MULTIVARIATE EXTREMES AND RISK

Richard L. Smith
Department of Statistics and Operations Research
University of North Carolina
Chapel Hill, NC 27599-3260
rls@email.unc.edu

Interface 2008
RISK: Reality
Durham, N.C.
May 22, 2008

http://www.stat.unc.edu/faculty/rs/talks/interface.pdf

I. OVERVIEW OF UNIVARIATE EXTREME VALUE THEORY

II. MULTIVARIATE EXTREME VALUE THEORY

III. ALTERNATIVE FORMULATIONS

IV. MAX-STABLE PROCESSES

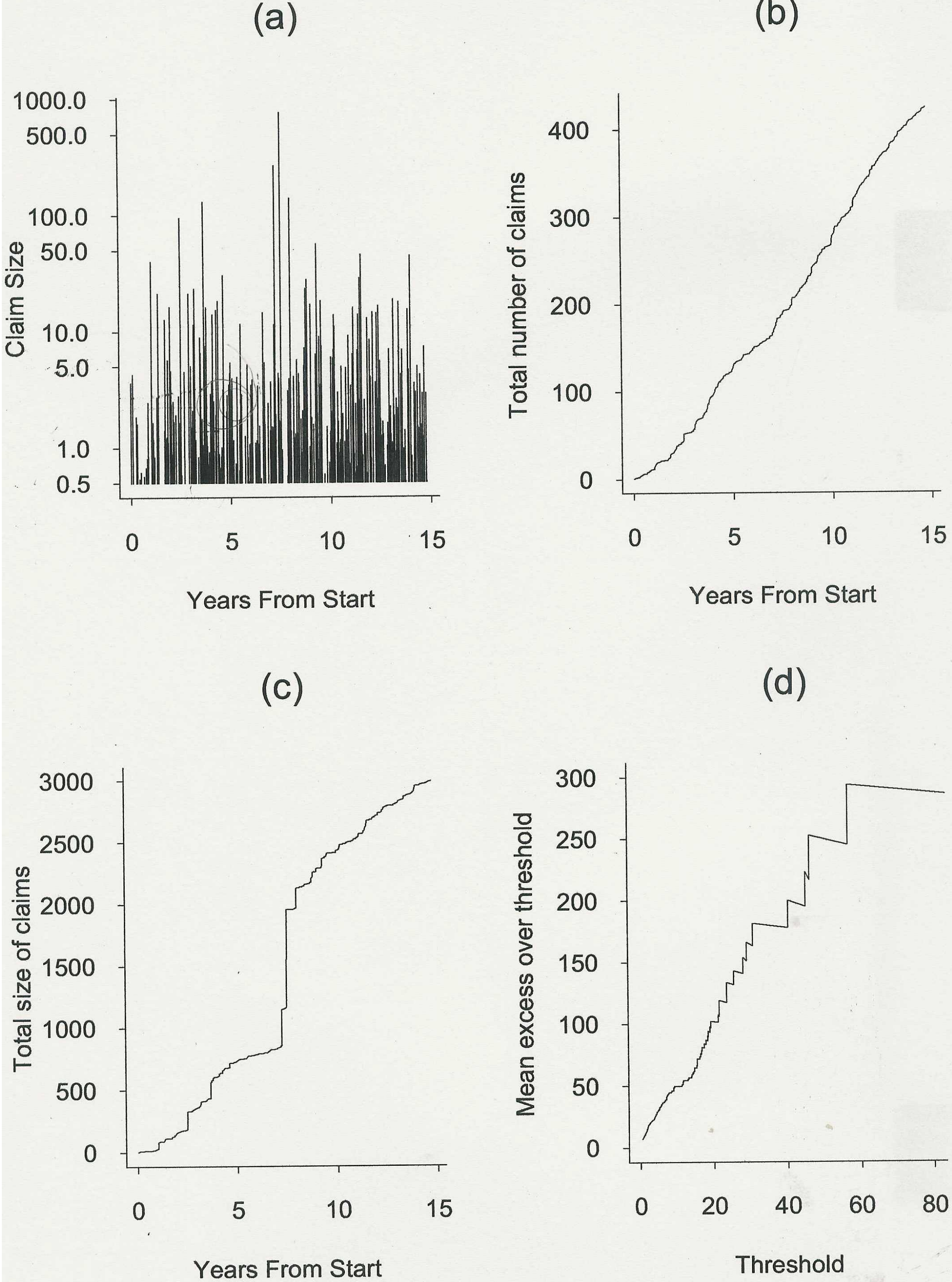
I. OVERVIEW OF UNIVARIATE EXTREME VALUE THEORY

INSURANCE RISK EXAMPLE
(Smith and Goodman, 2000)

The data consist of all insurance claims experienced by a large international oil company over a threshold during a 15-year period — a total of 393 claims.

Total of all 393 claims: 2989.6

10 largest claims: 776.2, 268.0, 142.0, 131.0, 95.8, 56.8, 46.2, 45.2, 40.4, 30.7.



Some plots of the insurance data.

Some problems:

1. What is the distribution of very large claims?

2. Is there any evidence of a change of the distribution over time?

3. What is the influence of the different types of claim?

4. How should one characterize the risk to the company? More precisely, what probability distribution can one put on the amount of money that the company will have to pay out in settlement of large insurance claims over a future time period of, say, three years?

EXTREME VALUE DISTRIBUTIONS

$X_1, X_2, ...,$ i.i.d., $F(x) = \Pr\{X_i \le x\}$, $M_n = \max(X_1, ..., X_n)$,

$$\Pr\{M_n \le x\} = F(x)^n.$$

For non-trivial results must *renormalize*: find $a_n > 0, b_n$ such that

$$\Pr\left\{\frac{M_n - b_n}{a_n} \le x\right\} \;=\; F(a_n x + b_n)^n \to H(x).$$

The *Three Types Theorem* (Fisher-Tippett, Gnedenko) asserts that if nondegenerate H exists, it must be one of three types:

$$H(x) \;=\; \exp(-e^{-x}),\ \text{all } x \quad \text{(Gumbel)}$$

$$H(x) \;=\; \begin{cases} 0 & x < 0 \\ \exp(-x^{-\alpha}) & x > 0 \end{cases} \text{(Fréchet)}$$

$$H(x) \;=\; \begin{cases} \exp(-|x|^{\alpha}) & x < 0 \\ 1 & x > 0 \end{cases} \text{(Weibull)}$$

In Fréchet and Weibull, $\alpha > 0$.

The three types may be combined into a single *generalized extreme value* (GEV) distribution:

$$H(x) = \exp\left\{ -\left(1 + \xi \frac{x-\mu}{\psi}\right)_{+}^{-1/\xi} \right\},$$

$(y_+ = \max(y, 0))$

where μ is a location parameter, $\psi > 0$ is a scale parameter and ξ is a shape parameter. $\xi \to 0$ corresponds to the Gumbel distribution, $\xi > 0$ to the Fréchet distribution with $\alpha = 1/\xi$, $\xi < 0$ to the Weibull distribution with $\alpha = -1/\xi$.

$\xi > 0$: "long-tailed" case, $1 - F(x) \propto x^{-1/\xi}$,

$\xi = 0$: "exponential tail"

$\xi < 0$: "short-tailed" case, finite endpoint at $\mu - \xi/\psi$

EXCEEDANCES OVER THRESHOLDS

Consider the distribution of X conditionally on exceeding some high threshold u:

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}.$$

As $u \to \omega_F = \sup\{x : \ F(x) < 1\}$, often find a limit

$$F_u(y) \approx G(y; \sigma_u, \xi)$$

where G is *generalized Pareto distribution* (GPD)

$$G(y; \sigma, \xi) = 1 - \left(1 + \xi \frac{y}{\sigma}\right)_+^{-1/\xi}.$$

Equivalence to three types theorem established by Pickands (1975).

The Generalized Pareto Distribution

$$G(y;\sigma,\xi) = 1 - \left(1 + \xi\frac{y}{\sigma}\right)_+^{-1/\xi}.$$

$\xi > 0$: long-tailed (equivalent to usual Pareto distribution), tail like $x^{-1/\xi}$,

$\xi = 0$: take limit as $\xi \to 0$ to get

$$G(y;\sigma,0) = 1 - \exp\left(-\frac{y}{\sigma}\right),$$

i.e. exponential distribution with mean σ,

$\xi < 0$: finite upper endpoint at $-\sigma/\xi$.

POINT PROCESS APPROACH

Two-dimensional plot of exceedance times and exceedance levels forms a nonhomogeneous Poisson process with

$$\begin{aligned}\Lambda(A) &= (t_2 - t_1)\Psi(y;\mu,\psi,\xi) \\ \Psi(y;\mu,\psi,\xi) &= \left(1+\xi\frac{y-\mu}{\psi}\right)^{-1/\xi}\end{aligned}$$

$(1 + \xi(y-\mu)/\psi > 0)$.

