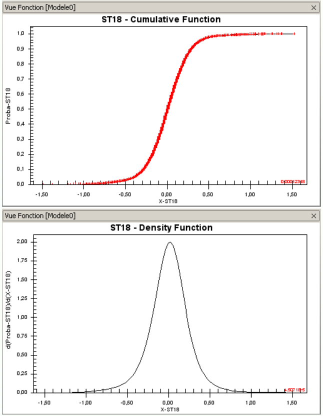
Distributions à queues épaisses et package FatTailsR

patrice.kiener@inmodelia.com

MNHN Semin-R – 16 juin 2017

Je m'intéresse aux distributions à queues épaisses à gauche ET à droite.



⇒ Quelle loi de probabilité ?

Queues épaisses : un exemple asymétrique

Applications : Marchés financiers, Neurosciences, Signaux antennes–satellites, etc...

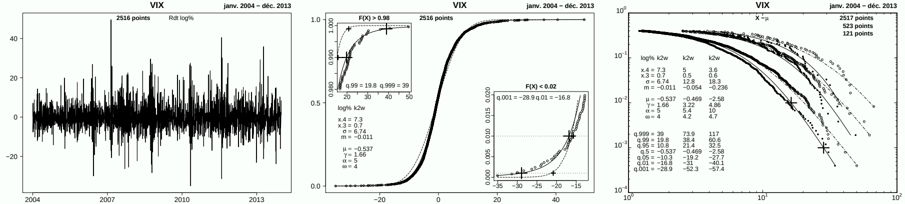
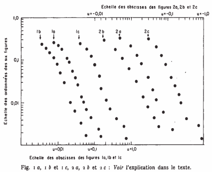


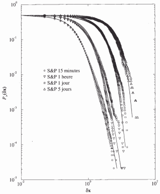
FIGURE – Indice VIX de janvier 2004 à décembre 2013 :
2517 valeurs journalières, 523 valeurs hebdomadaires, 121 valeurs mensuelles.

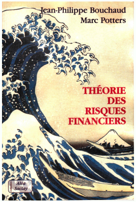
Le graphique log-log de droite met en évidence des fonctions en lois de puissance $|x|^{-\alpha}$.

Première loi de probabilité à queues épaisses (1962)

- **Mandelbrot (1962)** fut le premier à utiliser des lois à queues épaisses via une combinaison de deux lois de puissance unilatérales (en $|x|^{-\alpha}$) tronquées en $x = 0$.
- **Bouchaud et Potters (1997)** approfondissent le sujet.







$$\log\{\mathrm{Fr}[L(t, T) > u)]\} \sim -\alpha \log u + \log C'(T).$$
$$\log\{\mathrm{Fr}[L(t, T) < -u)]\} \sim -\alpha \log u + \log C''(T).$$

$$L_\mu(x) \simeq \frac{\mu A_\pm^\mu}{|x|^{1+\mu}} \text{ pour } x \to \pm\infty,$$

FIGURE – (a+b) Prix du coton jour-semaine-mois - (c+d) Indice SP500 15min-jour-semaine-mois

Lois de probabilité à queues épaisses (1972–2009)

Parmi les lois suivantes, seule la loi de Student décentrée est quelquefois utilisée.

(Elle a une singularité en $x = 0$)

Lambda généralisée (1972)	$x(p) = \lambda_1 + \frac{1}{\lambda_2}\left[-(1-p)^{\lambda_4} + p^{\lambda_3}\right]$	Inutilisable
L$_U$ (1982)	$x(p) = \xi + \lambda \sinh\left(-\frac{\gamma}{\delta} + \frac{1}{\delta}\,\text{logit}(p)\right)$	Inusitée
Student décentrée (1997, 2003)	$F_{v,\mu}(x) = \begin{cases} \frac{1}{2}\sum_{j=0}^{\infty}\frac{1}{j!}(-\mu\sqrt{2})^j e^{\frac{-\mu^2}{2}}\frac{\Gamma(\frac{j+1}{2})}{\Gamma(1/2)} I\left(\frac{v}{v+x^2};\frac{v}{2},\frac{j+1}{2}\right), & x \geq 0 \\ 1-\frac{1}{2}\sum_{j=0}^{\infty}\frac{1}{j!}(-\mu\sqrt{2})^j e^{\frac{-\mu^2}{2}}\frac{\Gamma(\frac{j+1}{2})}{\Gamma(1/2)} I\left(\frac{v}{v+x^2};\frac{v}{2},\frac{j+1}{2}\right), & x < 0 \end{cases}$	Correcte
Sinh-arcsinh (2009)	$z = \sinh\left(\epsilon + \delta\,\text{asinh}\left(\frac{x-\mu}{\sigma}\right)\right) \qquad f(x) = \frac{\delta}{\sqrt{2\pi}}\frac{\sqrt{1+z^2}}{\sqrt{1+x^2}}\exp\left(-\frac{1}{2}z^2\right)$	Gauss modifiée

