Ecrire une procédure qui calcule $\mathbf{G}^{-1}(\alpha)$.*

```
proc cdfGEVi(alpha,mu,sigma,xi);
  local cdfi;

  alpha = missex(missex(alpha,alpha .<= 0), alpha .>= 1);
  cdfi = mu - sigma.*(1 - (-ln(alpha))^-xi)./xi;

  retp(cdfi);
endp;
```

Exercice 14 *Ecrire une procédure qui simule des nombres aléatoires de la distribution $\mathcal{GEV}(\mu, \sigma, \xi)$.*

```
proc rndGEV(r,c,mu,sigma,xi);
  local u;

  u = cdfGEVi(rndu(r,c),mu,sigma,xi);

  retp(u);
endp;
```

2.3 Estimation par la méthode du maximum de vraisemblance

Soit θ le vecteur des paramètres. Nous avons

$$\theta = \begin{bmatrix} \mu \\ \sigma \\ \xi \end{bmatrix}$$

Nous considérons un échantillon de données $\{X_t\}$ de dimension $T = \varphi N$ avec $\varphi \in \mathbb{R}$. Nous divisons cet échantillon en N blocs et nous définissons χ_n^+ de la façon suivante

$$\chi_n^+ = \max(\{X_{1+n(t-1)}, t = 1, \dots, \varphi\}) \quad (2.4)$$

L'expression de la vraisemblance de l'observation n est donc

$$L(\chi_n^+; \theta) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{\chi_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1+\xi}{\xi}} \exp \left(- \left[1 + \xi \left(\frac{\chi_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right) \quad (2.5)$$

Nous en déduisons l'expression suivante pour la log-vraisemblance :

$$\ell(\chi_n^+; \theta) = -\ln \sigma - \left(\frac{1+\xi}{\xi} \right) \ln \left(1 + \xi \left(\frac{\chi_n^+ - \mu}{\sigma} \right) \right) - \left[1 + \xi \left(\frac{\chi_n^+ - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \quad (2.6)$$

L'estimateur du maximum de vraisemblance correspond alors à

$$\hat{\theta}_{\text{ML}} = \arg \max_{\theta \in \Theta} \sum_{n=1}^N \ell(\chi_n^+; \theta) \quad (2.7)$$

avec Θ l'espace des paramètres. L'estimation de l'expression (2.7) est relativement aisée à condition de prendre quelques précautions, notamment pour la spécification de Θ (à cause de la singularité au point $\xi = 0$).

Exercice 15 *Montrer que l'expression de la matrice jacobienne $[\partial_\mu \ell(\chi_n^+; \theta) \quad \partial_\sigma \ell(\chi_n^+; \theta) \quad \partial_\xi \ell(\chi_n^+; \theta)]$ correspond à*

$$\begin{aligned} \partial_\mu \ell(\chi_n^+; \theta) &= \frac{1 + \xi - \omega^{-\frac{1}{\xi}}}{\sigma \omega} \\ \partial_\sigma \ell(\chi_n^+; \theta) &= \frac{\left((1 + \xi) - \omega^{-\frac{1}{\xi}} \right) (\chi_n^+ - \mu) - \sigma \omega}{\sigma^2 \omega} \\ \partial_\xi \ell(\chi_n^+; \theta) &= \left(1 - \omega^{-\frac{1}{\xi}} \right) \left(\frac{1}{\xi^2} \ln \omega - \frac{(\chi_n^+ - \mu)}{\xi \sigma \omega} \right) - \frac{(\chi_n^+ - \mu)}{\sigma \omega} \end{aligned} \quad (2.8)$$

avec

$$\omega = 1 + \xi \left(\frac{\chi_n^+ - \mu}{\sigma} \right) \quad (2.9)$$

Exercice 16 *Ecrire une procédure qui estime les paramètres de distribution $\mathcal{GEV}(\mu, \sigma, \xi)$ par la méthode du maximum de vraisemblance avec un jacobien analytique. Afin de simplifier le code, traiter uniquement le cas $\xi > 0$.*

```

declare matrix _mleGEV_data;
declare matrix _mleGEV_algr = 1;

proc (3) = mleGEV(data,sv);
  local mu0,sigma0,xi0;
  local g,retcode;
  local theta,stderr,Mcov,Logl;
  local mu,sigma,xi,ParNames;

  if sv == 0;
  sigma0 = sqrt(6/pi) * stdc(data);
  mu0 = meanc(data) - 0.5772 * sigma0;
  xi0 = 0.10;
  sv = mu0|sigma0|xi0;
  endif;

  _mleGEV_data = data;

  if _mleGEV_algr == 1;

  _opgdprc = &_mleGEV_gradient;
  {theta,LogL,g,retcode} = optimum(&_mleGEV_fn,sv);

  else;

  _qn_gradProc = &_mleGEV_gradient;
  {theta,LogL,g,retcode} = Qnewton(&_mleGEV_fn,sv);

  endif;

  {mu,sigma,xi} = _mleGEV_parameters(theta);

  retp(mu,sigma,xi);
endp;

proc _mleGEV_fn(theta);
  local mu,sigma,xi,w,logl;

  {mu,sigma,xi} = _mleGEV_parameters(theta);

  w = 1 + xi*(_mleGEV_data-mu)/sigma;
  w = missex(w,w.<=0);

  logl = - ln(sigma) - (1+xi)*ln(w)/xi - w^(-1/xi);

  retp( -sumc(packr(logl)) );
endp;

proc _mleGEV_gradient(theta);
  local mu,sigma,xi,g;
  local w,Xc,w_xi,sigma_w,Xc_w;

  {mu,sigma,xi} = _mleGEV_parameters(theta);

  Xc = (_mleGEV_data-mu)/sigma;
  w = 1 + xi*Xc;
  w = missex(w,w.<=0);

  w_xi = w^(-1/xi);
  sigma_w = sigma * w;
  Xc_w = Xc ./ w;

  g = zeros(rows(_mleGEV_data),3);
  g[.,1] = ((1+xi) - w_xi) ./ sigma_w;
  g[.,2] = (g[.,1]*(_mleGEV_data-mu) - 1)/sigma;
  g[.,3] = (1-w_xi) .* ( ln(w)/xi - Xc_w) / xi - Xc_w;

  if _mleGEV_algr == 1;

```

```

    retp( -sumc(packr(g))' );
else;
    retp( -sumc(packr(g)) );
endif;
endp;

proc (3) = _mleGEV_parameters(theta);
    local mu,sigma,xi;

    mu = theta[1];
    sigma = sqrt(theta[2]^2); /* positive parameter */
    xi = sqrt(theta[3]^2); /* positive parameter */

    retp(mu,sigma,xi);
endp;

```

Exercice 17 *Simuler 100 nombres aléatoires de la distribution $\mathcal{GEV}(1,1,1)$ et estimer les paramètres (μ, σ, ξ) avec la procédure `mleGEV`.*

```

new;
library optnum;

rndseed 123;

#include gev.src;
#include mlegev.src;

mu = 1;
sigma = 1;
xi = 1;

x = rndGEV(100,1,mu,sigma,xi);

sv = mu|sigma|xi;

output file = gev1.out reset;

_mleGEV_algr = 0;
call mleGEV(x,sv);

output off;

```

```

=====
QNewton Version 3.2.39                      1/09/2001  4:07 pm
=====

```

```

return code = 0
normal convergence

```

```

Value of objective function      217.861602

```

Parameters	Estimates	Gradient
P01	1.0336	0.0005
P02	1.0249	-0.0002
P03	1.0095	0.0004

```

Number of iterations      9
Minutes to convergence    0.00000

```

Exercice 18 Compléter la procédure `mleGEV` afin d'obtenir l'ensemble des statistiques (matrice de covariance, test de student, pvalue, etc.). De plus, ajouter un affichage afin que la sortie ressemble à celle-ci :

```
=====
mleGEV - Generalized Extreme Value                1/09/2001   3:59 pm
=====

Total observations:                               100
Missing observations:                             0
Usable observations:                              100
Valid cases:                                      100

Number of parameters:                             3
Number of unrestricted parameters:                3
Degrees of freedom:                               97

Value of the maximized log-likelihood function:    -217.86160
Mean log-likelihood:                              -2.17862

Parameters      estimates      std.err.      t-statistic      p-value
-----
mu              1.033561      0.117272      8.813355         0.000000
sigma           1.024917      0.156417      6.552446         0.000000
xi              1.009484      0.133448      7.564601         0.000000

Covariance matrix: inverse of the negative Hessian.
```

2.4 Applications

Dans la pratique, nous cherchons à modéliser la distribution de χ_n^+ à partir d'un échantillon. Dans le cadre d'applications météorologiques, la définition d'un bloc est souvent très facile et correspond par exemple à une longueur égale à un an. Dans ce cas, on modélise la distribution du maximum annuel. Cependant, en général, le choix de la taille des blocs est difficile. Il faut que la période de référence du bloc soit suffisamment longue pour que la théorie des extrêmes puisse s'appliquer. Il faut aussi que nous disposions d'assez de données pour obtenir des estimations convenables.