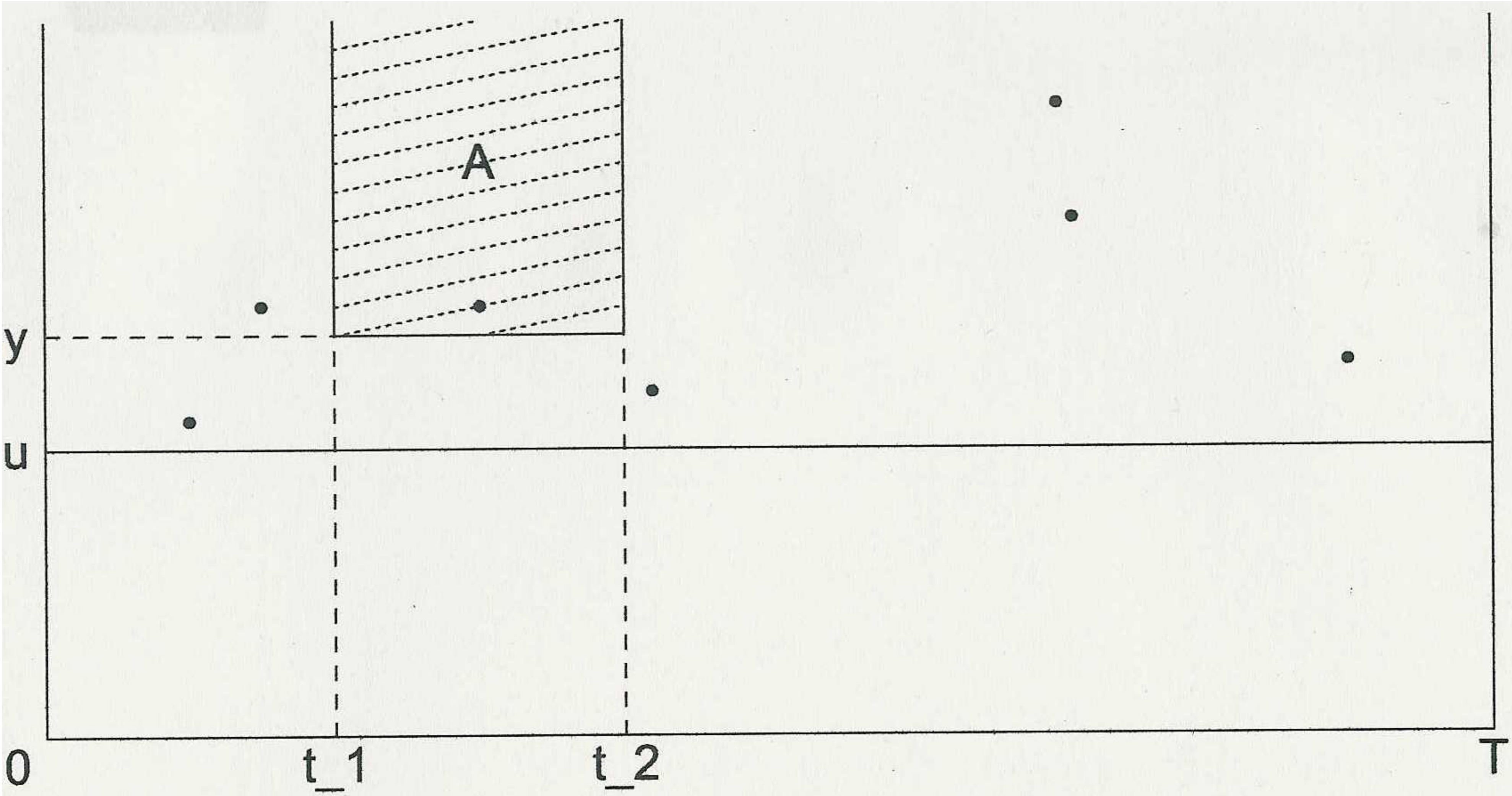


Illustration of point process model.

APPLICATION TO INSURANCE DATA

We apply the GPD and point process approaches to the 393 insurance claims described at the beginning of the talk.

GPD fits to various thresholds:

u	N_u	Mean Excess	σ	ξ
0.5	393	7.11	1.02	1.01
2.5	132	17.89	3.47	0.91
5	73	28.9	6.26	0.89
10	42	44.05	10.51	0.84
15	31	53.60	5.68	1.44
20	17	91.21	19.92	1.10
25	13	113.7	74.46	0.93
50	6	37.97	150.8	0.29

Point process approach:

u	N_u	μ	$\log \psi$	ξ
0.5	393	26.5 (4.4)	3.30 (0.24)	1.00 (0.09)
2.5	132	26.3 (5.2)	3.22 (0.31)	0.91 (0.16)
5	73	26.8 (5.5)	3.25 (0.31)	0.89 (0.21)
10	42	27.2 (5.7)	3.22 (0.32)	0.84 (0.25)
15	31	22.3 (3.9)	2.79 (0.46)	1.44 (0.45)
20	17	22.7 (5.7)	3.13 (0.56)	1.10 (0.53)
25	13	20.5 (8.6)	3.39 (0.66)	0.93 (0.56)

Standard errors are in parentheses

Predictive Distributions of Future Losses

What is the probability distribution of future losses over a specific time period, say 1 year?

Let Y be future total loss. Distribution function $G(y; \mu, \psi, \xi)$ — in practice this must itself be simulated.

Traditional frequentist approach:

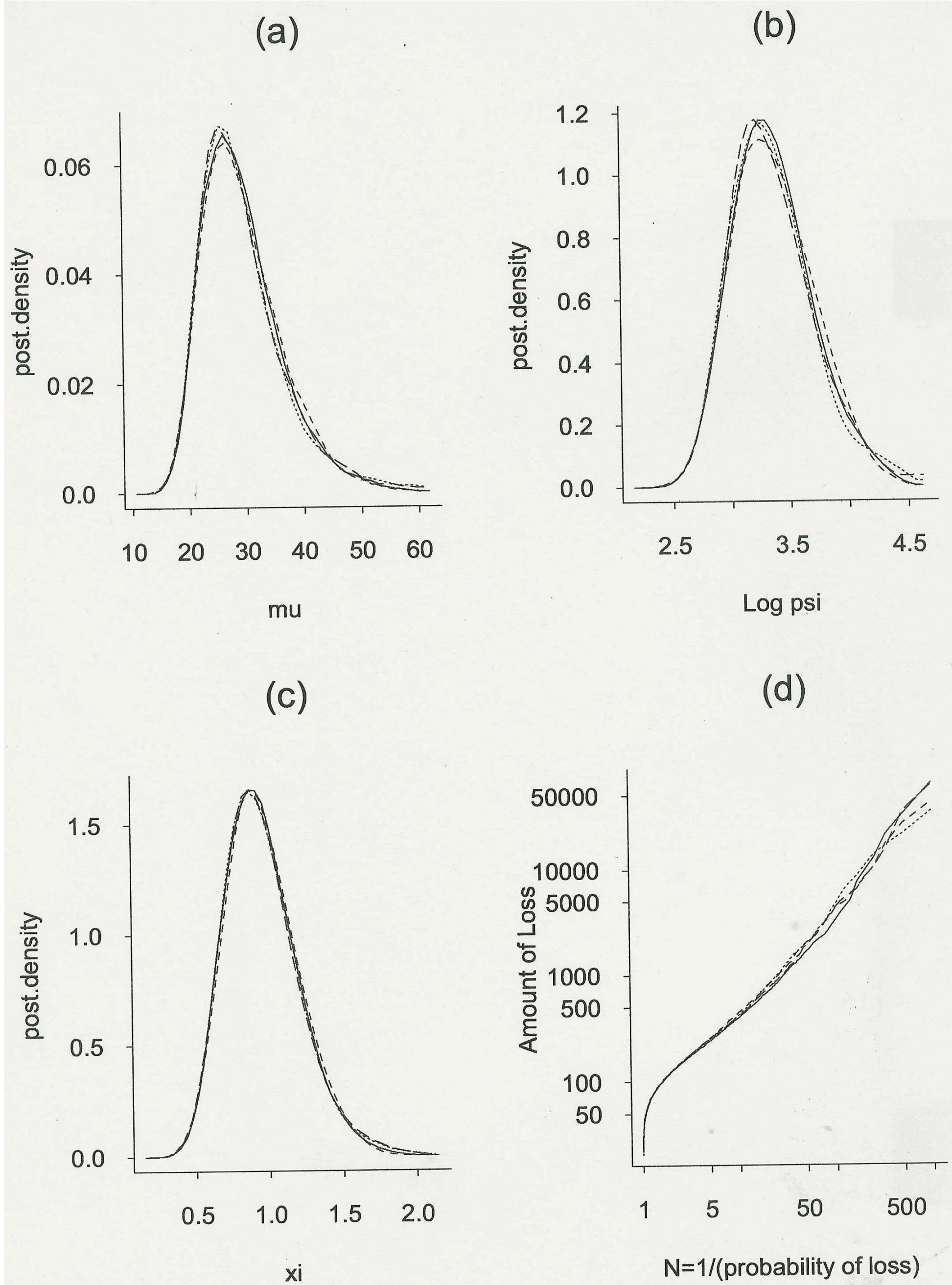
$$\widehat{G}(y) = G(y; \widehat{\mu}, \widehat{\psi}, \widehat{\xi})$$

where $\widehat{\mu}$, $\widehat{\psi}$, $\widehat{\xi}$ are MLEs.