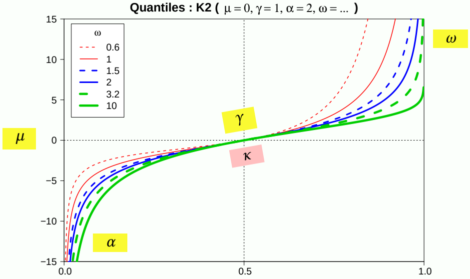
Loi de probabilité K2

En 2014 [1], j'ai proposé une nouvelle famille de 4 lois de probabilité à queues épaisses. Elles combinent deux fonctions log-logistique, une pour chaque queue. Ici, la loi K2 :

- médiane μ, échelle γ, queue gauche α, queue droite ω, queue moyenne $\frac{1}{\kappa} = \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\omega}\right)$

Fonction quantile de la loi K2

$$K2(\mu,\gamma,\alpha,\omega) \qquad x(p) = \mu + \gamma\kappa\left[-\left(\frac{p}{1-p}\right)^{-\frac{1}{\alpha}} + \left(\frac{p}{1-p}\right)^{\frac{1}{\omega}}\right] = \mu + \gamma\kappa\left(-e^{-\frac{\mathrm{logit}(p)}{\alpha}} + e^{\frac{\mathrm{logit}(p)}{\omega}}\right)$$



$$\frac{dx(p{=}0.5)}{dp} = 8\gamma$$

1. P. Kiener, Explicit models for bilateral fat-tailed distributions and applications in finance with the package FatTailsR, 8th R/Rmetrics Workshop, Paris, 2014. http://www.inmodelia.com/fattailsr-en.html

Lois de probabilité K1 (symétrique), K2, K3, K4 (asymétriques)

En pratique, 4 paramétrisations sont possibles : K1 (symétrique) et K2, K3, K4 (asymétriques)

- médiane μ, échelle γ, queue gauche α, queue droite ω
- queue moyenne $\frac{1}{\kappa} = \frac{1}{2}\left(\frac{1}{\alpha} + \frac{1}{\omega}\right)$, excentricité ϵ, distorsion $\delta = \frac{\epsilon}{\kappa} = \frac{1}{2}\left(-\frac{1}{\alpha} + \frac{1}{\omega}\right)$ avec $-1 < \epsilon < 1$
  et $\frac{1}{\alpha} = \frac{1}{\kappa} - \delta \quad \frac{1}{\omega} = \frac{1}{\kappa} + \delta \quad \epsilon = \frac{\alpha-\omega}{\alpha+\omega} \quad \alpha = \frac{\kappa}{1-\epsilon} \quad \omega = \frac{\kappa}{1+\epsilon}$

Fonctions quantiles des lois K2, K3, K4, K1

$$K2(\mu,\gamma,\alpha,\omega) \qquad x(p) = \mu + \gamma\kappa\left[-\left(\frac{p}{1-p}\right)^{-\frac{1}{\alpha}} + \left(\frac{p}{1-p}\right)^{\frac{1}{\omega}}\right] = \mu + \gamma\kappa\left(-e^{-\frac{\mathrm{logit}(p)}{\alpha}} + e^{\frac{\mathrm{logit}(p)}{\omega}}\right)$$

$$K3(\mu,\gamma,\kappa,\delta) \qquad x(p) = \mu + \gamma\kappa\left[-\left(\frac{p}{1-p}\right)^{-\frac{1}{\kappa}} + \left(\frac{p}{1-p}\right)^{\frac{1}{\kappa}}\right]\left(\frac{p}{1-p}\right)^{\delta} = \mu + 2\gamma\kappa\sinh\left(\frac{\mathrm{logit}(p)}{\kappa}\right)e^{\delta\,\mathrm{logit}(p)}$$