En finance, il n'existe pas de consensus sur la taille φ des blocs. Dans les exemples qui suivent, nous prenons φ égal à un mois (par exemple 22 jours de trading).

Remarque 6 Modifier la taille des blocs des exemples qui suivent afin de voir la sensibilité des résultats à φ .

2.4.1 Le concept de temps de retour

Pour appliquer la théorie des extrêmes à la finance, il est nécessaire d'introduire la notion de **temps de retour**. Considérons un événement E ayant une probabilité d'occurrence égale à α . La loi d'apparition de cet événement est alors une variable aléatoire géométrique de probabilité α . Son espérance mathématique est

$$\tau = \frac{1}{\alpha} \tag{2.10}$$

τ est le temps de retour de E .

Exemple 1 *Considérons une VaR 99%. Le temps de retour associé à l'événement "dépassement de la VaR" est*

$$\tau = \frac{1}{1 - 0.99} = 100 \quad (2.11)$$

Si la période de détention est de 1 jour, le temps de retour de la VaR 99% est 100 jours. Cela veut dire qu'on dépasse la VaR journalière en moyenne tous les 100 jours.

Définition 1 *Soit $\alpha = \Pr\{\chi_n^+ \leq x\}$. Le temps de retour de l'événement $\Pr\{\chi_n^+ > x\}$ est égal à $(1 - \alpha)^{-1}$ périodes de référence. Le temps de retour en jours est donc*

$$\tau = \varphi \frac{1}{1 - \alpha} \quad (2.12)$$

Remarque 7 *Nous pouvons comparer les résultats entre une VaR traditionnelle et une VaR GEV à partir de la relation suivante :*

$$\alpha_{\text{GEV}} = 1 - \varphi(1 - \alpha_{\text{VAR}}) \quad (2.13)$$

2.4.2 Construction de la VaR GEV

Remarque 8 *La procédure `qnewton` n'est pas assez performante. C'est pourquoi nous utilisons la bibliothèque `Optmum`.*

Exercice 19 *Calculer la VaR GEV pour les portefeuilles suivants :*

Portefeuille	CAC40	DowJones
P1	1	0
P2	0	1
P3	1	1
P4	1	-1

Comparer celle-ci avec les VaRs gaussienne et historique. Afin d'éviter une profusion de résultats, considérer uniquement les positions courtes.

```
new;
library optmum;

rndseed 123;

#include mlegev.src;
#include gev.src;

data = loadd('cac-dj');

CAC40 = data[,2];      /* CAC 40 */
DJ = data[,3];        /* Dow Jones */

y = packr(CAC40~DJ);
let Portfolio[4,2] = 1 0
                    0 1
                    1 1
                    1 -1;
data = ln(y*Portfolio);
r = real(data - lag1(data));
Nobs = rows(r);

varphi = 22;          /* one month of trading */
```

```

let alphaVaR[1,3] = 0.99 0.996 0.999;
tau = 1./(1-alphaVaR);
alphaGEV = 1 - (1./tau*varphi);

let sv = 0;

__output = 1;

VaR1 = zeros(4,3);
VaR2 = zeros(4,3);
VaR3 = zeros(4,3);

i = 1;
do until i > 4;

    mu = meanc(packr(r[.,i]));
    sigma = stdc(packr(r[.,i]));

    VaR1[i,.] = sigma*cdfni(alphaVaR);          /* Gaussian VaR */

    VaR2[i,.] = -quantile(packr(r[.,i]),1-alphaVaR)'; /* Historical VaR */

    block = reshape(r[.,i],Nobs/varphi,varphi)';
    min = -minc(block);
    {mu,sigma,xi} = mleGEV(min,sv);
    VaR3[i,.] = cdfGEVi(alphaGEV,mu,sigma,xi); /* GEV VaR */

    i = i + 1;
endo;

output file = gev2.out reset;

call printfm(100*VaR1,1,'%lf'~8^3); print;
call printfm(100*VaR2,1,'%lf'~8^3); print;
call printfm(100*VaR3,1,'%lf'~8^3); print;

output off;

```

Nous obtenons les résultats suivants :

VaR Gaussienne			
Portefeuille	VaR 99%	VaR 99.6%	VaR 99.9%
P1	3.780	4.310	5.022
P2	2.896	3.302	3.847
P3	2.473	2.819	3.285
P4	7.604	8.669	10.101

VaR Historique			
Portefeuille	VaR 99%	VaR 99.6%	VaR 99.9%
P1	4.168	7.174	12.221
P2	3.296	5.740	7.408
P3	2.862	4.392	6.773
P4	8.118	14.393	27.726

VaR GEV			
Portefeuille	VaR 99%	VaR 99.6%	VaR 99.9%
P1	3.421	5.153	8.929
P2	2.778	4.482	8.784
P3	2.375	3.691	6.883
P4	6.578	10.501	19.774

Remarque 9 Les résultats précédents sont exprimés en pourcentage. A partir de la dernière valeur des cours boursiers, exprimer les résultats en euros. En déduire le montant consommé des fonds propres pour une période de détention de 10 jours. Discuter de la pertinence de la VaR GEV par rapport aux autres modèles statistiques, à l'évolution de la CAD et de la mise en place de projets d'allocation de fonds propres sur base économique.

Exercice 20 Reprendre le programme et adopter une modélisation en terme de variations de prix et non plus en terme de rendements. Discuter les résultats.

Exercice 21 Reproduire les résultats de BEZAT et NIKEGHBALI [2000] (tables 2.1 et 2.2 pages 21-22).

2.4.3 Construction de scénarios de crise unidimensionnels

Exercice 22 Construire les échelles de risque pour les positions courtes et longues sur les indices CAC40 et DowJones.

```

new;
library optmum,pgraph;

rndseed 1234565;

#include mlegev.src;
#include gev.src;

data = loadadd('cac-dj');

CAC40 = data[:,2];      /* CAC 40 */
DJ = data[:,3];        /* Dow Jones */

y = packr(CAC40~DJ);
data = ln(y);
r = data - lag1(data);
Nobs = rows(r);

varphi = 22;           /* one month of trading */

year = seqa(1,1,100);
tau = 265*year;        /* return time (in trading days) */
alpha = 1 - varphi/tau; /* corresponding confidence level */

let sv = 0;
__output = 0;

stress_Short = zeros(100,2);
stress_Long = zeros(100,2);

i = 1;
do until i > 2;

    block = reshape(r[:,i],Nobs/varphi,varphi)';
    min = -minc(block);
    {mu,sigma,xi} = mleGEV(min,sv);

    stress_Long[:,i] = cdfGEVi(alpha,mu,sigma,xi);

    max = maxc(block);
    {mu,sigma,xi} = mleGEV(max,sv);

    stress_Short[:,i] = cdfGEVi(alpha,mu,sigma,xi);

    i = i + 1;

```

```

endo;

let indx = 5 10 25 50 75 100;

output file = gev3.out reset;

call printfm(-100*stress_Long[indx,.],1,'%lf'~8^2); print;
call printfm(100*stress_Short[indx,.],1,'%lf'~8^2); print;

output off;

graphset;
  _pdate = ''; _pnum = 2; _pframe = 0;
  _pltype = 6|3; _pstype = 8|9; _plctrl = 10;

  xtics(0,100,10,2);
  ytics(-40,0,10,2);
  xlabel('Return time (in years)');
  ylabel('Risk scale (in %)');

  _plegstr = 'CAC40\000DowJones';
  _plegctl = {2 5 6.5 4.5};

graphprt('-c=1 -cf=gev3.ps');

xy(year,-100*stress_Long);