Bayesian:

$$\tilde{G}(y) = \int G(y; \mu, \psi, \xi) d\pi(\mu, \psi, \xi \mid \mathbf{X})$$

where $\pi(\cdot \mid \mathbf{X})$ denotes posterior density given data $\mathbf{X}$.



Estimated posterior densities for the three parameters, and for the predictive distribution function. Four independent Monte Carlo runs are shown for each plot.

II. MULTIVARIATE EXTREME VALUE THEORY

Multivariate extreme value theory applies when we are interested in the joint distribution of extremes from several random variables.

Examples:

- Winds and waves on an offshore structure
- Meteorological variables, e.g. temperature and precipitation
- Air pollution variables, e.g. ozone and sulfur dioxide
- Finance, e.g. price changes in several stocks or indices
- Spatial extremes, e.g. joint distributions of extreme precipitation at several locations

LIMIT THEOREMS FOR MULTIVARIATE SAMPLE MAXIMA

Let $\mathbf{Y}_i = (Y_{i1} ... Y_{iD})^T$ be i.i.d. D-dimensional vectors, $i = 1, 2, ...$

$M_{nd} = \max\{Y_{1d}, ..., Y_{nd}\} (1 \le d \le D)$ — d'th-component maximum

Look for constants a_{nd}, b_{nd} such that

$$\Pr\left\{\frac{M_{nd} - b_{nd}}{a_{nd}} \le x_d,\ d = 1, ..., D\right\} \ \to\ G(x_1, ..., x_D).$$

Vector notation:

$$\Pr\left\{\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n} \le \mathbf{x}\right\} \ \to\ G(\mathbf{x}).$$

Two general points:

1. If G is a multivariate extreme value distribution, then each of its marginal distributions must be one of the univariate extreme value distributions, and therefore can be represented in GEV form

2. The form of the limiting distribution is invariant under monotonic transformation of each component. Therefore, without loss of generality we can transform each marginal distribution into a specified form. For most of our applications, it is convenient to assume the Fréchet form:

$$\Pr\{X_d \leq x\} = \exp(-x^{-\alpha}), \; x > 0, \; d = 1, ..., D.$$

Here $\alpha > 0$. The case $\alpha = 1$ is called *unit Fréchet.*

Basics of Multivariate Regular Variation

(following Resnick (2007), Chapter 6)

After transformation of margins,

$$\lim_{t\to\infty} t\,\mathrm{Pr}\left\{\frac{\mathbf{Y}_i}{b(t)} \in A\right\} \;=\; \nu(A)$$

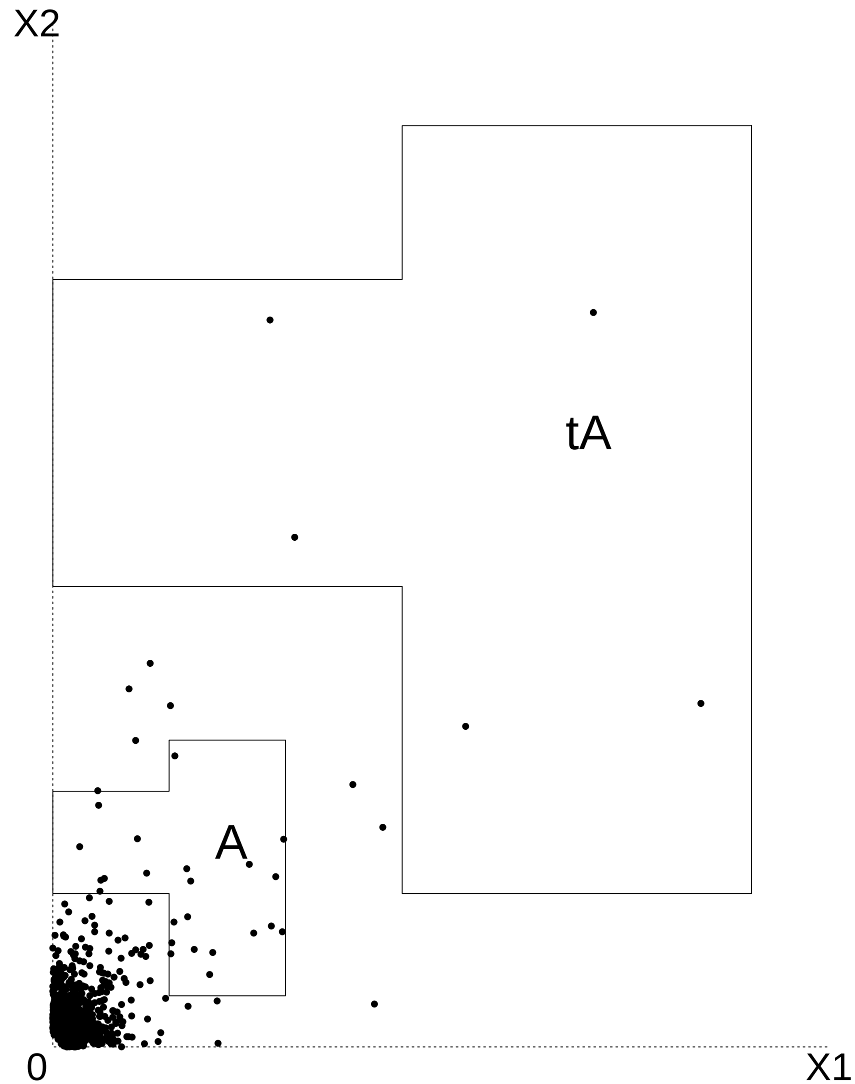
b regularly varying function of index $\alpha > 0$ (i.e. $\frac{b(tx)}{b(t)} \to x^\alpha$ as $t \to \infty$ for fixed $x > 0$), ν a measure on the cone

$$\mathcal{E} \;=\; [0,\infty]^D - \{0\}$$

satisfying

$$\nu(tA) \;=\; t^{-\alpha}\nu(A)$$

for any scalar $t > 0$.


X2
tA
A
0
X1

The last statement implies that ν can be decomposed into a product of *radial* and *angular* components. Define

$$\mathcal{S}_D \;=\; \{(x_1,...,x_D):\; x_1 \geq 0,...,x_D \geq 0, x_1+...+x_D = 1\}.$$

Consider sets A of form

$$A \;=\; \left\{\mathbf{x} \in \mathcal{E}:\; ||\mathbf{x}|| > r,\; \frac{\mathbf{x}}{||\mathbf{x}||} \in S\right\}$$

for some $S \in \mathcal{S}_D$, $||\mathbf{x}|| = \sum_{j=1}^{D} x_j$.

Then

$$\nu(A) \;=\; r^{-\alpha} H(S)$$

for some measure H on $\mathcal{S}_D$.

First Interpretation:

Consider i.i.d. vectors $\{\mathbf{X}_i, \ i = 1, 2, ...\}$ whose distribution is MRV.

Let P_n be a measure on $[0, \infty]^D$ consisting of the points $\left\{\frac{\mathbf{X}_1}{b(n)}, ..., \frac{\mathbf{X}_n}{b(n)}\right\}$.

Let A be a measurable set on $\mathcal{E}$, then the expected number of points of P_n in A is

$$n \,\mathrm{Pr}\left\{\frac{\mathbf{X}_i}{b(n)} \in A\right\} \ \rightarrow \ \nu(A) \ \text{ as } n \rightarrow \infty.$$

With some measure-theoretic formalities, this shows that P_n converges vaguely to a nonhomogeneous Poisson process on $\mathcal{E}$ with intensity measure ν.

Second Interpretation:

Fix $x_1 \geq 0, ..., x_D \geq 0, \sum_{d=1}^{D} x_d > 0$. Let A be the complement of

$$[0, x_1] \times [0, x_2] \times ... \times [0, x_D].$$

Then

$$\Pr^n \left\{ \frac{X_{i1}}{b(n)} \leq x_1, ..., \frac{X_{iD}}{b(n)} \leq x_D \right\} \tag{1}$$

is the probability that P_n places no points in the set A. By Poisson limit theorem, this probability tends to $e^{-\nu(A)}$ as $n \to \infty$. Therefore, the limit of (1) is

$$G(\mathbf{x}) \;=\; \exp\{-V(\mathbf{x})\} \tag{2}$$

where $V(\mathbf{x}) = \nu(A)$.