$$K4(\mu,\gamma,\kappa,\epsilon) \qquad x(p) = \mu + \gamma\kappa\left[-\left(\frac{p}{1-p}\right)^{-\frac{1}{\kappa}} + \left(\frac{p}{1-p}\right)^{\frac{1}{\kappa}}\right]\left(\frac{p}{1-p}\right)^{\frac{\epsilon}{\kappa}} = \mu + 2\gamma\kappa\sinh\left(\frac{\mathrm{logit}(p)}{\kappa}\right)e^{\frac{\epsilon}{\kappa}\,\mathrm{logit}(p)}$$

$$K1(\mu,\gamma,\kappa) \qquad x(p) = \mu + \gamma\kappa\left[-\left(\frac{p}{1-p}\right)^{-\frac{1}{\kappa}} + \left(\frac{p}{1-p}\right)^{\frac{1}{\kappa}}\right] = \mu + 2\gamma\kappa\sinh\left(\frac{\mathrm{logit}(p)}{\kappa}\right) = \mu + \gamma\,\mathrm{kogit}(p,\kappa)$$

Lois K1, K2, K3, K4 : une grande variété de forme. Ici K4$(\mu, \gamma, \kappa, \epsilon)$

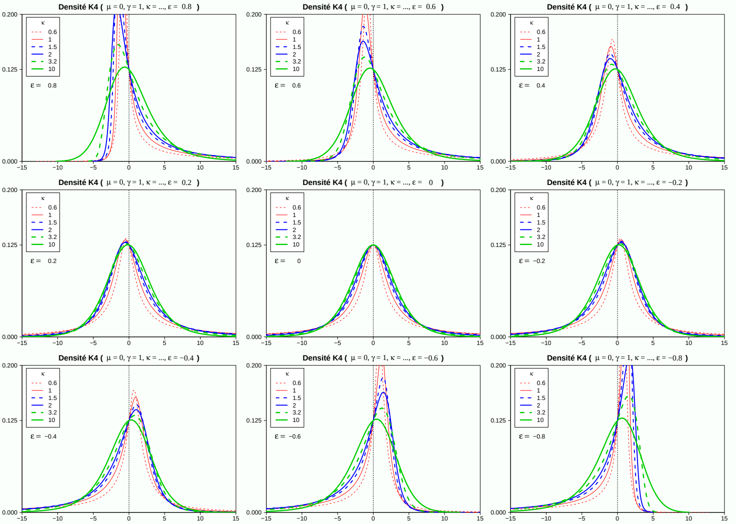


FIGURE – Fonctions de densité de la loi K4 pour diverses valeurs de κ et ϵ
(a) $\epsilon = 0.8$ (b) $\epsilon = 0.6$ (c) $\epsilon = 0.4$
(d) $\epsilon = 0.2$ (e) $\epsilon = 0$ (f) $\epsilon = -0.2$
(g) $\epsilon = -0.4$ (h) $\epsilon = -0.6$ (i) $\epsilon = -0.8$

Loi symétrique K1

Loi du quantile, fonction de répartition et densité de la loi K1 : $\quad t = \operatorname{asinh}\left(\frac{x-\mu}{2\gamma\kappa}\right)$

$$x(p) = \mu + 2\gamma\kappa \sinh\left(\frac{\operatorname{logit}(p)}{\kappa}\right) \qquad F(x) = \frac{1}{1+e^{-\kappa t}} \qquad f(x) = \frac{1}{4\gamma \cosh(t)\,(1+\cosh(\kappa t))}$$

