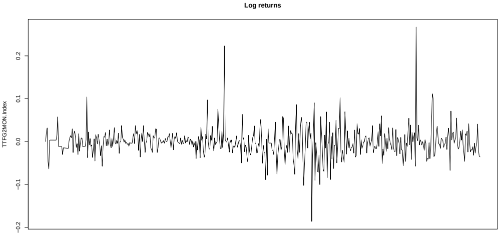
Are copulas in finance time-dynamic?

Moving Window Maximum Likelihood Estimation of Copula Parameters

"Dynamic Copula Methods in Finance"
Bologna (Italy)
September 2010

Magda Mroz
Ulm University
10/09/2010

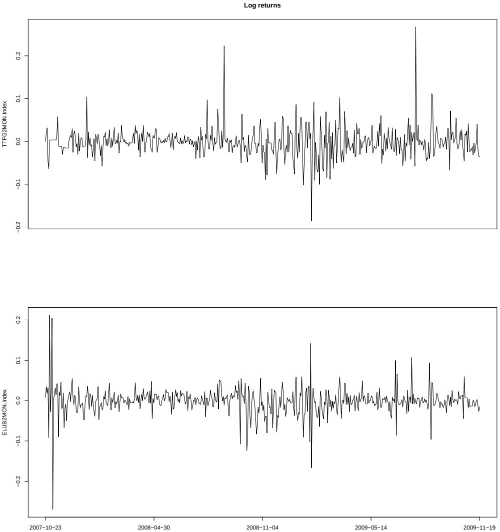
Multivariate Financial Time Series



$$X_t = m + \sigma_t \varepsilon_t$$
$$\sigma_t^2 = \omega + \alpha\, \sigma_{t-1}^2 + \beta\, \varepsilon_{t-1}^2$$

$$\varepsilon_t \sim F \text{ iid}$$

Multivariate Financial Time Series



$$X_{jt} = m_j + \sigma_{jt}\varepsilon_{jt}$$
$$\sigma^2_{jt} = \omega_j + \alpha_j \sigma^2_{jt-1} + \beta_j \varepsilon^2_{jt-1}$$

$$(\varepsilon_{1t}, \ldots, \varepsilon_{dt})' \sim H \text{ iid}$$
$$H = C(\theta, F_1, \ldots, F_d)$$

$$\tilde{\varepsilon}_{jt} = \frac{X_{jt} - \hat{m}_j}{\sqrt{\hat{\sigma}^2_{jt}}}$$
$$\tilde{\varepsilon}_t = (\tilde{\varepsilon}_{1t}, \ldots, \tilde{\varepsilon}_{dt})'$$

Outline

Maximum likelihood on copula data

For $u_1, \dots, u_T$ iid realisations of a d-variate parametric copula $C(\theta)$ with density c

$$\hat{\theta}_T = \underset{\theta}{\operatorname{argmax}} \sum_{t=1}^{T} \log c(\theta, u_t) = \underset{\theta}{\operatorname{argmax}} \sum_{t=1}^{T} l(\theta, u_t).$$

Under the usual regularity conditions

$$\sqrt{T}(\hat{\theta}_T - \theta) \xrightarrow[T\to\infty]{d} N(0, v)$$

with variance

$$v = I(\theta)^{-1} = \mathbb{E}(l_\theta^2(\theta, U))^{-1}.$$

A consistent variance estimator is given by the observed Fisher information

$$\hat{v} = \left(\frac{1}{T}\sum_{t=1}^{T} l_\theta^2(\hat{\theta}_T, u_t)\right)^{-1}.$$

Multivariate data

For d-variate data $x_1, \dots, x_T$ being iid realisations of

$$H = C(\theta, F_1(\delta_1), \dots, F_d(\delta_d)).$$

1. Full maximum likelihood

$$(\hat{\theta}, \hat{\delta}_1, \dots, \hat{\delta}_d) = \text{argmax } L(\theta, \delta_1, \dots, \delta_d)$$

2. Inference for margins
   - $\hat{\delta}_1, \dots, \hat{\delta}_d$ separately for each margin
   - $\hat{\theta}$ on pseudo likelihood

3. Canonical maximum likelihood (semi-parametric approach)

Canonical maximum likelihood

First, obtain pseudo-observations

$$u_t = (\hat{F}_1(x_{1t}), \ldots, \hat{F}_d(x_{dt}))'$$

through a rank transformation

$$\hat{F}_j(x) = \frac{1}{T+1} \sum_{t=1}^{T} \mathbb{1}\{x_{jt} \leqslant x\}.$$

Then $\hat{\theta}_T = \underset{\theta}{\operatorname{argmax}} \sum_{t=1}^{T} l(\theta, u_t)$ is again asymptotically normal, but with a greater variance $v = \frac{\sigma^2}{\beta^2}$. A consistent variance estimator $\hat{v} = \frac{\hat{\sigma}^2}{\hat{\beta}^2}$ can be given.