Moreover, using the radial-spectral decomposition of ν,

$$V(\mathbf{x}) \;=\; \int_{\mathcal{S}_D} \max_{d=1,...,D} \left(\frac{w_j}{x_j} \right) dH(w). \tag{3}$$

The function $V(\mathbf{x})$ is called the *exponent measure* and formula (3) is the *Pickands representation*. If we fix $d' \in \{1, ..., D\}$ with $0 < x_{d'} < \infty$, and define $x_d = +\infty$ for $d \neq d'$, then

$$\begin{aligned} V(\mathbf{x}) &= \int_{\mathcal{S}_D} \max_{d=1,...,D} \left(\frac{w_d}{x_d} \right) dH(w) \\ &= \frac{1}{x_{d'}} \int_{\mathcal{S}_D} w_{d'} dH(w) \end{aligned}$$

so we must have

$$\int_{\mathcal{S}_D} w_d dH(w) = 1, \quad d = 1, ..., D, \tag{4}$$

to ensure that the marginal distributions are correct.

Note that

$$kV(\mathbf{x}) = V\left(\frac{\mathbf{x}}{k}\right)$$

(which is in fact another characterization of V) so

$$\begin{aligned} G^k(\mathbf{x}) &= \exp(-kV(\mathbf{x})) \\ &= \exp\left(-V\left(\frac{\mathbf{x}}{k}\right)\right) \\ &= G\left(\frac{\mathbf{x}}{k}\right). \end{aligned}$$

Hence G is *max-stable*. In particular, if $\mathbf{X}_1, ..., \mathbf{X}_k$ are i.i.d. from G, then $\max\{\mathbf{X}_1, ..., \mathbf{X}_k\}$ (vector of componentwise maxima) has the same distribution as $k\mathbf{X}_1$.

EXAMPLES

Logistic (Gumbel and Goldstein, 1964)

$$V(\mathbf{x}) = \left(\sum_{d=1}^{D} x_d^{-r} \right)^{1/r}, \quad r \geq 1.$$

Check:

1. $V(\mathbf{x}/k) = kV(\mathbf{x})$
2. $V((+\infty, +\infty, ..., x_d, ..., +\infty, +\infty) = x_d^{-1}$
3. $e^{-V(\mathbf{x})}$ is a valid c.d.f.

Limiting cases:

- $r = 1$: independent components
- $r \to \infty$: the limiting case when $X_{i1} = X_{i2} = ... = X_{iD}$ with probability 1.

Asymmetric logistic (Tawn 1990)

$$V(\mathbf{x}) = \sum_{c \in C} \left\{ \sum_{i \in c} \left(\frac{\theta_{i,c}}{x_i} \right)^{r_c} \right\}^{1/r_c},$$

where C is the class of non-empty subsets of $\{1, ..., D\}$, $r_c \geq 1$, $\theta_{i,c} = 0$ if $i \notin c$, $\theta_{i,c} \geq 0$, $\sum_{c \in C} \theta_{i,c} = 1$ for each i.

Negative logistic (Joe 1990)

$$V(\mathbf{x}) = \sum \frac{1}{x_j} + \sum_{c \in C:\ |c| \geq 2} (-1)^{|c|} \left\{ \sum_{i \in c} \left(\frac{\theta_{i,c}}{x_i} \right)^{r_c} \right\}^{1/r_c},$$

$r_c \leq 0$, $\theta_{i,c} = 0$ if $i \notin c$, $\theta_{i,c} \geq 0$, $\sum_{c \in C} (-1)^{|c|} \theta_{i,c} \leq 1$ for each i.

Tilted Dirichlet (Coles and Tawn 1991)

A general construction: Suppose h^* is an arbitrary positive function on $\mathcal{S}_d$ with $m_d = \int_{\mathcal{S}_D} u_d h^*(\mathbf{u}) d\mathbf{u} < \infty$, then define

$$h(\mathbf{w}) \;=\; \left(\sum m_k w_k\right)^{-(D+1)} \prod_{d=1}^{D} m_d h^* \left(\frac{m_1 w_1}{\sum m_k w_k}, \ldots, \frac{m_D w_D}{\sum m_k w_k} \right).$$

h is density of positive measure H, $\int_{\mathcal{S}_D} u_d dH(\mathbf{u}) = 1$ each d.

As a special case of this, they considered Dirichlet density

$$h^*(\mathbf{u}) \;=\; \frac{\Gamma(\sum \alpha_j)}{\prod_d \Gamma(\alpha_d)} \prod_{d=1}^{D} u_d^{\alpha_d - 1}.$$

Leads to

$$h(\mathbf{w}) \;=\; \prod_{d=1}^{D} \frac{\alpha_d}{\Gamma(\alpha_d)} \cdot \frac{\Gamma(\sum \alpha_d + 1)}{(\sum \alpha_d w_d)^{D+1}} \prod_{d=1}^{D} \left(\frac{\alpha_d w_d}{\sum \alpha_k w_k} \right)^{\alpha_d - 1}.$$

Disadvantage: need for numerical integration

ESTIMATION

- Parametric

- Non/semi-parametric

Both approaches have problems, e.g. nonregular behavior of MLE even in finite-parameter problems; curse of dimensionality if D large.

Typically proceed by transforming margins to unit Fréchet first, though there are advantages in joint estimation of marginal distributions and dependence structure (Shi 1995)

III. ALTERNATIVE FORMULATIONS OF MULTIVARIATE EXTREMES

- Ledford-Tawn-Ramos approach

- Heffernan-Tawn approach

However the Heffernan-Tawn approach will be the subject of a later talk so we concentrate here on the Ledford-Tawn-Ramos approach

The first paper to suggest that multivariate extreme value theory (as defined so far) might not be general enough was Ledford and Tawn (1996).

Suppose (Z_1, Z_2) are a bivariate random vector with unit Fréchet margins. Traditional cases lead to

$$\text{Pr}\{Z_1 > r,\ Z_2 > r\} \sim \begin{cases} \text{const.} \times r^{-1} & \text{dependent cases} \\ \text{const.} \times r^{-2} & \text{exact independent} \end{cases}$$

The first case covers all bivariate extreme value distributions except the independent case. However, Ledford and Tawn showed by example that for a number of cases of practical interest,

$$\text{Pr}\{Z_1 > r,\ Z_2 > r\} \sim \mathcal{L}(r) r^{-1/\eta},$$

where $\mathcal{L}$ is a slowly varying function ($\frac{\mathcal{L}(rx)}{\mathcal{L}(r)} \to 1$ as $r \to \infty$) and $\eta \in (0, 1]$.

Estimation: used fact that $1/\eta$ is Pareto index for $\min(Z_1, Z_2)$.

More general case (Ledford and Tawn 1997):

$$\Pr\{Z_1 > z_1, Z_2 > z_2,\} \;=\; \mathcal{L}(z_1, z_2) z_1^{-c_1} z_2^{-c_2},$$

$0 < \eta \leq 1$; $c_1 + c_2 = \frac{1}{\eta}$; $\mathcal{L}$ slowly varying in sense that

$$g(z_1, z_2) \;=\; \lim_{t \to \infty} \frac{\mathcal{L}(tz_1, tz_2)}{\mathcal{L}(t,t)}.$$

They showed $g(z_1, z_2) = g_* \left(\frac{z_1}{z_1 + z_2}\right)$ but were unable to estimate g_* directly — needed to make parametric assumptions about this.

More recently, Resnick and co-authors were able to make a more rigorous mathematical theory using concept of *hidden regular variation* (see e.g. Resnick 2002, Maulik and Resnick 2005, Heffernan and Resnick 2005; see also Section 9.4 of Resnick (2007)).