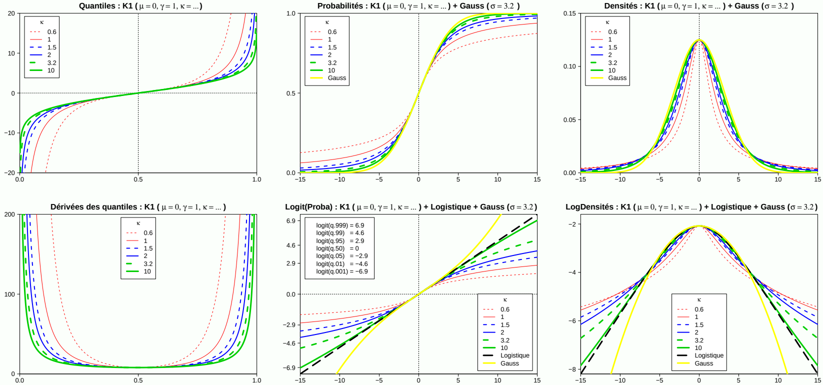
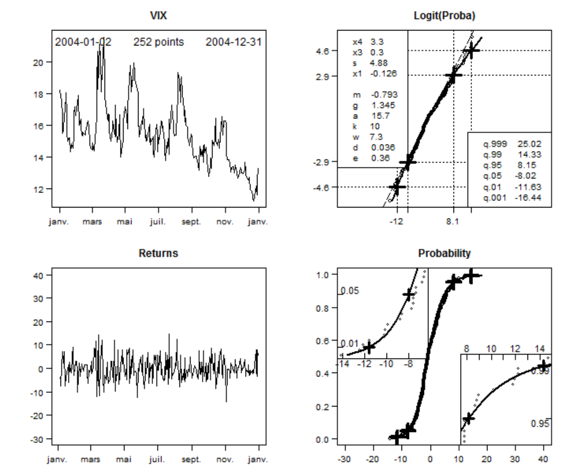


FIGURE – Loi de probabilité K1 : (a) Fonctions de répartition (b) Densités (c) Logit de la foncion de répartition - (d) Log-densités

Vidéo 1 : Fenêtres glissantes sur VIX

Fenêtres glissantes 252 jours (1 an) rafraichies chaque mois entre janvier 2004 et décembre 2013

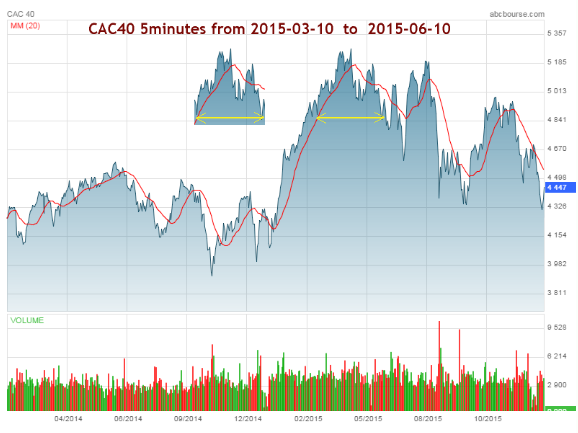


url : VIX VIX-Garch

Vidéo 2 : Fenêtres glissantes sur CAC40

CAC40 5 minutes pendant les journées du 10 mars au 10 juin 2015 :

(102 points x 64 fenêtres indépendantes)



url : CAC40-5min

Packages FatTailsR et FatTailsRplot

- FatTailsR (Mathématiques) ⇒ https://CRAN.R-project.org/package=FatTailsR
- FatTailsRplot (Graphiques) ⇒ package privé

Signets

FatTailsR-package
aw2k
checkcoefk
checkquantiles
ckiener1234
dfData
dimdim
elevate
elevenprobs
estimkiener11
exfit0
exphp
extractData
fatreturns
fitkienerX
getDSdata
getnamesk
kashp
kiener1
kiener2
kiener3
kiener4
kiener7
kmoments
laplacegaussnorm
loghp
logishp
logit
mData
pk2pk
pprobs0
regkienerLX
roundcoefk
tData
xData
zData
Index

Package 'FatTailsR'

May 16, 2017

Title Kiener Distributions and Fat Tails in Finance

Description Kiener distributions K1, K2, K3, K4 and K7 to characterize distributions with left and right, symmetric or asymmetric fat tails in market finance, neuroscience and other disciplines. Two algorithms to estimate with a high accuracy distribution parameters, quantiles, value-at-risk and expected shortfall. Include power hyperbolas and power hyperbolic functions.