$\beta = \mathbb{E}\left(l_\theta^2(\theta, F_1(X_1), \ldots, F_d(X_d))\right)$ and
$\sigma^2 = \operatorname{var}\left(l_\theta(\theta, F_1(X_1), \ldots, F_d(X_d)) + \sum_{j=1}^{d} W_j(X_j)\right)$ as in Genest *et al.* (1995).

Moving window estimation

- given a sample $u_1, \dots, u_T$ of copula data respectively pseudo-observations
- time-variation in the dependence structure through time-varying parameter θ_t for all $t = 1, \dots, T$
- assume that the parameter series is actually constant on small time intervals (”local homogeneity”, see e.g. Giacomini, Härdle, and Spokoiny (2009))

moving windows

Estimation on overlapping subsamples of length $b = b(T)$, $b/T \xrightarrow[T\to\infty]{} 0$,

$$\hat{\theta}_t = \operatorname*{argmax}_{\theta} \sum_{\tau=t-b}^{t} l(\theta, u_\tau).$$

results in a parameter series $\{\hat{\theta}_t\}_{t=b,\dots,T}$.

Null hypothesis of parameter constancy

Under the null hypothesispdftk main.pdf cat 1 2 4 5 8 12 14 16-18 21 23-25 27 29 33-end output main4print.pdf

$$\mathbf{H}_0 : \{\theta_t = \theta \; \forall \, t\}$$

$$\frac{\sqrt{b}(\hat{\theta}_t - \theta)}{\sqrt{v}} \xrightarrow[T\to\infty]{d} N(0,1).$$

Null hypothesis of parameter constancy

Under the null hypothesispdftk main.pdf cat 1 2 4 5 8 12 14 16-18 21 23-25 27 29 33-end output main4print.pdf

$$\mathbf{H}_0 : \{\theta_t = \theta \ \forall\, t\}$$

and with consistent estimators for θ and v it holds that

$$\frac{\sqrt{b}(\hat{\theta}_t - \hat{\theta}_T)}{\sqrt{\hat{v}}} \xrightarrow[T\to\infty]{d} N(0,1).$$

local confidence interval

For a given confidence level α we obtain

$$\mathbf{I}_t = \left[\hat{\theta}_T - \sqrt{\frac{\hat{v}}{b}} z_{1-\alpha/2},\ \hat{\theta}_T + \sqrt{\frac{\hat{v}}{b}} z_{1-\alpha/2}\right],$$

Thus, $\hat{\theta}_t \in \mathbf{I}_t$ with probability $1-\alpha$ (asymptotically and pointwise for all t).

Joint probability

We are interested in

$$\mathbb{P}\left(\hat{\theta}_t \in \mathbf{I}_t \ \forall t = b, \ldots, T\right).$$

For $\{\hat{\theta}_t\}$ we can show

- b-dependence
- joint normality with

$$\operatorname{cov}\left(\hat{\theta}_t, \hat{\theta}_s\right) = \begin{cases} 0, & |s-t| > b \\ v\mu_l, & |s-t| = l < b, \ b/l \to \mu_l \end{cases}$$

- (strong) stationarity

$\Longrightarrow$ too strong dependence to apply extreme value theory

Thinning of parameter series

Consider $c \in (0{,}1)$ and for $k = 1, \ldots, N = \frac{T-b}{cb}$

$$\hat{\theta}_{t_k} \text{ with } t_k = b + (k-1)cb.$$

The asymptotic normality under $\mathbf{H}_0$

$$\sqrt{b}\begin{pmatrix} \hat{\theta}_{t_{k_1}} - \theta \\ \hat{\theta}_{t_{k_2}} - \theta \end{pmatrix} \xrightarrow[T\to\infty]{d} N\left(\mathbf{0}, v \cdot \begin{pmatrix} 1 & 1-kc \\ 1-kc & 1 \end{pmatrix}\right)$$

exhibits a more favorable covariance structure

$$\mathrm{cov}\left(\hat{\theta}_{t_{k_1}}, \hat{\theta}_{t_{k_2}}\right) = v(1-kc),\ |k_1 - k_2| = k < 1/c$$

i.e. the covariance doesn't degenerate in the limit.