```

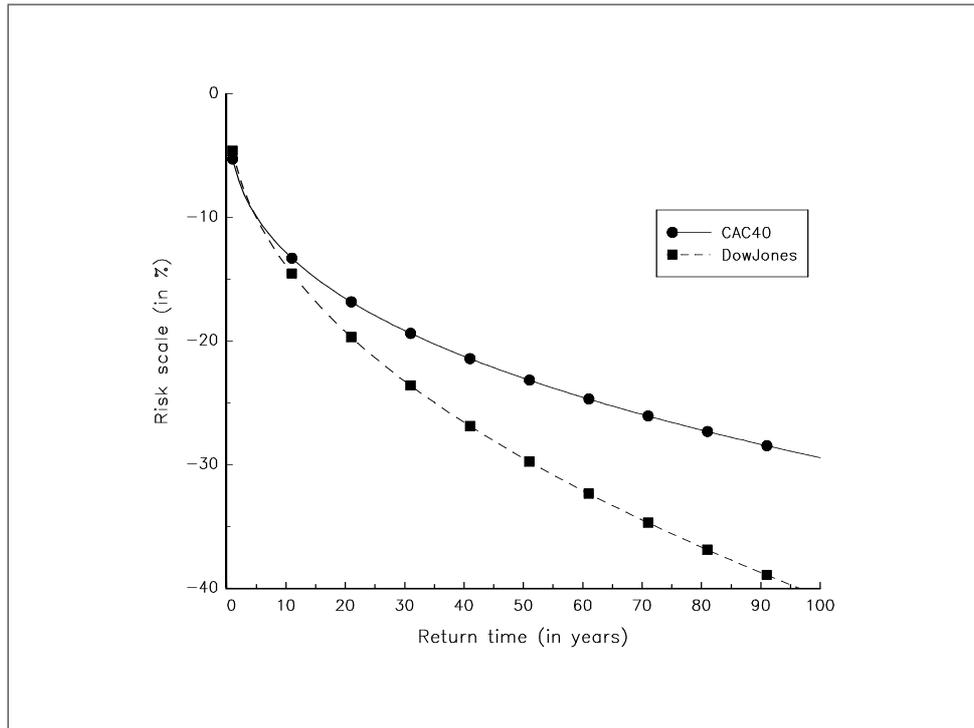
Nous obtenons les résultats suivants :

Echelles de Risque pour des positions longues (en %)		
Temps de Retour	CAC40	DowJones
5	-9.93	-10.04
10	-12.85	-13.91
25	-17.95	-21.35
50	-23.01	-29.48
75	-26.59	-35.59
100	-29.44	-40.67

Echelles de Risque pour des positions courtes (en %)		
Temps de Retour	CAC40	DowJones
5	9.95	6.13
10	13.06	7.38
25	18.66	9.36
50	24.41	11.13
75	28.55	12.30
100	31.91	13.19

Remarque 10 *Si nous comparons ces résultats avec ceux de [CRR], nous voyons qu'ils sont très différents. Montrer alors que le choix de φ n'est pas neutre. Discuter des valeurs pertinentes de φ pour faire des scénarios de crise. Quel est alors l'impact sur la VaR GEV précédente ? Conclure.*

Exercice 23 *Comment pouvons nous adapter le programme précédent pour créer des scénarios de crise pour des portefeuilles d'actifs financiers ?*



Graphique 2.1. Temps de retour pour les positions longues

Exercice 24 A partir des valeurs $(\hat{\mu}_{ML}, \hat{\sigma}_{ML}, \hat{\xi}_{ML})$ estimées précédemment, simuler des trajectoires d'un processus de distribution GEV. Calculer numériquement la puissance de l'estimateur ML en fonction de la taille des trajectoires. A partir de combien d'observations pouvons-nous considérer que nous avons des estimations "convenables"? Discuter alors de la pertinence de l'utilisation de la théorie des extrêmes pour la gestion des risques de marché.

2.5 Complément bibliographique

Vous trouverez la dérivation de la matrice d'information de Fisher dans [PW 80]. Les trois autres articles [Hosking], [Macleod] et [PW 83] sont consacrés à l'implémentation de l'estimateur du maximum de vraisemblance.

- [Hosking] HOSKING, J.R.M. [1985], Maximum likelihood estimation of the parameters of the generalized extreme value distribution, *Applied Statistics*, **34**, 301-310
- [Macleod] MACLEOD, A.J. [1989], Comment on "Maximum likelihood estimation of the parameters of the generalized extreme value distribution", *Applied Statistics*, **38**, 198-199
- [PW 80] PRESCOTT, P. et A.T. WALDEN [1980], Maximum likelihood estimation of the parameters of the generalized extreme value distribution, *Biometrika*, **67**, 723-724
- [PW 83] PRESCOTT, P. et A.T. WALDEN [1983], Maximum likelihood estimation of the parameters of the three-parameter generalized extreme value distribution from censored samples, *Journal of statistical computation and simulation*, **16**, 241-250

3

Les distributions multidimensionnelles des extrêmes

3.1 Les copules

Exercice 25 *Programmer la pdf et la cdf de la copule Normale à deux dimensions.*

```
proc cdfNormal2Copula(u1,u2,rho);
  local cdf;

  cdf = cdfbvn(cdfni(u1),cdfni(u2),rho);

  retp(cdf);
endp;

proc pdfNormal2Copula(u1,u2,rho);
  local x1,x2,x1sqr,x2sqr,rho2,pdf;

  x1 = cdfni(u1);
  x2 = cdfni(u2);

  x1sqr = x1^2;
  x2sqr = x2^2;
  rho2 = rho^2;

  pdf = exp(-0.5 ./ (1-rho2) .* (x1sqr + x2sqr - 2*rho.*x1.*x2)) ./ sqrt(1-rho2) ;
  pdf = pdf .* exp(0.5 .* (x1sqr + x2sqr));

  retp(pdf);
endp;
```

Exercice 26 *Programmer la pdf et la cdf de la copule Gumbel.*

```
proc cdfGumbelCopula(u,v,alpha);
  local cdf;

  u = -ln(u);
  v = -ln(v);
```

```

cdf = exp( - ( u^alpha + v^alpha )^(1./alpha) );

retp(cdf);
endp;

proc pdfGumbelCopula(u,v,alpha);
  local utilde,vtilde,cdf,pdf,beta,w;

  utilde = -ln(u);
  vtilde = -ln(v);
  w = utilde^alpha + vtilde^alpha;

  beta = 1./alpha;
  cdf = exp( - (w^beta) );
  pdf = ((utilde.*vtilde)^(alpha-1)) .* ((w^beta) + alpha - 1) ./
        (w^(2-beta)) ./ (u .* v);

  retp(pdf .* cdf);
endp;

```

Exercice 27 Comparer les lignes de niveau des densités des distributions avec des marginales uniformes et gaussiennes. Les paramètres des copules Normale et Gumbel correspondront à un tau de Kendall égal à 0.5.

```

new;
library pgraph;

#include copula.src;

u = 0.001|seqa(0.01,0.01,99)|0.999; v = u';
x = seqa(-3,0.01,601); y = x';

tau = 0.5;                /* Kendall's tau */
rho = sin(pi/2*tau);      /* Corresponding parameter for the Normal copula */
alpha = 1/(1-tau);        /* Corresponding parameter for the Gumbel copula */

pdf1 = pdfNormal2Copula(u,v,rho);
pdf2 = pdfNormal2Copula(cdfn(x),cdfn(y),rho) .* pdfn(x) .* pdfn(y);

pdf3 = pdfGumbelCopula(u,v,alpha);
pdf4 = pdfGumbelCopula(cdfn(x),cdfn(y),alpha) .* pdfn(x) .* pdfn(y);

graphset;
  begwind;
  window(2,2,0);
  _pnum = 2; _paxht = 0.25; _pnumht = 0.20; _pframe = 0;
  fonts('simplex simgrma'); _ptitlht = 0.25;

  _plev = seqa(0.5,0.25,10);
  xtics(0,1,0.2,2);
  ytics(0,1,0.2,2);
  xlabel('U@[0,1@]');
  ylabel('U@[0,1@]');

  setwind(1);
  title('Normal Copula');
  contour(u',v',pdf1');

  setwind(3);
  title('Gumbel Copula');
  contour(u',v',pdf3');

  _plev = seqa(0.02,0.03,6);