Ramos-Ledford (2007a, 2007b) approach:

$$\begin{aligned}
\text{Pr}\{Z_1 > z_1, Z_2 > z_2,\} &= \tilde{\mathcal{L}}(z_1, z_2)(z_1 z_2)^{-1/(2\eta)}, \\
\lim_{u \to \infty} \frac{\tilde{\mathcal{L}}(ux_1, ux_2)}{\tilde{\mathcal{L}}(u, u)} &= g(x_1, x_2) \\
&= g_*\left(\frac{x_1}{x_1 + x_2}\right)
\end{aligned}$$

Limiting joint survivor function

$$\begin{aligned}
\text{Pr}\{X_1 > x_1, X_2 > x_2\} &= \int_0^1 \eta \left\{\min\left(\frac{w}{x_1}, \frac{1-w}{x_2}\right)\right\}^{1/\eta} dH_\eta^*(w), \\
1 &= \eta \int_0^1 \{\min(w, 1-w)\}^{1/\eta} dH_\eta^*(w).
\end{aligned}$$

Multivariate generalization to $D > 2$:

$$\Pr\{X_1 > x_1, ..., X_D > x_D\} = \int_{\mathcal{S}_D} \eta \left\{ \min_{1 \le d \le D} \left(\frac{w_d}{x_d} \right) \right\}^{1/\eta} dH_\eta^*(w),$$

$$1 = \int_{\mathcal{S}_D} \eta \left\{ \min_{1 \le d \le D} w_j \right\}^{1/\eta} dH_\eta^*(w).$$

Open problems:

- How to find sufficiently rich classes of $H_\eta^*(w)$ to derive parametric families suitable for real data

- How to do the subsequent estimation

IV. MAX-STABLE PROCESSES

Suppose $Y_t,\ t \in \mathcal{T}$ is a stochastic process (e.g. spatial, temporal)

Normalization condition: for each $n \geq 1$, there exist constants $a_{nt},\ b_{nt},\ t \in \mathcal{T}$ such that, for any $t_1, ..., t_m \in \mathcal{T}$,

$$\text{Pr}^n \left\{ \frac{Y_{t_j} - b_{nt_j}}{a_{nt_j}} \leq x_{t_j},\ j = 1, ..., m \right\} \ \rightarrow \ G_{t_1,...,t_m}(x_{t_1}, ..., x_{t_m}). \quad (5)$$

Then $G_{t_1,...,t_m}(x_{t_1}, ..., x_{t_m})$ is a MEVD. When (5) holds in a stochastic process sense the limiting process is called *max-stable*.

For this section of the talk, I describe a particular class of max-stable processes called *M4 processes* (Smith and Weissman (1996),Zhang (2002),Smith (2003) Chamú Morales (2005),Zhang and Smith (2007)). Alternative approaches will be described in the talk by D. Cooley.

Suppose $\{X_{id},\ i=0,\pm 1,\pm 2,...,\ d=1,...,D\}$ form a stationary time series in D dimensions.

Suppose we transform each X_{id} to a variable Y_{id} which is unit Fréchet ($P\{Y \le y\} = e^{-1/y}$). For example, a common method would be to fit a GPD to all exceedances of X_{id} above a high threshold, use the empirical distribution function for all values below the threshold, and apply the probability integral transformation.

The process $\{Y_{id}\}$ is called a *multivariate maxima of moving maxima* process (M4 for short), if it is defined by

$$Y_{id} \ = \ \max_{\ell} \max_{k} a_{\ell,k,d} Z_{\ell,i-k}.$$

where $a_{\ell,k,d} \ge 0$ are coefficients such that $\sum_{\ell}\sum_{k} a_{\ell,k,d} = 1$ for each d, and $\{Z_{\ell,i}, \ell = 1,...,L,\ i = 0,\pm 1,\pm 2,...,\}$ is a double array of independent unit Fréchet random variables.

Elementary Properties

$$\begin{aligned}
\Pr\{Y_{id} \le y_{id} \text{ for all } i,d\} &= \Pr\{a_{\ell,k,d} Z_{\ell,i-k} \le y_{id} \text{ for all } \ell,k,i,d\} \\
&= \Pr\{a_{\ell,k,d} Z_{\ell,i} \le y_{i+k,d} \text{ for all } \ell,k,i,d\} \\
&= \Pr\{Z_{\ell,i} \le \min_{k,d} \frac{y_{i+k,d}}{a_{\ell,k,d}} \text{ for all } \ell,i\} \\
&= \exp\left(-\sum_{\ell}\sum_{i} \max_{k,d} \frac{a_{\ell,k,d}}{y_{i+k,d}}\right).
\end{aligned}$$

So

$$\Pr\{Y_{id} \le n y_{id} \text{ for all } i,d\}^n = \Pr\{Y_{id} \le y_{id} \text{ for all } i,d\} \qquad (6)$$

and for a particular (i,d),

$$\Pr\{Y_{id} \le y_{id}\} = \exp\left(-\frac{\sum_\ell \sum_k a_{\ell,k,d}}{y_{i,d}}\right) = \exp\left(-\frac{1}{y_{i,d}}\right). \qquad (7)$$

(7) shows that the margins are unit Fréchet while (6) shows that the joint distributions are *max-stable*.

If we ignore the time series component and just treat this as a model for a single random vector $(Y_1, ..., Y_D)$, the equation simplifies to

$$\Pr\{Y_d \le y_d,\ d = 1, ..., D\} \;=\; \exp\left(-\sum_{\ell} \max_{d} \frac{a_{\ell,d}}{y_d}\right). \tag{8}$$

Recall the formula for the spectral measure:

$$\Pr\{Y_d \le y_d,\ d = 1, ..., D\} \;=\; \exp\left\{\int_{\mathcal{S}_D} \max_{j=1,...,d} \left(\frac{w_j}{y_j}\right) dH(w)\right\}$$

where $\mathcal{S}_D$ is simplex. Thus, (8) is equivalent to defining a spectral measure H that is discrete with mass at points

$$\frac{(a_{\ell,1}, ..., a_{\ell,D})}{\sum_d a_{\ell,d}} \in \mathcal{S}_D.$$

The condition $\int_{\mathcal{S}_D} w_d H(w) = 1$, for $d = 1, ..., D$, is equivalent to the statement that $\sum_\ell a_{\ell,d} = 1$ for all d.

Difficulties of Estimation

$$Y_{id} = \max_{\ell} \max_{k} a_{\ell,k,d} Z_{\ell,i-k}.$$

Problem of *signature patterns*. If a specific $Z_{\ell,k}$ is much larger than its neighbors, then for i close to k,

$$Y_{id} \propto a_{\ell,i-k,d}.$$

Under the model, this pattern will be replicated throughout the process. Such patterns cannot be expected in real data, so we must devise methods of estimation that take account of the fact that the M4 process is not an exact model.

Davis and Resnick (1989, 1993) developed many properties of "max-ARMA" processes but did not devise a good statistical approach.

Hall, Peng and Yao (2002) proposed an alternative estimation scheme for max-ARMA processes, based on the empirical distribution distribution.

Chamú Morales (PhD thesis, 2005) developed an approximate MCMC technique

Zhang and Smith (2007) (but originally Zhang (2002)) generalized the Hall-Peng-Yao technique to M4 processes

APPLICATION OF M4 PROCESSES TO FINANCIAL DATA

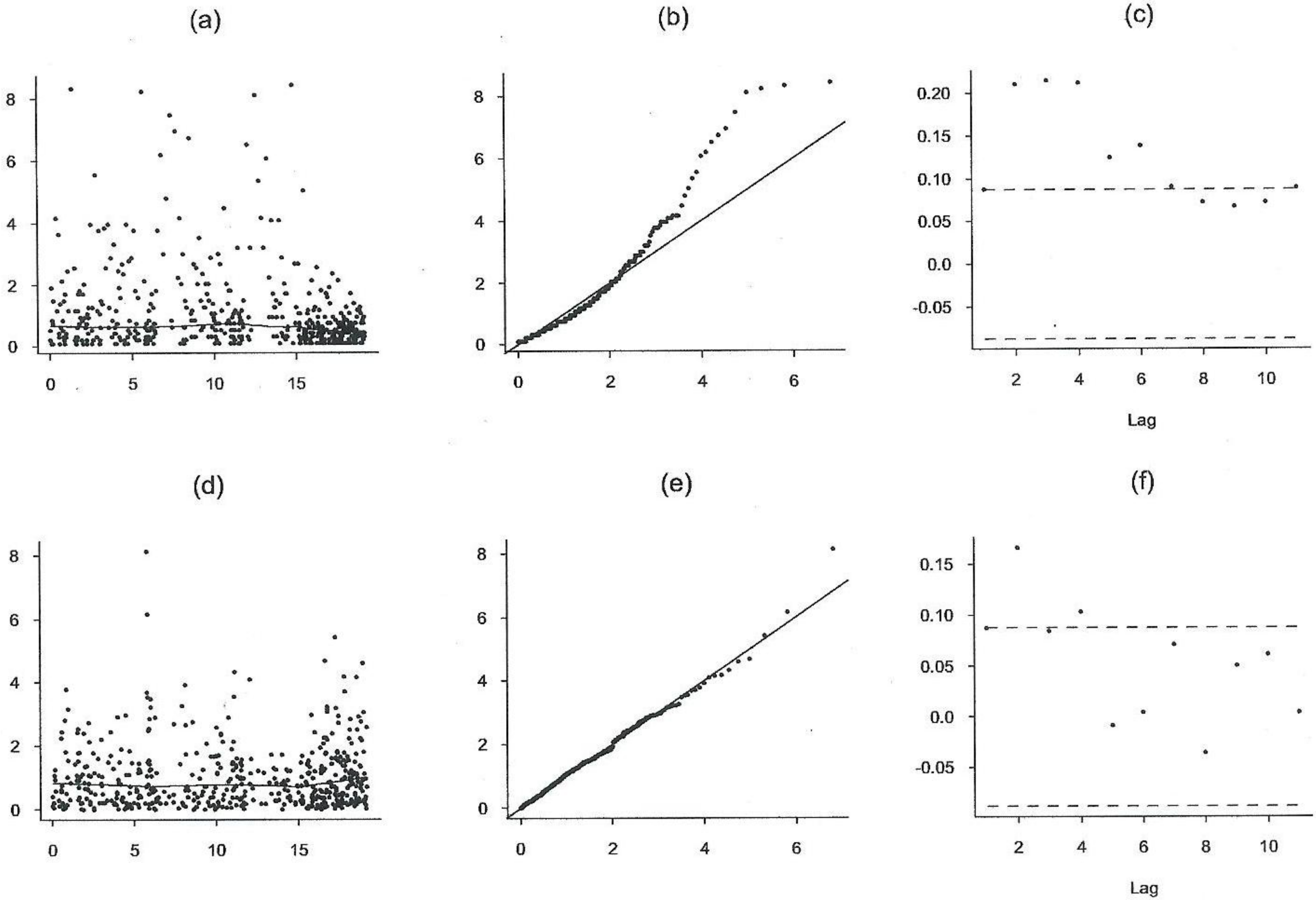
(Smith 2003)