URL http://www.inmodelia.com/fattailsr-en.html

Version 1.7-5

Date 2017-05-16

Author Patrice Kiener

Maintainer Patrice Kiener <fattailsr@inmodelia.com>

Depends R (>= 3.1.0)

Imports minpack.lm, timeSeries, parallel, methods, stats

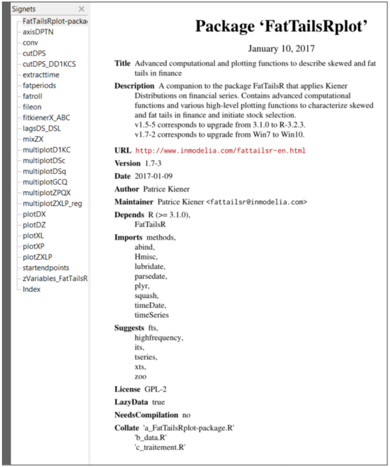
Suggests zoo, xts

License GPL-2

LazyData true

NeedsCompilation no

Collate 'a_FatTailsR-package.R' 'b_data.R' 'c_trigohp.R' 'd_logishp.R' 'e_conversion.R' 'f_kiener1.R' 'g_kiener2.R' 'h_kiener3.R' 'i_kiener4.R' 'j_kiener7.R' 'k_kienerES.R' 'l_moments.R' 'm_laplaceroll.R' 'n_estimation.R'



Signets

FatTailsRplot-packa
axisDPTN
conv
cutDPS
cutDPS_DD1KCS
extracttime
fatperiods
fatroll
fileon
fitkienerX_ABC
lagsDS_DSL
mixZX
multiplotD1KC
multiplotDSc
multiplotDSq
multiplotGCQ
multiplotZPQX
multiplotZXLP_reg
plotDX
plotDZ
plotXL
plotXP
plotZXLP
startendpoints
zVariables_FatTailsR
Index

Package 'FatTailsRplot'

January 10, 2017

Title Advanced computational and plotting functions to describe skewed and fat tails in finance

Description A companion to the package FatTailsR that applies Kiener Distributions on financial series. Contains advanced computational functions and various high-level plotting functions to characterize skewed and fat tails in finance and initiate stock selection. v1.5-5 corresponds to upgrade from 3.1.0 to R-3.2.3. v1.7-2 corresponds to upgrade from Win7 to Win10.

URL http://www.inmodelia.com/fattailsr-en.html

Version 1.7-3

Date 2017-01-09

Author Patrice Kiener

Maintainer Patrice Kiener <fattailsr@inmodelia.com>

Depends R (>= 3.1.0), FatTailsR

Imports methods, abind, Hmisc, lubridate, parsedate, plyr, squash, timeDate, timeSeries

Suggests fts, highfrequency, its, tseries, xts, zoo

License GPL-2

LazyData true

NeedsCompilation no

Collate 'a_FatTailsRplot-package.R' 'b_data.R' 'c_traitement.R'

url : FatTailsR

Package FatTailsR

Les 4 lois de probabilité, plus une fonction qui réunit les 7 paramètres (m, g, a, k, w, d, e)

- d,p,q,r kiener1(xqpn, m = 0, g = 1, k = 3.2)
- d,p,q,r kiener2(xqpn, m = 0, g = 1, a = 3.2, w = 3.2)
- d,p,q,r kiener3(xqpn, m = 0, g = 1, k = 3.2, d = 0)
- d,p,q,r kiener4(xqpn, m = 0, g = 1, k = 3.2, e = 0)
- d,p,q,r kiener7(xqpn, coefk = c(0, 1, 3.2, 3.2, 3.2, 0, 0))

3 fonctions d'apprentissage et 2 algorithmes d'estimation paramétrique (régression, estimation)

- paramkienerX(X, algo = c("reg", "estim"), ...)
- fitkienerX(X, algo = c("reg", "estim"), ...)
- regkienerLX(X, model = "K4", ...)

Les moments empiriques, les **vrais** moments, 2 mesures de risque (VaR, ES)

- xmoments(x, ...)
- kmoments(coefk, model = "K2", ...)
- varkiener7(p, coefk = c(0, 1, 3.2, 3.2, 3.2, 0, 0), ...)
- eskiener7(p, coefk = c(0, 1, 3.2, 3.2, 3.2, 0, 0), ...)