```

```

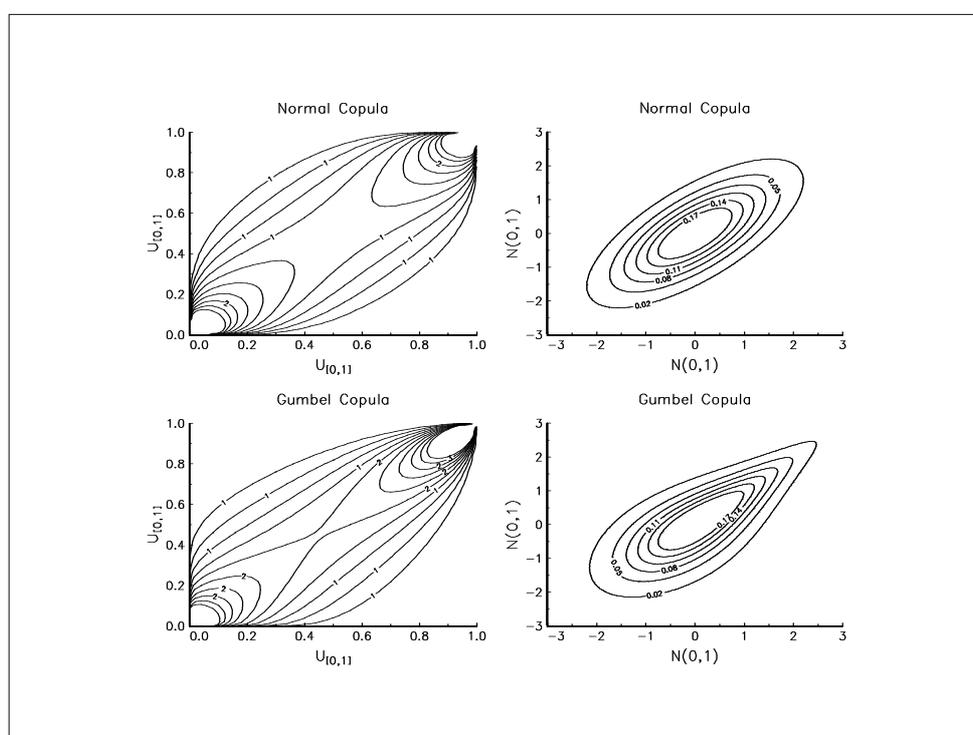
xtics(-3,3,1,2);
ytics(-3,3,1,2);
xlabel('N(0,1)');
ylabel('N(0,1)');

setwind(2);
title('Normal Copula');
contour(x',y',pdf2');

setwind(4);
title('Gumbel Copula');
contour(x',y',pdf4');

graphprt('-c=1 -cf=copula1.ps');
endwind;

```



Graphique 3.1. Lignes de niveau des densités

Exercice 28 *Ecrire une procédure qui simule des nombres aléatoires issus de la copule Normale.*

```

proc (2) = rndNormal2Copula(r, c, rho);
local n1, n2, u1, u2;

n1 = rndn(r, c);
n2 = rho .* n1 + sqrt(1-rho^2) .* rndn(r, c);

u1 = cdfn(n1);
u2 = cdfn(n2);

retp(u1, u2);
endp;

```

Exercice 29 Calculer la distribution conditionnelle $\Pr\{U_2 \leq u_2 \mid U_1 = u_1\}$ lorsque la copule de (U_1, U_2) est Gumbel. En déduire un algorithme de simulation des nombres aléatoires pour la copule Gumbel.

```

declare matrix _copula_u1;
declare matrix _copula_w;
declare matrix _copula_alpha;
declare matrix _FindZero_tol = 1e-5;

proc (2) = rndGumbelCopula(r,c,alpha);
  local Ns,u1,u2,w;

  Ns = r*c;
  alpha = alpha .* ones(r,c);

  u1 = rndu(Ns,1);
  w = rndu(Ns,1);

  _copula_alpha = vecr(alpha);
  _copula_u1 = u1;
  _copula_w = w;

  u2 = FindZero(&_rndGumbelCopula,0+__macheps,1-__macheps);

  u1 = reshape(u1,r,c);
  u2 = reshape(u2,r,c);

  retp(u1,u2);
endp;

proc _rndGumbelCopula(u2);
  retp( _cdfGumbelCopula_1(_copula_u1,u2,_copula_alpha) - _copula_w );
endp;

proc _cdfGumbelCopula_1(u1,u2,alpha);
  local u1tilde,u2tilde,cdf,beta,w;

  u1tilde = -ln(u1);
  u2tilde = -ln(u2);
  w = u1tilde^alpha + u2tilde^alpha;

  beta = 1./alpha;
  cdf = exp(-(w^beta)) .* ( 1+ (u2tilde./u1tilde)^alpha )^(-1+beta) ./ u1;

  retp(cdf);
endp;

proc FindZero(f,a,b);
  local f:proc;
  local ya,yb,Nobs,c,yc,indx1,indx2;
  local indx,s,diff,const;

  ya = f(a); yb = f(b);
  Nobs = rows(ya);
  a = a .* ones(Nobs,1);
  b = b .* ones(Nobs,1);

  indx = (ya .< 0) .and (yb .> 0);
  if sumc(indx) == 0;
    retp(miss(zeros(Nobs,1),0));
  endif;

  if sumc(indx) == Nobs;

```

```

do while maxc(abs(a-b)) > _FindZero_tol;
  c = (a+b)/2;
  yc = f(c);
  indx1 = yc.<0;
  indx2 = 1 - indx1;
  a = indx1.*c + indx2.*a;
  b = indx1.*b + indx2.*c;
enddo;

else;

s = delif(seqa(1,1,Nobs),indx);
diff = selif(a-b,indx);
const = miss(zeros(Nobs-sumc(indx),1),0);

do while maxc(abs(diff)) > _FindZero_Tol;
  c = (a+b)/2;
  c[s] = const;
  yc = f(c);
  indx1 = yc.<0;
  indx2 = 1 - indx1;
  a = indx1.*c + indx2.*a;
  b = indx1.*b + indx2.*c;
  diff = selif(a-b,indx);
enddo;

endif;

c = (a+b)/2;

retp(c);
endp;

```

Exercice 30 *Simuler des nombres aléatoires avec les copules Normale et Gumbel et avec des marginales t_3 et t_7 . Considérer à chaque fois une structure de dépendance faible — $\tau = 0.20$ — et une structure de dépendance forte — $\tau = 0.80$.*

```

new;
library pgraph;

#include copula.src;

rndseed 123456;

tau1 = 0.20; /* small dependence */
tau2 = 0.80; /* big dependence */
tau = tau1~tau2;

Ns = 500; Nr = 2;

rho = sin(pi/2*tau); /* Corresponding parameter for the Normal copula */
alpha = 1./(1-tau); /* Corresponding parameter for the Gumbel copula */

{u1,u2} = rndNormal2Copula(Ns,Nr,rho);
x1 = cdftci(1-u1,3);
x2 = cdftci(1-u2,7);

{u1,u2} = rndGumbelCopula(Ns,Nr,alpha);
x3 = cdftci(1-u1,3);
x4 = cdftci(1-u2,7);