We use neg daily returns from closing prices of 1982-2001 stock prices in three companies, Pfizer, GE and Citibank.

Univariate EVT models were fitted to threshold $u = .02$ with results:

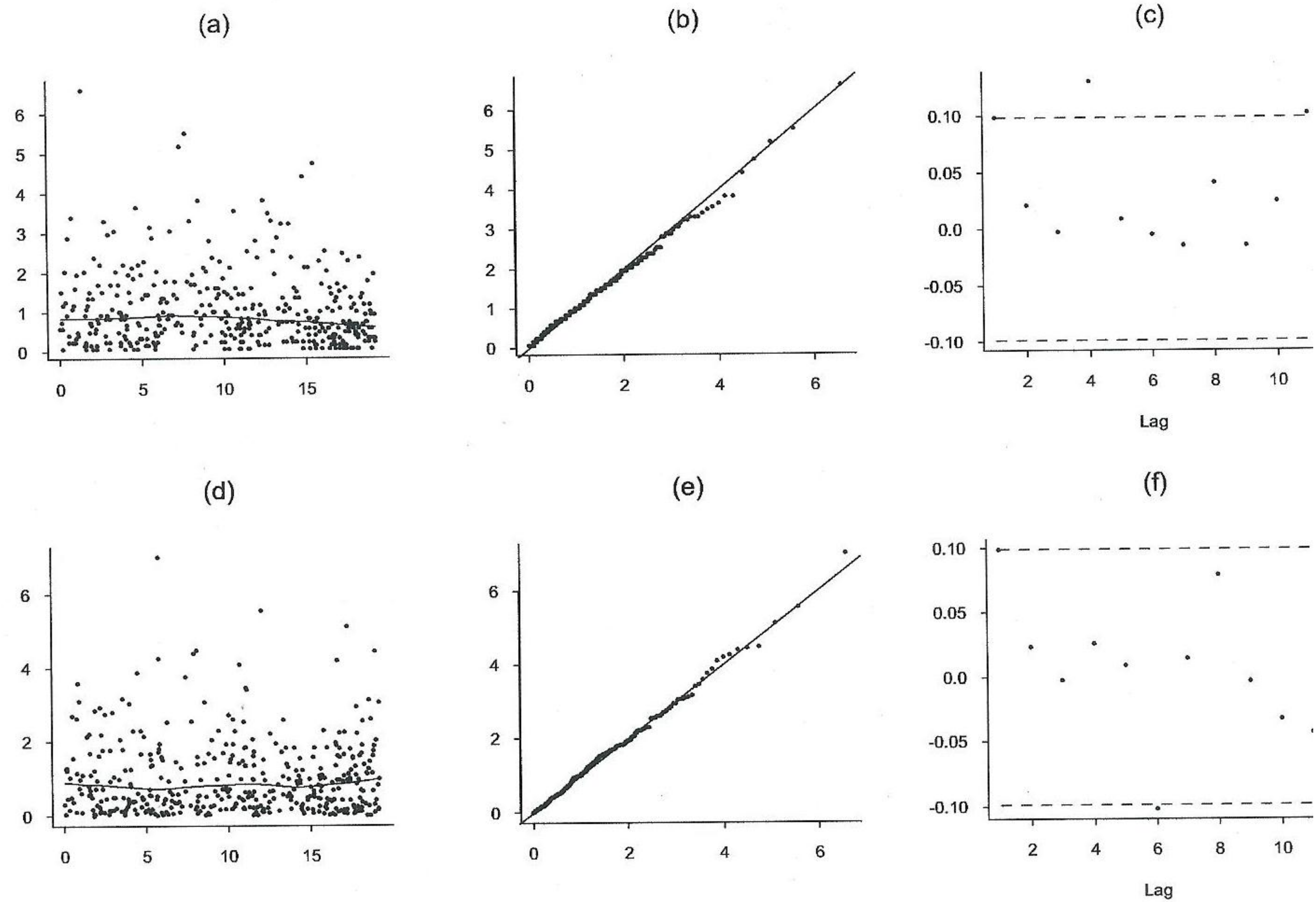
Series	N_u	μ (SE)	$\log \psi$ (SE)	ξ (SE)
Pfizer	518	.0623 (.0029)	−4.082 (.132)	.174 (.051)
GE	336	.0549 (.0029)	−4.139 (.143)	.196 (.062)
Citibank	587	.0743 (.0036)	−3.876 (.119)	.164 (.012)

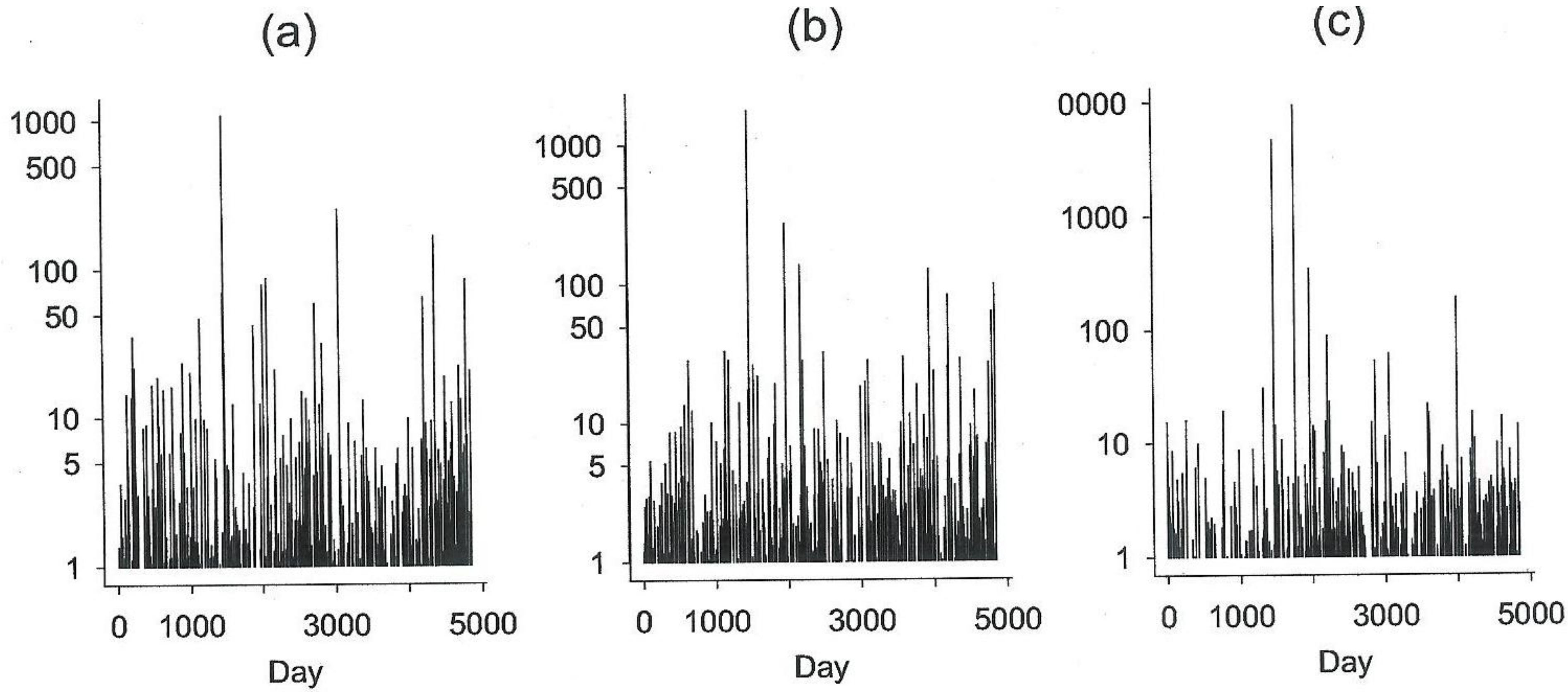
However, diagnostic plots show a problem — sign of volatility. (See in particular plot (b) — QQ plot based on intervals between threshold exceedances — these are clearly inconsistent with a simple Poisson processes.)