2 jeux de données et des fonctions de manipulation de données

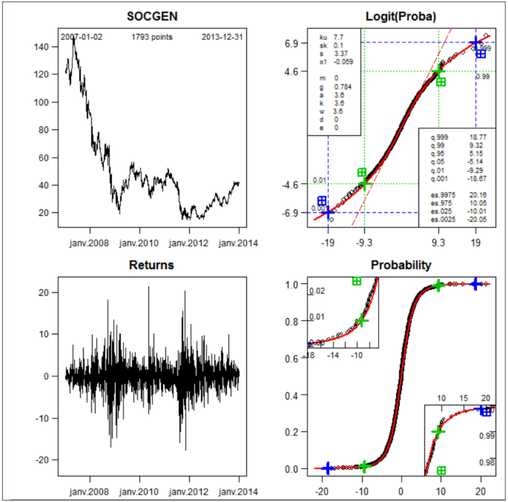
- getDSdata()
- extractData(...) aux formats data.frame, matrix, timeSeries, xts, zoo.
- dimdim(x), fatreturns(x, log = TRUE, ...), etc...

Code FatTailsR

```
>
> library("abind")
> library("timeDate")
> library("timeSeries")
> library("FatTailsR")
> library("FatTailsRplot")
> rm(list=ls(all=TRUE))
>
> head(Z <- extractData())
GMT
            GOLD SOCGEN VIVENDI EURUSD   VIX   CAC40     DJIA   SP500
2007-01-02 15400 123.00   30.04 1.3279    NA 5617.71       NA      NA
2007-01-03 15150 123.09   30.13 1.3165 12.04 5610.92 12474.52 1416.60
2007-01-04 15150 123.09   30.20 1.3084 11.51 5574.56 12480.69 1418.34
2007-01-05 15150 122.25   29.87 1.3007 12.14 5517.35 12398.01 1409.71
2007-01-08 15250 121.69   30.24 1.3022 12.00 5518.59 12423.49 1412.84
2007-01-09 15100 121.41   31.09 1.2998 11.91 5533.03 12416.60 1412.11
>
> round(head(X <- fatreturns(Z)), 3)
GMT
             GOLD SOCGEN VIVENDI EURUSD    VIX  CAC40   DJIA  SP500
2007-01-02  0.000  0.000   0.000  0.000  0.000  0.000  0.000  0.000
2007-01-03 -1.637  0.073   0.299 -0.862  0.000 -0.121  0.000  0.000
2007-01-04  0.000  0.000   0.232 -0.617 -4.502 -0.650  0.049  0.123
2007-01-05  0.000 -0.685  -1.099 -0.590  5.329 -1.032 -0.665 -0.610
2007-01-08  0.658 -0.459   1.231  0.115 -1.160  0.022  0.205  0.222
2007-01-09 -0.988 -0.230   2.772 -0.184 -0.753  0.261 -0.055 -0.052
>
> (coefk <- paramkienerX(X, dgts = 13))
             m     g   a   k   w      d     e
GOLD     0.000 0.361 4.3 4.4 4.5 -0.004 -0.02
SOCGEN   0.000 0.784 3.6 3.6 3.6  0.000  0.00
VIVENDI  0.000 0.430 3.9 4.0 4.1 -0.006 -0.02
EURUSD   0.007 0.173 6.1 6.0 5.9  0.003  0.02
VIX     -0.341 1.709 4.7 4.1 3.7  0.030  0.12
CAC40    0.000 0.367 3.4 3.4 3.4 -0.001  0.00
DJIA     0.037 0.272 2.8 2.9 2.9 -0.005 -0.02
SP500    0.050 0.298 2.8 2.8 2.9 -0.010 -0.03
>
> kmoments(coefk, dgts = 3)[,1:5]
            ku     ke     sk    sd     m1
GOLD    20.364 17.364 -0.209 1.496 -0.014
SOCGEN      NA     NA  0.000 3.508  0.000
VIVENDI     NA     NA -0.340 1.841 -0.021
EURUSD   6.668  3.668  0.070 0.672  0.010
VIX         NA     NA  1.519 7.257  0.027
CAC40       NA     NA  0.000 1.694  0.000
DJIA        NA     NA     NA 1.473  0.022
SP500       NA     NA     NA 1.558  0.034
>
```