```

```

graphset;
begwind;
window(2,2,0);

_pdate = ''; _pnum = 2; _pframe = 0;
_paxht = 0.20; _ptitlht = 0.20;
_plctrl = -1; _pcross = 1;
fonts('simplex simgrma');
xlabel('\214t]3['');
ylabel('\214t]7['');
xtics(-8,8,2,0);
ytics(-6,6,2,0);

setwind(1);
title('Normal copula with \202t\201 = 0.20');
xy(x1[.,1],x2[.,1]);

setwind(2);
title('Normal copula with \202t\201 = 0.80');
xy(x1[.,2],x2[.,2]);

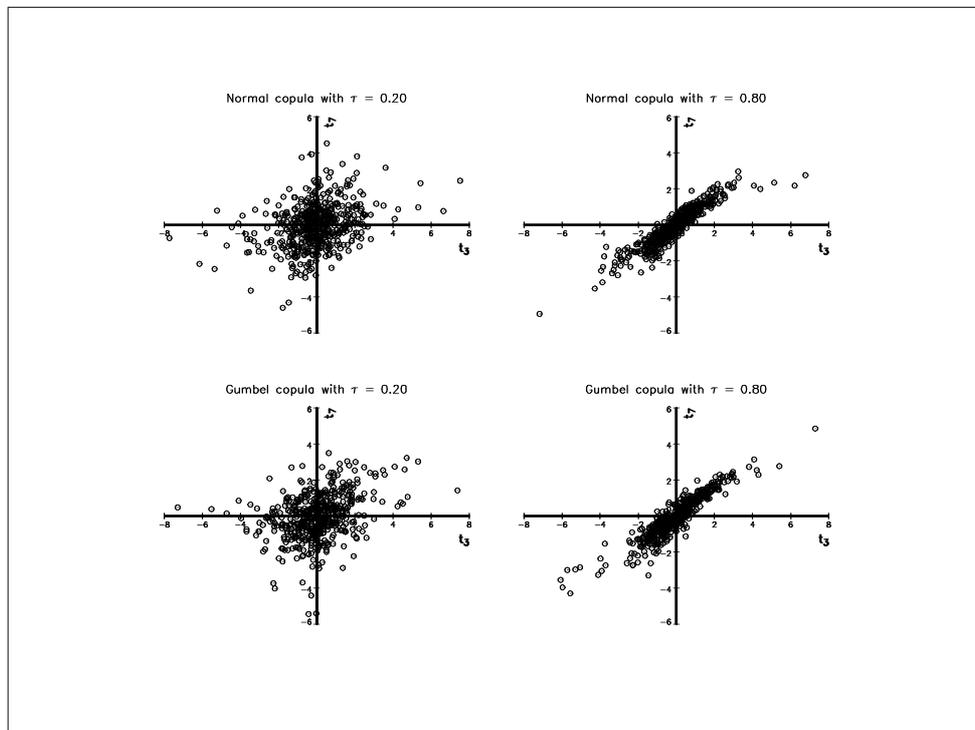
setwind(3);
title('Gumbel copula with \202t\201 = 0.20');
xy(x3[.,1],x4[.,1]);

setwind(4);
title('Gumbel copula with \202t\201 = 0.80');
xy(x3[.,2],x4[.,2]);

graphprt('-c=1 -cf=copula2.ps');

endwind;

```



Graphique 3.2. Simulation de nombres aléatoires

3.2 Distributions bidimensionnelles des extrêmes

3.2.1 Max-domaine d'attraction

Considérons un vecteur aléatoire (X, Y) de fonction de distribution \mathbf{F} et de marginales \mathbf{F}_1 et \mathbf{F}_2 . Notons $X_N^+ = \max(X_1, \dots, X_N)$.

Définition 2 On dit que \mathbf{F} appartient au max-domaine d'attraction d'une distribution à valeurs extrêmes \mathbf{G} avec des lois marginales non dégénérées \mathbf{G}_1 et \mathbf{G}_2 s'il existe des suites $a_N^1 > 0$, $a_N^2 > 0$, b_N^1 et b_N^2 telles que

$$\lim_{N \rightarrow \infty} \Pr \left(\frac{X_N^+ - b_N^1}{a_N^1} \leq x, \frac{Y_N^+ - b_N^2}{a_N^2} \leq y \right) = \lim_{N \rightarrow \infty} [\mathbf{F}(a_N^1 x + b_N^1, a_N^2 x + b_N^2)]^N = \mathbf{G}(x, y) \quad (3.1)$$

Exercice 31 Montrer que les constantes sont nécessairement celles données par le théorème de Fisher-Tippet, et que les marginales de \mathbf{G} sont les distributions unidimensionnelles des valeurs extrêmes de X et Y . Reformuler alors la définition précédente en terme de copules.

Théorème 2 Soit \mathbf{C}^* la copule correspondant à \mathbf{G} et soient \mathbf{G}_1 et \mathbf{G}_2 les marginales de \mathbf{G} . \mathbf{F} appartient au max-domaine d'attraction de \mathbf{G} si et seulement si \mathbf{F}_1 et \mathbf{F}_2 appartiennent aux max-domaines d'attraction de \mathbf{G}_1 et \mathbf{G}_2 respectivement et si \mathbf{C} satisfait

$$\mathbf{C}^*(u_1, u_2) = \lim_{N \rightarrow \infty} \mathbf{C}(u_1^{1/N}, u_2^{1/N})^N \quad (3.2)$$

3.2.2 Première caractérisation des copules à valeurs extrêmes

Théorème 3 $\mathbf{C}(u_1, u_2)$ est une copule à valeurs extrêmes si et seulement si

$$\mathbf{C}(u_1^t, u_2^t) = \mathbf{C}^t(u_1, u_2) \quad (3.3)$$

Exercice 32 Montrer que la copule Gumbel est une copule à valeurs extrêmes.

Exercice 33 Montrer que la copule Marshall-Olkin est une copule à valeurs extrêmes.

Exercice 34 Montrer que la copule Frank n'est pas une copule à valeurs extrêmes.

Exercice 35 Montrer que la copule \mathbf{C}^- n'est pas une copule à valeurs extrêmes.

Proposition 1 Pour toute copule \mathbf{C} à valeurs extrêmes, on a

$$\mathbf{C}^\perp \prec \mathbf{C} \prec \mathbf{C}^+ \quad (3.4)$$

Cela implique que les extrêmes ne présentent jamais de dépendance négative !

3.2.3 Seconde caractérisation des copules à valeurs extrêmes

Théorème 4 Soit \mathbf{C} une copule à valeurs extrêmes. Alors il existe une fonction convexe A définie sur $[0, 1]$ et qui satisfait la condition $\max(u, 1 - u) \leq A(u) \leq 1$ telle que

$$\mathbf{C}(u_1, u_2) = \exp \left[\ln(u_1 u_2) A \left(\frac{\ln(u_2)}{\ln(u_1 u_2)} \right) \right] \quad (3.5)$$

Exercice 36 Vérifier que la représentation précédente définit bien une copule à valeurs extrêmes.

Exercice 37 Trouver la fonction A des copules \mathbf{C}^+ , \mathbf{C}^\perp , Gumbel et Marshall-Olkin.

Exercice 38 Montrer que la copule \mathbf{C} appartient au max-domaine d'attraction de \mathbf{C}^* caractérisée par la fonction A si et seulement si

$$\lim_{u \rightarrow 0} \frac{1 - \mathbf{C}\left((1-u)^{1-t}, (1-u)^t\right)}{u} = A(t) \quad \forall t \in [0, 1] \quad (3.6)$$

En déduire que la copule *Kimeldorf-Sampson* (aussi appelée *Clayton* ou *Cook-Johnson*) appartient au max-domaine d'attraction de \mathbf{C}^\perp .

3.2.4 Le coefficient de dépendance de queue

Définition 3 Soient X_1 et X_2 deux variables aléatoires de fonction de distribution \mathbf{F}_1 et \mathbf{F}_2 respectivement. Le coefficient de **dépendance de queue supérieure** de X_1 et X_2 est défini par

$$\lambda = \lim_{u \rightarrow 1^-} \Pr \{X_2 \geq \mathbf{F}_2^{-1}(u) \mid X_1 > \mathbf{F}_1^{-1}(u)\} \quad (3.7)$$

si la limite $\lambda \in [0, 1]$ existe. Si $\lambda \in (0, 1]$, les v.a. X_1 et X_2 sont dites *asymptotiquement dépendantes au niveau supérieur de la queue de distribution*. Si $\lambda = 0$, les v.a. X_1 et X_2 sont dites *asymptotiquement indépendantes au niveau supérieur de la queue de distribution*.

Remarque 11 Si $\lambda > 0$, cela signifie qu'il y a une probabilité non nulle que l'une des variables dépasse u sachant que l'autre est déjà supérieure à u . **Dit autrement, λ est la probabilité d'avoir un extrême sachant qu'on a déjà un extrême.**

Exercice 39 Définissons la fonction $\bar{\mathbf{C}}$ par

$$\bar{\mathbf{C}}(u_1, u_2) = 1 - u_1 - u_2 + \mathbf{C}(u_1, u_2) \quad (3.8)$$

Montrer alors que

$$\lambda = \lim_{u \rightarrow 1^-} \frac{\bar{\mathbf{C}}(u, u)}{1 - u}$$

avec \mathbf{C} la copule des variables aléatoires X_1 et X_2 .

Exercice 40 Montrer que $\lambda = 0$ pour la copule Normale. En déduire que la copule Normale appartient au max-domaine d'attraction de \mathbf{C}^\perp .

Exercice 41 Montrer que $\lambda = 2 - 2^{1/\alpha}$ pour la copule Gumbel.