(a)
(b)
(c)
0.20
0.15
0.10
0.05
0.0
-0.05
Lag
(d)
(e)
(f)
0.15
0.10
0.05
0.0
-0.05
Lag

Estimate volatility by fitting a GARCH(1,1) model to each series. Standardize returns data, recompute threshold analysis ($u = 1.2$). Diagnostic plots now look OK (Pfizer shown, others similar)

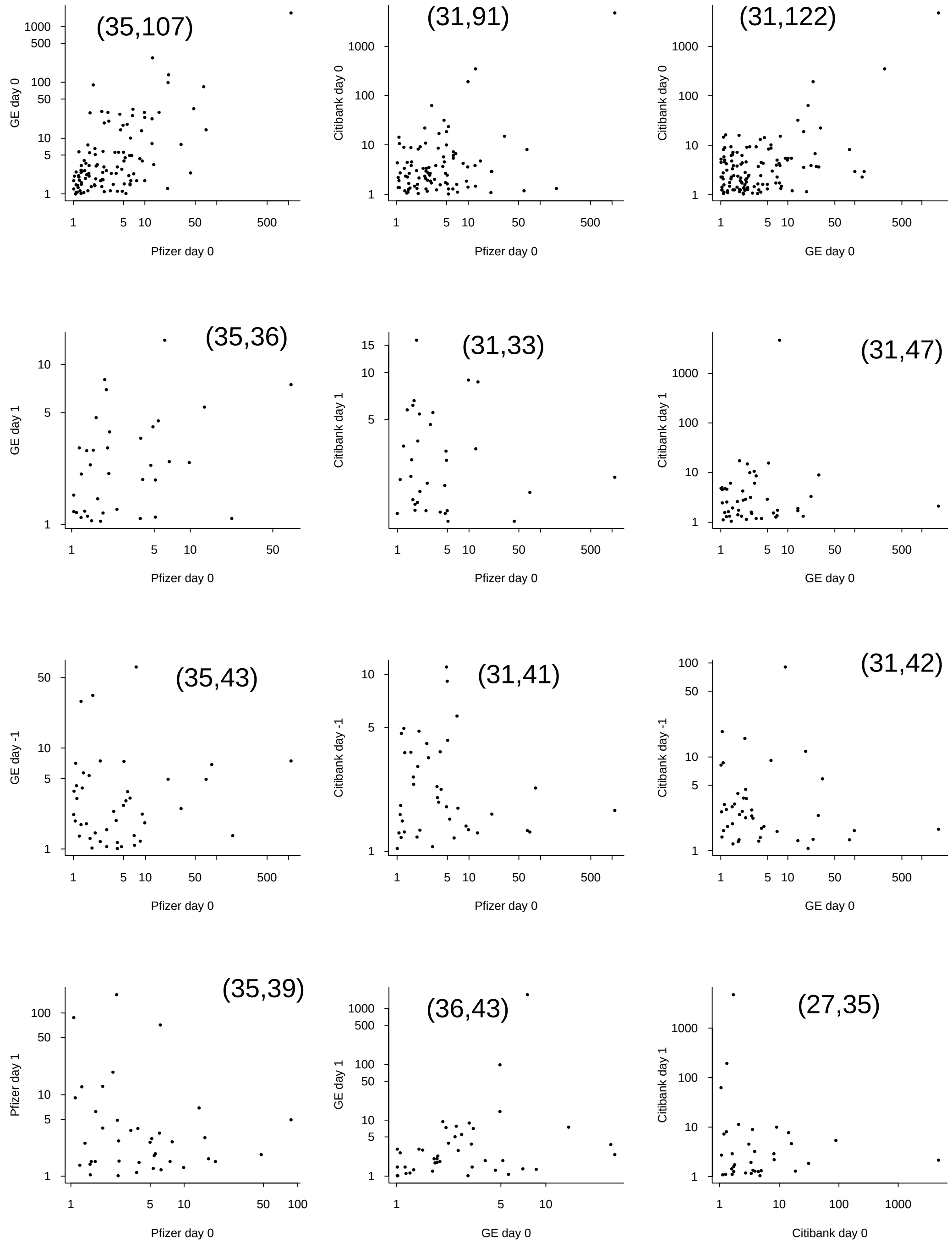
Series	N_u	μ (SE)	$\log \psi$ (SE)	ξ (SE)
Pfizer	411	3.118 (.155)	−.177 (.148)	.200 (.061)
GE	415	3.079 (.130)	−.330 (.128)	.108 (.053)
Citibank	361	3.188 (.157)	−0.118 (.126)	.194 (.050)


(a)
(b)
(c)
(d)
(e)
(f)
0 5 10 15
0 2 4 6
0.10
0.05
0.0
-0.05
-0.10
2 4 6 8 10
Lag

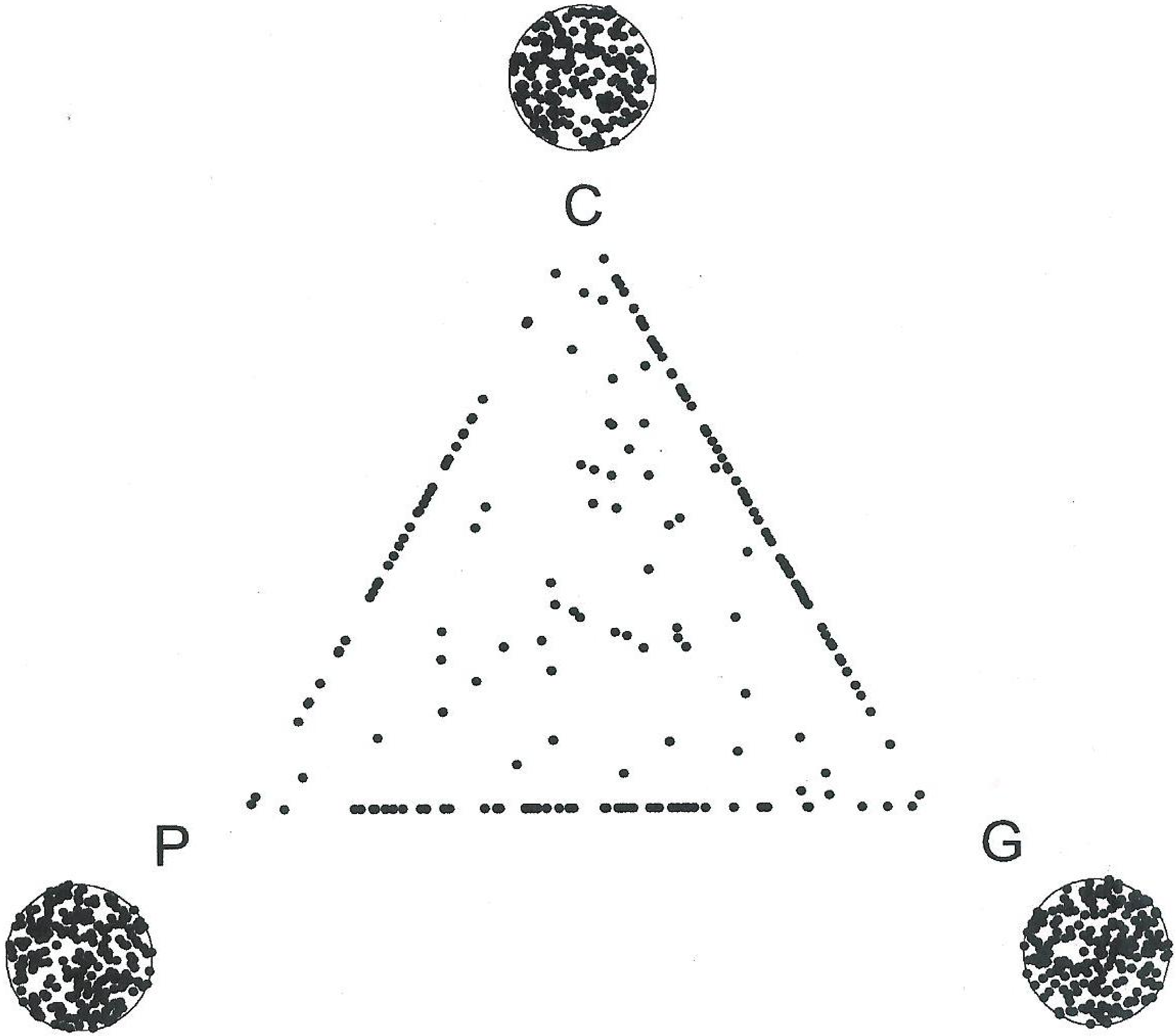

(a)
1000
500
100
50
10
5
1
0
1000
3000
5000
Day
(b)
1000
500
100
50
10
5
1
0
1000
3000
5000
Day
(c)
0000
1000
100
10
1
0
1000
3000
5000
Day

After fitting a univariate extreme value model to each series, the exceedances over the threshold for each series are transformed to have marginal Fréchet distributions. On the transformed scale, the data in each series consist of all values in excess of threshold 1.