Joint normality and weak dependence

Consider the normalized and centered estimators

$$\xi_{t_k} := \frac{\sqrt{b}(\hat{\theta}_{t_k} - \hat{\theta}_T)}{\sqrt{\hat{v}}}$$

which are

- asymptotically jointly normal
- stationary, i.e. cov $\left(\xi_{t_{k_1}}, \xi_{t_{k_2}}\right) = \gamma_k$ for $|k_1 - k_2| = k$
- $1/c$-dependent

Particularly, it holds that

$$\lim_{k\to\infty} \gamma_k \log k = 0.$$

Global confidence regions

Extreme value theory for weakly dependent sequences[1] yields

$$\mathbb{P}\left(M_N \leqslant \quad a_N x + d_N\right) \to e^{-e^{-x}} = \Lambda(x) \quad \forall x \in \mathbb{R}$$

with $a_N = \sqrt{2 \ln N}$, $d_N = a_N - \frac{\ln \ln N + \ln 4\pi}{2a_N}$ and $M_N = \max_k \xi_{t_k}$.

[1]See for example Leadbetter *et al.* (1983).

Global confidence regions

Extreme value theory for weakly dependent sequences[1] yields

$$\mathbb{P}\left(m_N > -a_N x - d_N\right) \to e^{-e^{-x}} = \Lambda(x) \quad \forall x \in \mathbb{R}$$

with $a_N = \sqrt{2 \ln N}$, $d_N = a_N - \frac{\ln \ln N + \ln 4\pi}{2 a_N}$ and $m_N = \min_k \xi_{t_k}$.

global confidence region with level α

For the level α we obtain

$$\mathbf{I} = \left[-a_N \lambda_{1-\alpha/2} - d_N, a_N \lambda_{1-\alpha/2} + d_N\right]$$

for the extrema of $\{\xi_{t_k}\}$, i.e.

$$\lim_{N \to \infty} \mathbb{P}\left(M_N \in I, m_N \in I\right) = 1 - \alpha.$$

[1]See for example Leadbetter *et al.* (1983).

Tests

Let α be a fixed level as before.

local test for independent estimators

For $\hat{\theta}_{t_k}$, $t_k = kb$, $n = (T-b)/b$ count the local exceedances

$$S_n = \sum_{k=1}^{n} \mathbb{1}\{\hat{\theta}_{t_k} \notin \mathbf{I}_{t_k}\} \sim \mathrm{Bin}(n, \alpha).$$

With q_α the $(1-\alpha)$-quantile of the $\mathrm{Bin}(n, \alpha)$ distribution, reject $\mathbf{H}_0$ if $S_n > q_\alpha$.

global test for weakly dependent estimators

Consider $t_k = b + (k-1)cb$, $N = (T-b)/cb$ and the extrema of

$$\xi_{t_k} = \frac{\sqrt{b}(\hat{\theta}_{t_k} - \hat{\theta}_T)}{\sqrt{\hat{v}}}.$$

If $M_N \notin \mathbf{I}$ or $m_N \notin \mathbf{I}$, then reject $\mathbf{H}_0$.

Simulation

Clayton copula

- parameters: $\theta = 1.5$, $d = 2, \ldots, 10$
- iid samples, variance $v = I(\theta)^{-1}$