Exercice 42 Interpréter la mesure suivante dans une analyse de type Valeur en Risque :

$$\lambda(u) = \frac{\bar{\mathbf{C}}(u, u)}{1 - u}$$

Représenter graphiquement $\lambda(u)$ et λ pour les copules Normale et Gumbel et pour différentes valeurs du tau de Kendall.

```
new;
library pgraph;

#include copula.src;

u = 0.001|seqa(0.01,0.01,99)|0.999|0.9999|0.99999|0.999999|0.99999999;

tau = 0.25~0.50~0.75~0.90;
rho = sin(pi/2*tau);      /* Corresponding parameter for the Normal copula */
```

```

alpha = 1./(1-tau);          /* Corresponding parameter for the Gumbel copula */

cdf1 = cdfNormal2Copula(u,u,rho);
cdf2 = cdfGumbelCopula(u,u,alpha);

l1 = (1 - 2.*u + cdf1) ./ (1-u);
l2 = (1 - 2.*u + cdf2) ./ (1-u);

lambda1 = 0 .* rho;
lambda2 = 2 - 2^(1./alpha);

u = u | 1;
l1 = l1 | lambda1;
l2 = l2 | lambda2;

graphset;
begwind;
window(2,2,0);

_pnum = 2; _pltype = 6|1; _pframe = 0;
_paxht = 0.25; _ptitlht = 0.25; _pnumht = 0.20;
fonts('simplex simgrma');
xtics(0,1,0.2,2);
ytics(0,1,0.2,2);
xlabel('u');
ylabel('\2021\201(u)');

i = 1;
setwind(i);
title(ftos(tau[i], '\202t\201 = %lf', 3, 2));
e = ones(2,1);
_pline = e.*(1~6~0.97)^(lambda1[i]|lambda2[i])~e~
(lambda1[i]|lambda2[i])~e.*(1~15~15);
xy(u, l1[.,i]~l2[.,i]);

i = 2;
setwind(i);
title(ftos(tau[i], '\202t\201 = %lf', 3, 2));
_pline = e.*(1~6~0.97)^(lambda1[i]|lambda2[i])~e~
(lambda1[i]|lambda2[i])~e.*(1~15~15);
xy(u, l1[.,i]~l2[.,i]);

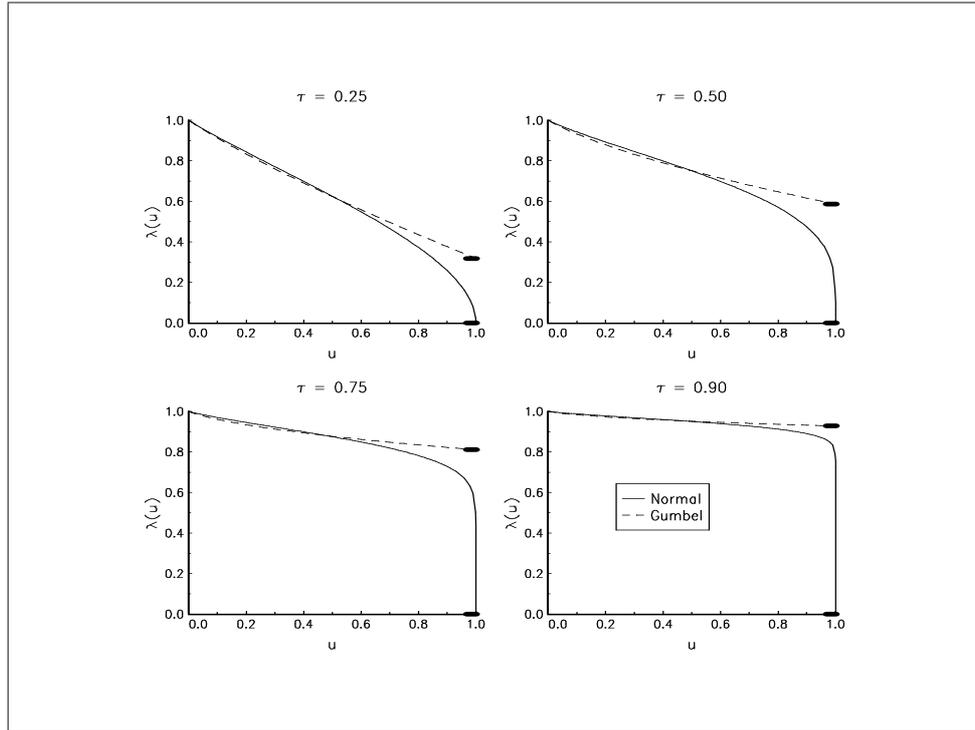
i = 3;
setwind(i);
title(ftos(tau[i], '\202t\201 = %lf', 3, 2));
_pline = e.*(1~6~0.97)^(lambda1[i]|lambda2[i])~e~
(lambda1[i]|lambda2[i])~e.*(1~15~15);
xy(u, l1[.,i]~l2[.,i]);

i = 4;
setwind(i);
title(ftos(tau[i], '\202t\201 = %lf', 3, 2));
_pline = e.*(1~6~0.97)^(lambda1[i]|lambda2[i])~e~
(lambda1[i]|lambda2[i])~e.*(1~15~15);
_plegstr = 'Normal\000Gumbel';
_plegctl = {2 6 3 3};
xy(u, l1[.,i]~l2[.,i]);

graphprt('-c=1 -cf=mev1.ps');

endwind;

```



Graphique 3.3. Comparaison des dépendances de queue des copules Normale et Gumbel

3.3 Estimation par la méthode du maximum de vraisemblance

La fonction de log-vraisemblance individuelle de la copule Gumbel est

$$\begin{aligned} \ell(u_1, u_2; \alpha) = & -(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{\alpha}} + \tilde{u}_1 + \tilde{u}_2 + (\alpha - 1) \ln(\tilde{u}_1 \tilde{u}_2) + (\alpha^{-1} - 2) \ln(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha) \\ & + \ln \left[(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{\alpha}} + \alpha - 1 \right] \end{aligned} \quad (3.9)$$

avec $\tilde{u}_1 = -\ln u_1$.

Exercice 43 *Écrire un programme qui estime le paramètre α de la copule Gumbel par maximum de vraisemblance.*

```

declare matrix _mleGumbel_data;
declare matrix _mleGumbel_algr = 1;

proc (3) = mleGumbel(u1,u2,sv);
  local alpha,logl,g,retcode,stderr;

  _mleGumbel_data = u1~u2;

  if _mleGumbel_algr == 1;

    {alpha,LogL,g,retcode} = optnum(&_mleGumbel_fn,sv-1);

  else;

    {alpha,LogL,g,retcode} = Qnewton(&_mleGumbel_fn,sv-1);

  endif;

  alpha = 1 + sqrt(alpha^2);

```

```

logl = - logl;
stderr = 1/hessp(_mleGumbel_fn,alpha);

retp(alpha,stderr,logl);
endp;

proc _mleGumbel_fn(alpha);
local logl;

alpha = 1 + sqrt(alpha^2); /* alpha > 1 */

logl = ln(pdfGumbelCopula(_mleGumbel_data[.,1],_mleGumbel_data[.,2],alpha));

retp( -sumc(packr(logl)) );
endp;

```

Exercice 44 *Illustrer la convergence de l'estimateur ML par Monte Carlo.*

```

new;
library optmum,pgraph;

#include copula.src;
#include gumbel.src;

__output = 0;

alpha = 1.75;
sv = 1.75;
Nr = 250;
theta = zeros(Nr,1);

Nobs = 100;
{u1,u2} = rndGumbelCopula(Nobs,Nr,alpha);
i = 1;
do until i > Nr;
  {theta[i],stderr,Logl} = mleGumbelCopula(u1[.,i],u2[.,i],sv);
  i = i + 1;
endo;
{est1,dens1} = Kernel(theta,1,3,256);

Nobs = 200;
{u1,u2} = rndGumbelCopula(Nobs,Nr,alpha);
i = 1;
do until i > Nr;
  {theta[i],stderr,Logl} = mleGumbelCopula(u1[.,i],u2[.,i],sv);
  i = i + 1;
endo;
{est2,dens2} = Kernel(theta,1,3,256);

Nobs = 300;
{u1,u2} = rndGumbelCopula(Nobs,Nr,alpha);
i = 1;
do until i > Nr;
  {theta[i],stderr,Logl} = mleGumbelCopula(u1[.,i],u2[.,i],sv);
  i = i + 1;
endo;
{est3,dens3} = Kernel(theta,1,3,256);

graphset;
_pframe = 0; _pdate = '''; _pnum = 2;
_pltype = 6|1|3;
xtics(1.25,2.25,0.25,2);
ytics(0,5,1,2);