tData : SOCGEN

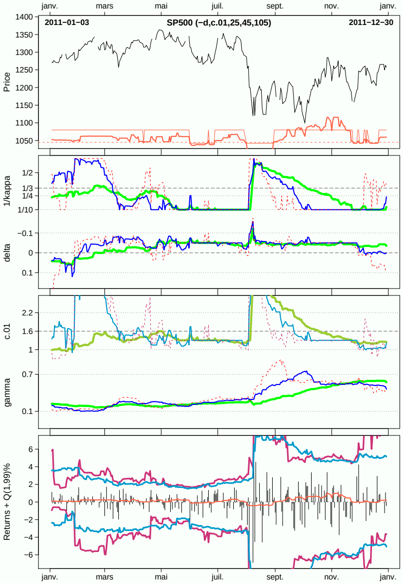
```
>
> reg  <- regkienerLX(X[,"SOCGEN"], model = "K3")
>
> attributes(reg)
$names
 [1] "dfrXP"    "dfrXL"    "dfrXR"    "dfrEP"    "dfrEL"
 [6] "dfrED"    "regk0"    "coefk0"   "vcovk0"   "vcovk0m"
[11] "mcork0"   "coefk"    "coefk1"   "coefk2"   "coefk3"
[16] "coefk4"   "quantk"   "quantes"  "coefr"    "coefr1"
[21] "coefr2"   "coefr3"   "coefr4"   "quantr"   "quantesr"
[26] "dfrQkPk"  "dfrQkLk"  "dfrESkPk" "dfrESkLk" "fitk"

$class
[1] "clregk"

>
> reg$coefr
    m     g     a     k     w     d     e
0.000 0.784 3.600 3.600 3.600 0.000 0.000
>
> reg$quantr
q.0001 q.0005  q.001  q.005   q.01   q.05   q.50   q.95   q.99
-35.83 -22.77 -18.67 -11.58  -9.29  -5.14   0.00   5.15   9.32
 q.995  q.999 q.9995 q.9999
 11.62  18.77  22.90  36.07
>
> ZX   <- mixZX(Z[,"SOCGEN"], X[,"SOCGEN"], "SOCGEN",
+               Xlim = c(-22, 22), atft = "%b%Y")
> plotZXLP(ZX, reg)
>
```