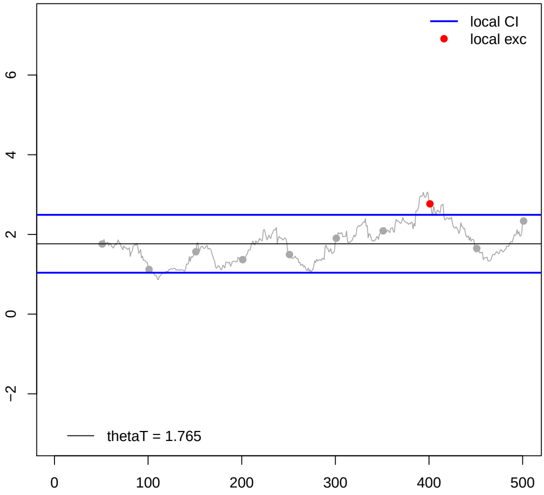
GARCH process

- parameters: $m = 0$, $\omega = 10^{-6}$, $\alpha = 0.1$, $\beta = 0.8$ (analogously to the `garchSim`)
- bivariate GARCH(1,1) processes, normal margins and Clayton dependence
- CML outperforms IFM, variance $v = \frac{\sigma^2}{\beta^2}$

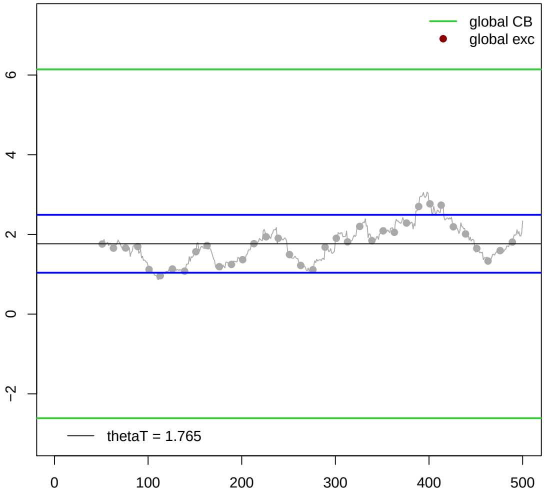
test parameters

- sample size $T = 500$, bandwidth $b = 50$ $\Rightarrow n = 9$
- confidence level $\alpha = 5\%$ $\Rightarrow$ rejection tolerance $q_\alpha = 2$
- thinning with $c = 0.25$ $\Rightarrow N = 36$

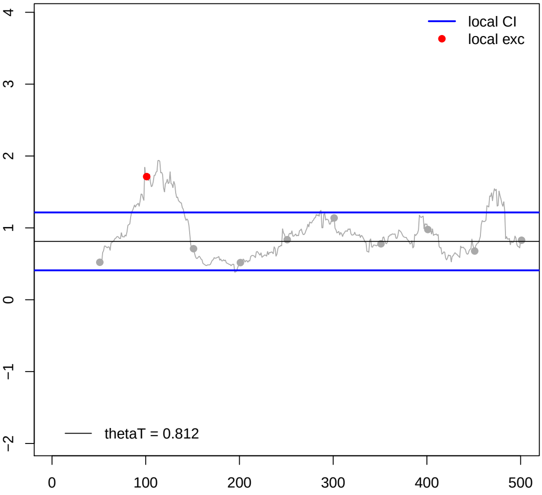
Bivariate Clayton copula



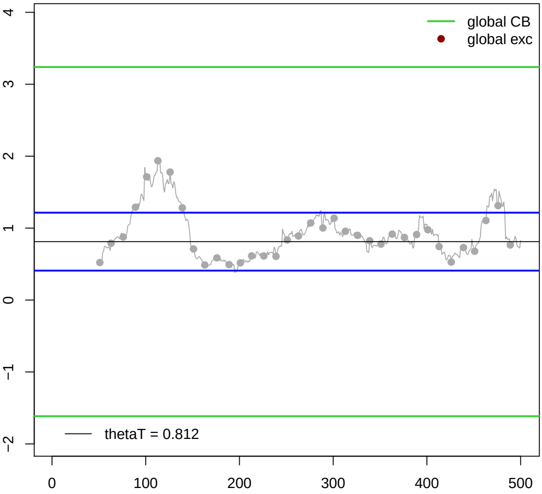
Bivariate Clayton copula



Simulated GARCH process



Simulated GARCH process



Commodity contracts

- 2nd front month future contracts on 11 different commodities (coal, gas, oil, electricity)
- univariate GARCH(1,1) processes for the deseasonalized log returns

$$X_{jt} = m_j + \sigma_{jt}\varepsilon_{jt}$$
$$\sigma^2_{jt} = \omega_j + \alpha_j\sigma^2_{jt-1} + \beta_j\varepsilon^2_{jt-1}$$