```

```

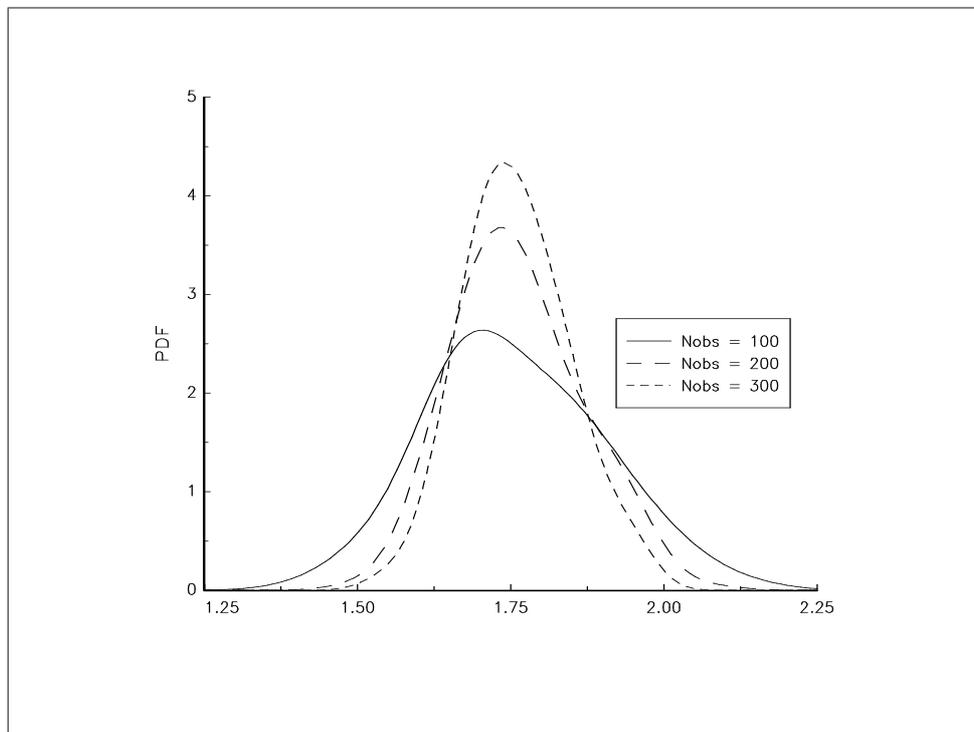
xlabel('\202a\201[ML]');
ylabel('PDF');
_plegstr = ''Nobs = 100\000Nobs = 200\000Nobs = 300'';
_plegct1 = {2 5 6 3};
graphprt('-c=1 -cf=gumbel1.ps');
xy(est1~est2~est3,dens1~dens2~dens3);

proc (2) = Kernel(data,xmin,xmax,Nx);
local N,s,h,x,d;

N = rows(data);
s = stdc(data);
h = 1.364*s*(N^(-0.2));
x = seqa(xmin,(xmax-xmin)/(Nx-1),Nx);
d = meanc(pdfn((x'-data)/h)/h);

retp(x,d);
endp;

```



Graphique 3.4. Convergence de l'estimateur ML du paramètre de la copule Gumbel

3.4 Application au stress-testing

Les scénarios bivariés de crise peuvent être vus comme la représentation d'une zone de défaillance. Nous avons

$$\begin{aligned}
\Pr(\chi_1^+ > \chi_1, \chi_2^+ > \chi_2) &= 1 - \Pr(\chi_1^+ \leq \chi_1) - \Pr(\chi_2^+ \leq \chi_2) + \Pr(\chi_1^+ \leq \chi_1, \chi_2^+ \leq \chi_2) \\
&= 1 - \mathbf{F}_1(\chi_1) - \mathbf{F}_2(\chi_2) + \mathbf{C}(\mathbf{F}_1(\chi_1), \mathbf{F}_2(\chi_2)) \\
&= \overline{\mathbf{C}}(\mathbf{F}_1(\chi_1), \mathbf{F}_2(\chi_2))
\end{aligned} \tag{3.10}$$

Soit τ le temps de retour. La zone de défaillance est représentée par

$$\left\{ (x_1, x_2) \in \mathbb{R}^2 \mid \overline{\mathbf{C}}(\mathbf{F}_1(x_1), \mathbf{F}_2(x_2)) < \frac{1}{\tau} \right\} \quad (3.11)$$

Exercice 45 Construire la zone de défaillance du (CAC40, Dow Jones) pour les minimums et pour un temps de retour égal à 5 ans. Comparer les résultats obtenus avec les copules Gumbel, \mathbf{C}^+ et \mathbf{C}^\perp .

```

new;
library optnum, pgraph;

rndseed 1234565;

#include gev.src;
#include mlegev.src;
#include copula.src;
#include gumbel.src;

data = loadd('cac-dj');

CAC40 = data[,2];      /* CAC 40 */
DJ = data[,3];        /* Dow Jones */

y = packr(CAC40~DJ);
data = ln(y);
r = data - lag1(data);
Nobs = rows(r);

varphi = 22;          /* one month of trading */
tau = 5;              /* Return time in years */

block_CAC40 = reshape(r[,1], Nobs/varphi, varphi)';
block_DJ = reshape(r[,2], Nobs/varphi, varphi)';

min_CAC40 = -minc(block_CAC40);
min_DJ = -minc(block_DJ);
max_CAC40 = maxc(block_CAC40);
max_DJ = maxc(block_DJ);

/*
**> Modelisation of the dependence of Min/Min
**
*/

x1 = min_CAC40;
x2 = min_DJ;

{mu, sigma, xi} = mleGEV(x1, 0);
theta1 = mu|sigma|xi;
u1 = cdfGEV(x1, mu, sigma, xi);
{mu, sigma, xi} = mleGEV(x2, 0);
theta2 = mu|sigma|xi;
u2 = cdfGEV(x2, mu, sigma, xi);
{alpha, stderr, logl} = mleGumbelCopula(u1, u2, 1.25);

x1 = seqa(0, 0.15/200, 201);
x2 = x1';
u1 = cdfGEV(x1, theta1[1], theta1[2], theta1[3]);
u2 = cdfGEV(x2, theta2[1], theta2[2], theta2[3]);

cdf = cdfGumbelCopula(u1, u2, alpha);
tau = 265 * tau / varphi;          /* return time in period of block */

```

```

stress1 = 100*cdfGEVi(1-1/tau,theta1[1],theta1[2],theta1[3]); /* Univariate stress testing */
stress2 = 100*cdfGEVi(1-1/tau,theta2[1],theta2[2],theta2[3]); /* Univariate stress testing */

e = (1 - u1 - u2 + cdf) .< (1/tau); /* failure area with Gumbel Copula */

u2 = (1 - u1 - 1/tau)/(1-u1);
u2 = missex(u2,u2 .<= 0);

y1 = 100*cdfGEVi(u1,theta1[1],theta1[2],theta1[3]);
y2 = 100*cdfGEVi(u2,theta2[1],theta2[2],theta2[3]);

graphset;
  _pdate = '''; _pnum = 2; _pframe = 0;
  _pcross = 1;
  fonts('simplex simgrma');
  xtics(0,15,5,5);
  ytics(0,15,5,5);
  xlabel('\215CAC40 (in %)'); ylabel('\215DowJones (in %)');
  {row,col} = SpyPlot(e,0);
  _psym = 100*(x1[row]~x2[col])~ones(rows(row),1).*(4~0.05~15~1~2);
  _pline = 1~6~stress1~0~stress1~15~1~11~5 |
           1~6~0~stress2~15~stress2~1~11~5 ;
  _pmsgstr = 'C[+]\000C[\202\74\201]\000Gumbel';
  _pmsgctl = 10.1~10.25~0.15~0~1~12~5 |
             4.1~4.1~0.15~0~1~12~5 |
             6.25~8.25~0.15~-38~1~12~5 ;
  graphprt('-c=1 -cf=stress1.ps');

xy(y1,y2);

proc (2) = spyPlot(x,eps);
  local r_x,c_x,r,c,y;

  if eps == 0;
    eps = __macheps;
  endif;
  eps = abs(eps);

  r_x = rows(x);
  c_x = cols(x);

  r = vecr(seqa(1,1,r_x) .* ones(r_x,c_x));
  c = vecr(seqa(1,1,c_x)' .* ones(r_x,c_x));
  x = vecr(x);

  y = selif(r~c~x, x .> eps .or x .< -eps );

  r = y[. ,1];
  c = y[. ,2];

  retp(r,c);
endp;