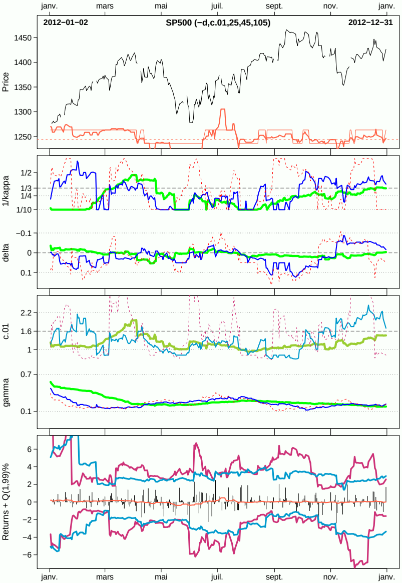


Fenêtres glissantes : fatroll

```
>
> Z    <- extractData("p", "tss", start = "2010-01-01", end = "2012-12-31")
> lstZ <- fatroll(Z, winL = -1, on = "years", offset = 0, out = "Date")
> lapply(lstZ, tail, 4)
$`2010-12-31`
GMT
            GOLD SOCGEN VIVENDI EURUSD   VIX   CAC40     DJIA   SP500
2010-12-28 33990 41.510  20.515 1.3118 17.52 3858.72 11575.54 1258.51
2010-12-29 34350 41.485  20.770 1.3223 17.28 3890.65 11585.38 1259.78
2010-12-30 34230 40.865  20.515 1.3287 17.52 3850.76 11569.71 1257.88
2010-12-31 34230 40.220  20.200 1.3392 17.75 3804.78 11577.51 1257.64

$`2011-12-30`
GMT
            GOLD SOCGEN VIVENDI EURUSD   VIX   CAC40     DJIA   SP500
2011-12-27 39490 16.860   16.57 1.3070 21.91 3103.11 12291.35 1265.43
2011-12-28 39450 16.345   16.36 1.2938 23.52 3071.08 12151.41 1249.64
2011-12-29 38890 16.695   16.76 1.2963 22.65 3127.56 12287.04 1263.02
2011-12-30 39390 17.205   16.92 1.2941 23.40 3159.81 12217.56 1257.60

$`2012-12-31`
GMT
            GOLD SOCGEN VIVENDI EURUSD   VIX   CAC40     DJIA   SP500
2012-12-26    NA     NA      NA 1.3218 19.48      NA 13114.59 1419.83
2012-12-27 39600  28.90   17.35 1.3237 19.47 3674.26 13096.31 1418.10
2012-12-28 40500  28.15   16.97 1.3215 22.72 3620.25 12938.11 1402.43
2012-12-31 40200  28.34   16.95 1.3181 18.02 3641.07 13104.14 1426.19

> lstX <- fatreturns(lstZ)
> lapply(lapply(lstX, round, 3), tail, 4)
$`2010-12-31`
GMT
             GOLD SOCGEN VIVENDI EURUSD    VIX  CAC40   DJIA  SP500
2010-12-28  0.000  0.410  -0.413 -0.350 -0.853 -0.090  0.177  0.077
2010-12-29  1.054 -0.060   1.235  0.797 -1.379  0.824  0.085  0.101
2010-12-30 -0.350 -1.506  -1.235  0.483  1.379 -1.031 -0.135 -0.151
2010-12-31  0.000 -1.591  -1.547  0.787  1.304 -1.201  0.067 -0.019

$`2011-12-30`
GMT
             GOLD SOCGEN VIVENDI EURUSD    VIX  CAC40   DJIA  SP500
2011-12-27 -0.278 -2.113   0.332  0.222  5.536  0.033 -0.022  0.008
2011-12-28 -0.101 -3.102  -1.275 -1.015  7.091 -1.038 -1.145 -1.256
2011-12-29 -1.430  2.119   2.416  0.193 -3.769  1.822  1.110  1.065
2011-12-30  1.277  3.009   0.950 -0.170  3.258  1.026 -0.567 -0.430

$`2012-12-31`
GMT
             GOLD SOCGEN VIVENDI EURUSD     VIX  CAC40   DJIA  SP500
2012-12-26  0.000  0.000   0.000  0.189   8.795  0.000 -0.187 -0.480
2012-12-27 -2.173  0.869   1.043  0.144  -0.051  0.591 -0.139 -0.122
2012-12-28  2.247 -2.629  -2.215 -0.166  15.437 -1.481 -1.215 -1.111
2012-12-31 -0.743  0.673  -0.118 -0.258 -23.176  0.573  1.275  1.680

>
```

Fenêtres glissantes 21, 41, 101 jours : SP500 en 2011 et 2012





SP500 — Paramètres c05 et q05 — Fenêtre 41 jours

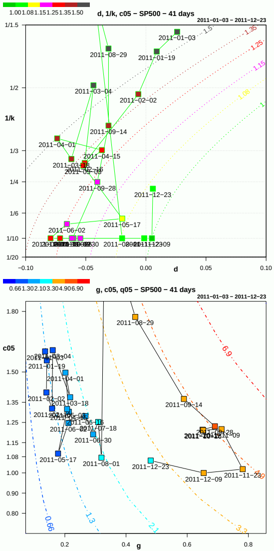


FIGURE – Paramètres c05 (haut) et q05 (bas) de janvier 2011 (gauche) à décembre 2012 (droite). SP500 : Fenêtres glissantes de 41 jours recalculées tous les 10 jours

Fenêtre glissante 100 jours : c.01 et q.01

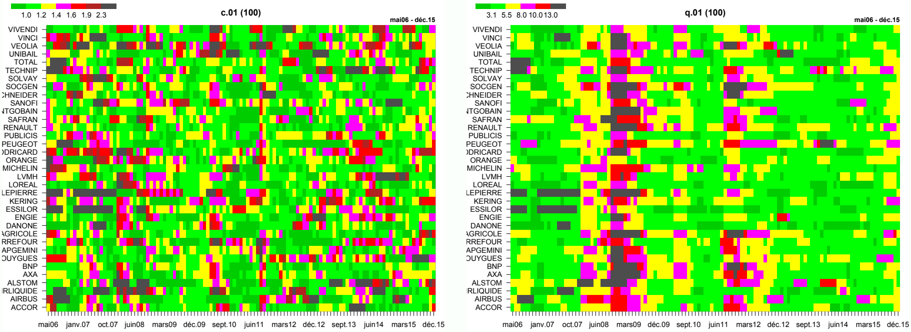


FIGURE – Paramètres c.01 et q.01 de 35 valeurs du CAC 40 de janvier 2006 à décembre 2015. Fenêtres glissantes de 100 jours (≈ 5 mois) rafraichies tous les mois.

Conclusion : Un très bon package. Essayez-le !

patrice.kiener@inmodelia.com

Tél. : +33.9.53.45.07.38