- empirical residuals bivariately coupled by a Clayton copula

$$C(\theta, u_1, u_2) = \max\{0, u_1^{-\theta} + u_2^{-\theta} - 1\}^{-1/\theta}$$

- copula estimation via CML

Estimation results

- GARCH parameters estimated with `garchFit`

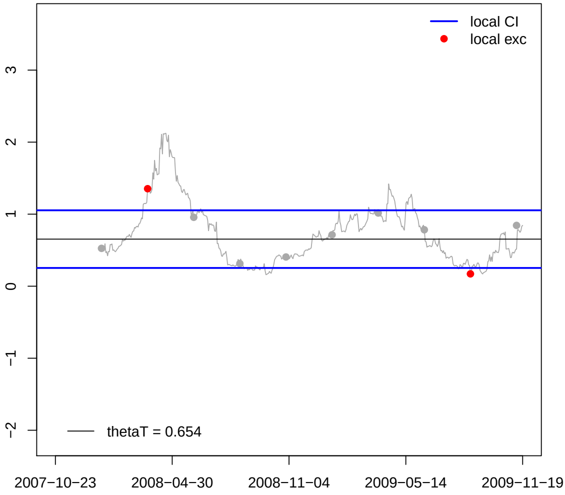
	$\hat{m}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
$\varnothing$	0.00059	0.00028	0.19362	0.58903
std	0.00056	0.00042	0.11019	0.38651

- copula parameters estimated with `fitCopula`

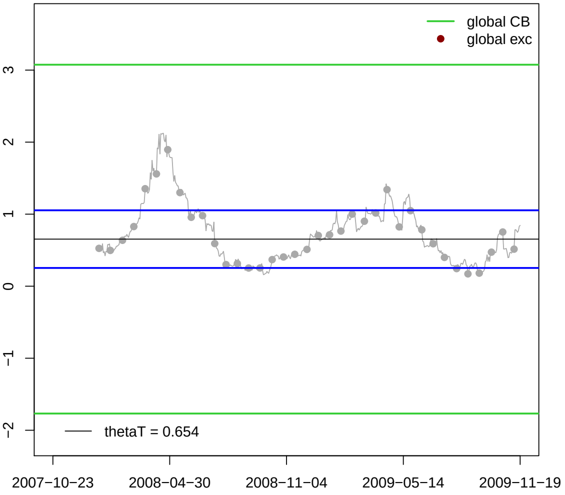
	$\hat{\theta}_T$	$\hat{\nu}$	$\hat{\tau}$	$\hat{\lambda}_L$
$\varnothing$	0.46	2.28	0.19	0.22
std	0.4129	1.2966	–	–
min	0.11	1.31	0.05	0.00
max	1.69	8.24	0.46	0.66

- sample size $T = 539$, bandwidth $b = 54 \ \Rightarrow n = 9$
- confidence level $\alpha = 5\% \ \Rightarrow$ rejection tolerance $q_\alpha = 2$
- thinning with $c = 0.25 \ \Rightarrow N = 36$