```

Exercice 46 *Calculer le temps de retour implicite d'une baisse simultanée de 10% du CAC40 et 10% du DowJones.*

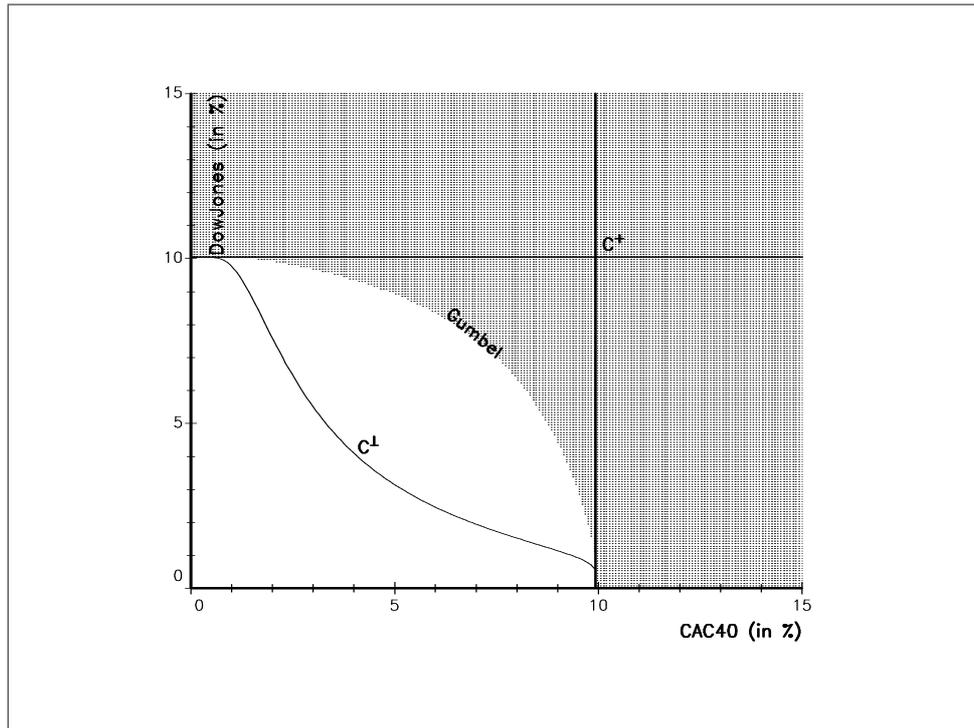
```

new;
library optmum,pgraph;

rndseed 1234565;

#include gev.src;

```



Graphique 3.5. Zone de défaillance à 5 ans

```

#include mlev. src;
#include copula. src;
#include gumbel. src;

data = load('cac-dj');

CAC40 = data[:,2];      /* CAC 40 */
DJ = data[:,3];        /* Dow Jones */

y = packr(CAC40~DJ);
data = ln(y);
r = data - lag1(data);
Nobs = rows(r);

varphi = 22;           /* one month of trading */

block_CAC40 = reshape(r[:,1],Nobs/varphi,varphi)';
block_DJ = reshape(r[:,2],Nobs/varphi,varphi)';

min_CAC40 = -minc(block_CAC40);
min_DJ = -minc(block_DJ);
max_CAC40 = maxc(block_CAC40);
max_DJ = maxc(block_DJ);

/*
**> Modelisation of the dependence of Min/Min
**
*/

x1 = min_CAC40;
x2 = min_DJ;

{mu,sigma,xi} = mleGEV(x1,0);
theta1 = mu|sigma|xi;
u1 = cdfGEV(x1,mu,sigma,xi);
{mu,sigma,xi} = mleGEV(x2,0);

```

```

theta2 = mu|sigma|xi;
u2 = cdfGEV(x2,mu,sigma,xi);
{alpha,stderr,logl} = mleGumbelCopula(u1,u2,1.25);

x1 = 0.10;
x2 = 0.10;
u1 = cdfGEV(x1,theta1[1],theta1[2],theta1[3]);
u2 = cdfGEV(x2,theta2[1],theta2[2],theta2[3]);

C = cdfGumbelCopula(u1,u2,alpha);
tau1 = 1./(1 - u1 - u2 + C)*varphi/265;

C = u1.*u2;
tau2 = 1./(1 - u1 - u2 + C)*varphi/265;

e = u1 .> u2;
C = e .* u2 + (1-e) .* u1;
tau3 = 1./(1 - u1 - u2 + C)*varphi/265;

cls;

print tau1;
print tau2;
print tau3;

```

3.5 Complément bibliographique

Les chercheurs citent souvent [Pickands] comme étant l'article de référence sur la théorie multidimensionnelle des extrêmes. Je vous conseille plutôt l'article de DEHEUVELS [1978]. Dans le cas bivarié, [GKR] et [Tawn 88] sont des lectures plaisantes et abordables. Concernant les domaines d'attraction, vous trouverez des éléments dans [CFG 00], [MO] et [Yun]. A noter aussi qu'il existe d'autres méthodes pour construire des scénarios de crise multidimensionnels (voir LEGRAS et SOUPÉ [2000]).

- [CFG97] CAPÉRAÀ, P., A-L. FOUGÈRES et C. GENEST [1997], Estimation of bivariate extreme value copulas, *Biometrika*, **84**, 567-577
- [CFG 00] CAPÉRAÀ, P., A-L. FOUGÈRES et C. GENEST [2000], Bivariate distributions with given extreme value attractor, *Journal of Multivariate Analysis*, **72**, 30-49
- [CCT] COLES, S., J. CURRIE et J. TAWN [1999], Dependence measures for extreme value analyses, Department of Mathematics and Statistics, Lancaster University, *Working Paper*
- [Deheuvels 78] DEHEUVELS, P. [1978], Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes, *Publications de l'Institut de Statistique de l'Université de Paris*, **23**, 1-36
- [GKR] GHOUDI, K., A. KHOUDRAJI et L.P. RIVEST [1998], Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles, *Canadian Journal of Statistics*, **26**, 187-197
- [LS] LEGRAS, J. et F. SOUPÉ [2000], Designing multivariate stress scenarios : the extreme value approach, CCF, *Research and Innovation Notes*, march
- [MO] MARSHALL, A.W. et I. OLKIN [1983], Domains of attraction of multivariate extreme value distributions, *Annals of Probability*, **11**, 168-177
- [Pickands] PICKANDS, J. [1981], Multivariate extreme value distributions, *Bull. Int. Statist. Inst.*, **49**, 859-878

- [Tawn 88] TAWN, J.A. [1988], Bivariate extreme value theory : models and estimation, *Biometrika*, **75**, 397-415
- [Tawn 90] TAWN, J.A. [1990], Modelling multivariate extreme value distributions, *Biometrika*, **77**, 245-253
- [Yun] YUN, S. [1997], On domains of attraction of multivariate extreme value distribution under absolute continuity, *Journal of Multivariate Analysis*, **63**, 277-295

Conclusion

Il est difficile en 5 heures de cours d'aborder les différentes méthodes relatives à la théorie des extrêmes et l'ensemble des applications possibles en finance. En particulier, je n'ai pas abordé les approches estimateur de Hill, distribution généralisée de Pareto, processus points, etc. Et d'autres applications financières sont possible. Voici donc de nouvelles références qui peuvent vous intéresser.

- [JSW] JOE, H., R.L. SMITH et I. WEISSMANN [1992], Bivariate threshold methods for extremes, *Journal of the royal statistical society — Series B : methodological*, **54**, 171-183
- [Legras] LEGRAS, J. [1999], The 1998 financial crisis : how serious was it ?, Crédit Commercial de France, *Working Paper*
- [LS 95] LONGIN, F. et B. SOLNIK [1995], Is the correlation in international equity returns constant : 1960-1990 ?, *Journal of International Money and Finance*, **14**, 3-26
- [LS 00] LONGIN, F. et B. SOLNIK [2000], Correlation structure of international equity markets during extremely volatile periods, CEPR, *Discussion Paper*, **2538**
- [Starica] STĂRICĂ, C. [1999], Multivariate extremes for models with constant conditional correlations, *Journal of Empirical Finance*, **6**, 515-553
- [Straetmans] STRAETMANS, S. [1999], Extreme financial returns and their comovements, Erasmus University Rotterdam's Thesis, *Tinbergen Institute Research Series*, **181**