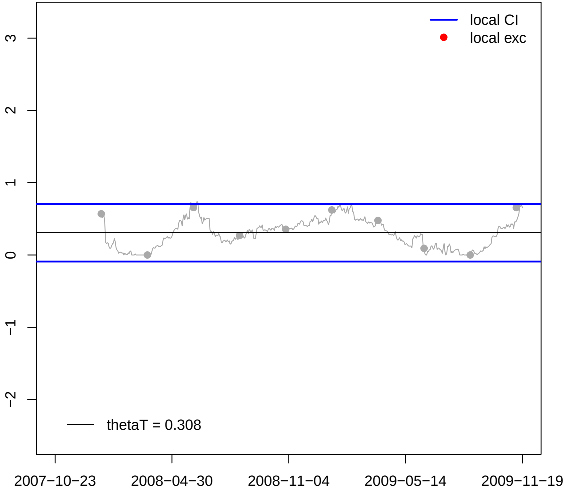
Gas vs. Electricity



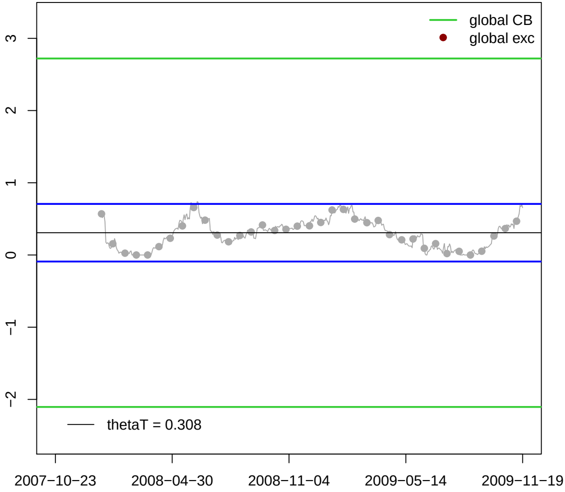
Gas vs. Electricity



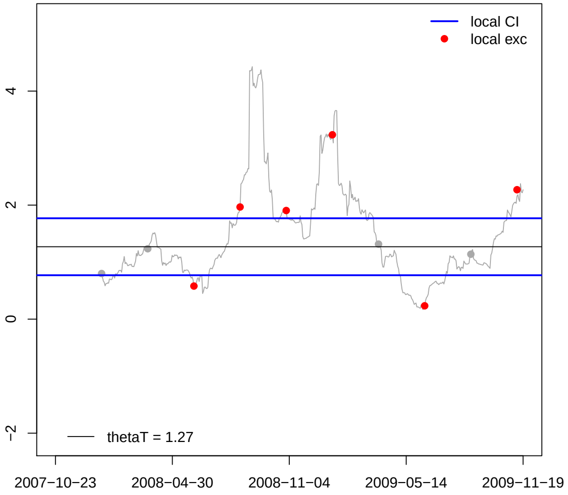
Coal vs. Oil


local CI
local exc
thetaT = 0.308
3
2
1
0
−1
−2
2007−10−23
2008−04−30
2008−11−04
2009−05−14
2009−11−19

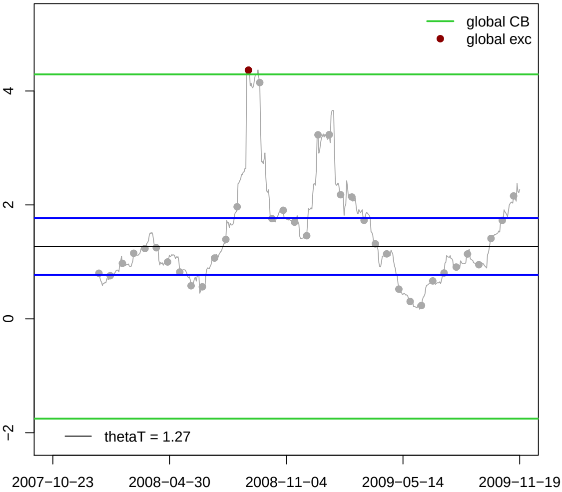
Coal vs. Oil


global CB
global exc
thetaT = 0.308
3
2
1
0
−1
−2
2007−10−23
2008−04−30
2008−11−04
2009−05−14
2009−11−19

Coal with different delivery places



Coal with different delivery places



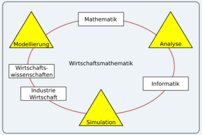
Summary

- GARCH-time series model as a tractable financial model
  - rather for risk management applications
  - not common in option pricing or hedging
- maximum likelihood estimation is most common and well-understood
  - the available implementations (e.g. `copula` package) even contain the suitable variance estimators
- applied to empirical data which consists of future contracts - these are more liquid and comparable than spot prices
- the moving window approach considers the local homogeneity
- thinning allows for a global view

Conclusions

- unbiased test of copula parameter constancy
- global exceedances rarely occur
- time variation can be considered as inherent in the statistical approach

Contact



Research Training Group 1100
Ulm University

Magda Mroz
magda.mroz@uni-ulm.de

Thank you for your attention

References

Christian Genest, Kilani Ghoudi, and Louis-Paul Rivest.
A semiparametric estimation procedure of dependence parameters in multivariate families of distributions.
Biometrika, 82(3):543–552, 1995.

Enzo Giacomini, Wolfgang Härdle, and Vladimir Spokoiny.
Inhomogeneous dependency modelling with time varying copulae.
Journal of Business and Economic Statistics, 27:224–234, 2009.

M. R. Leadbetter, Georg Lindgren, and Holger Rootzen.
Extremes and Related Properties of Random Sequences and Processes.
Springer, 1983.