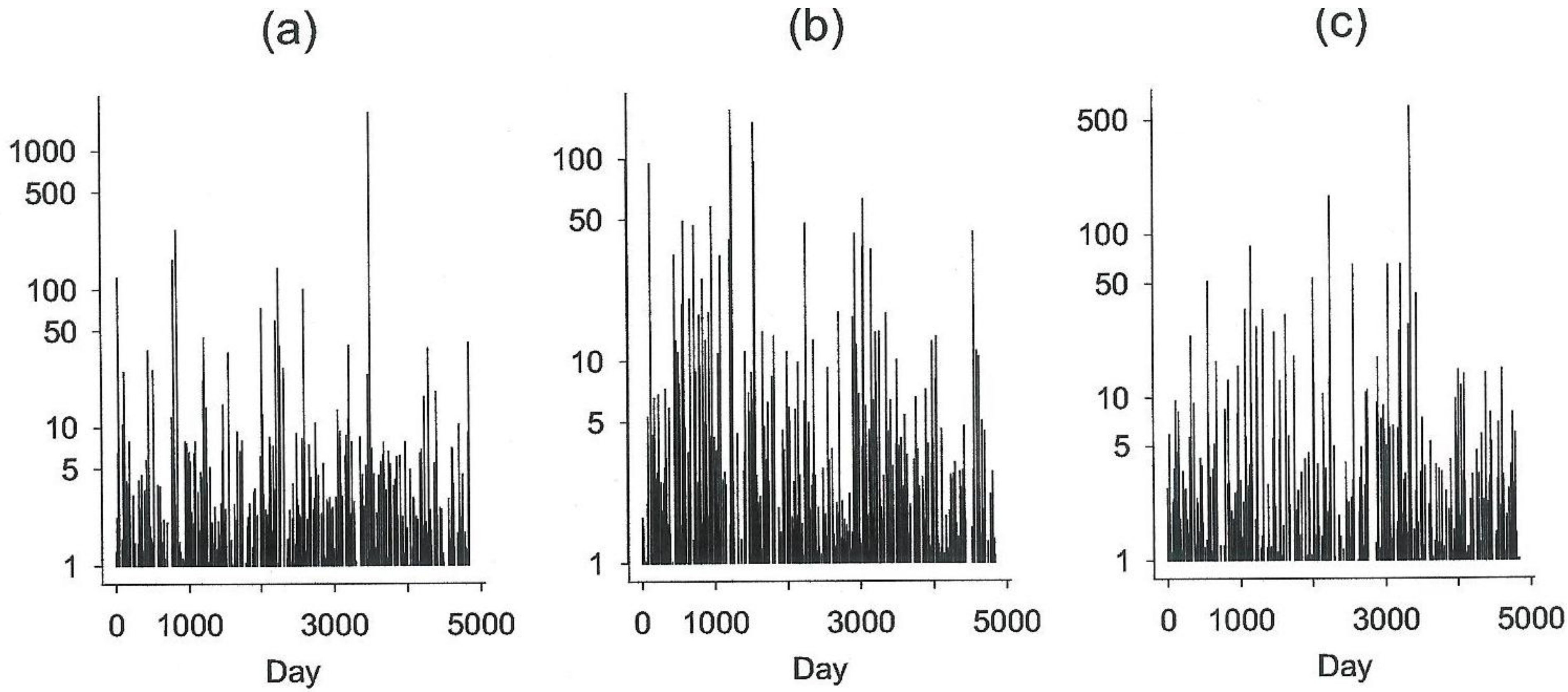
On this transformed scale, pairwise scatterplots are shown of the three series against each other on the same day (top 3 plots), and against series on neighboring days. The two numbers on each plot show the expected number of joint exceedances based on an independence assumption, and the observed number of joint exceedances.


(35,107)
GE day 0
Pfizer day 0
(31,91)
Citibank day 0
Pfizer day 0
(31,122)
Citibank day 0
GE day 0
(35,36)
GE day 1
Pfizer day 0
(31,33)
Citibank day 1
Pfizer day 0
(31,47)
Citibank day 1
GE day 0
(35,43)
GE day -1
Pfizer day 0
(31,41)
Citibank day -1
Pfizer day 0
(31,42)
Citibank day -1
GE day 0
(35,39)
Pfizer day 1
Pfizer day 0
(36,43)
GE day 1
GE day 0
(27,35)
Citibank day 1
Citibank day 0

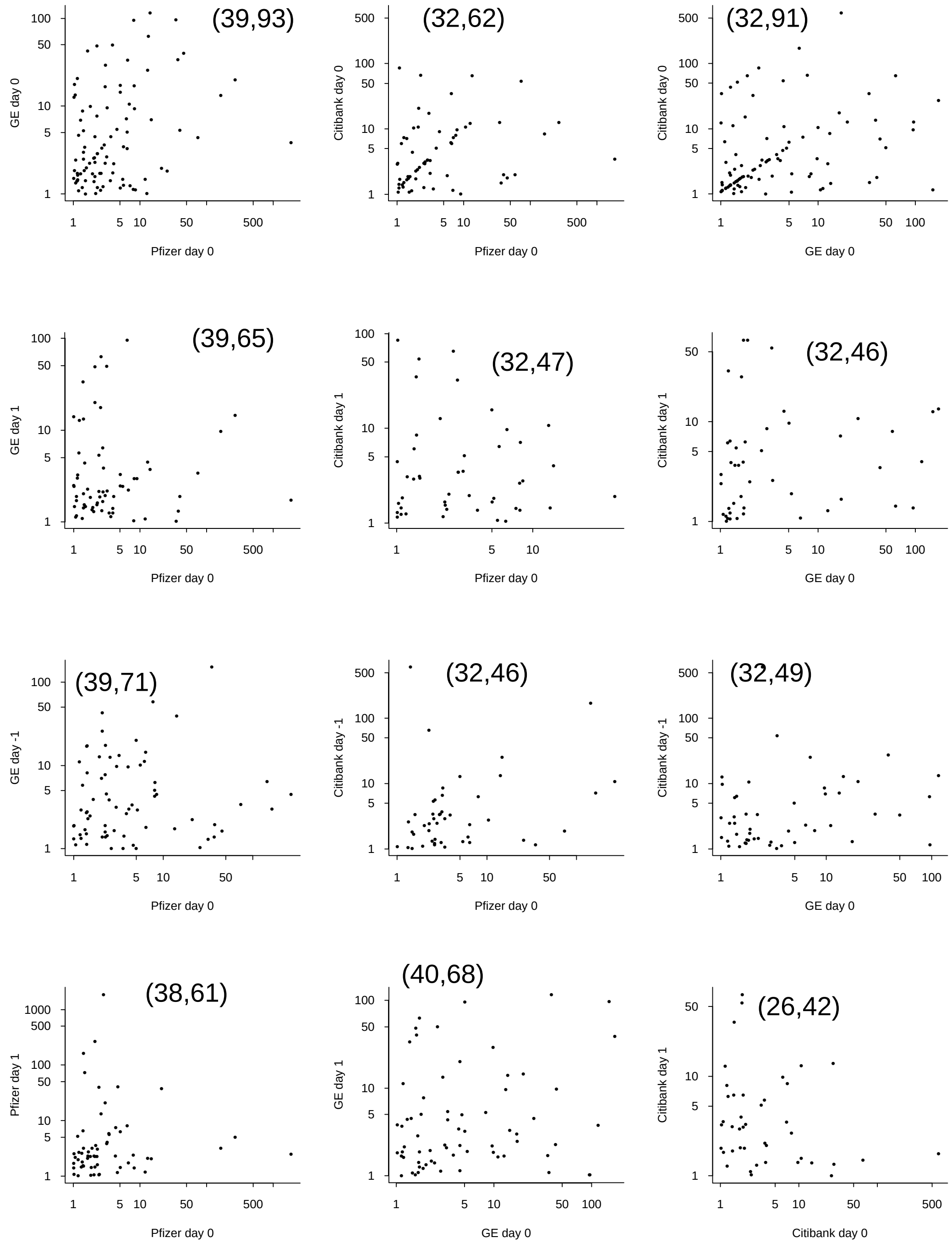
Also shown is a plot of Fréchet exceedances for the three series on the same day, normalized to have total 1, plotted in barycentric coordinates. The three circles near the corner points P, G and C correspond to days for which that series along had an exceedance.

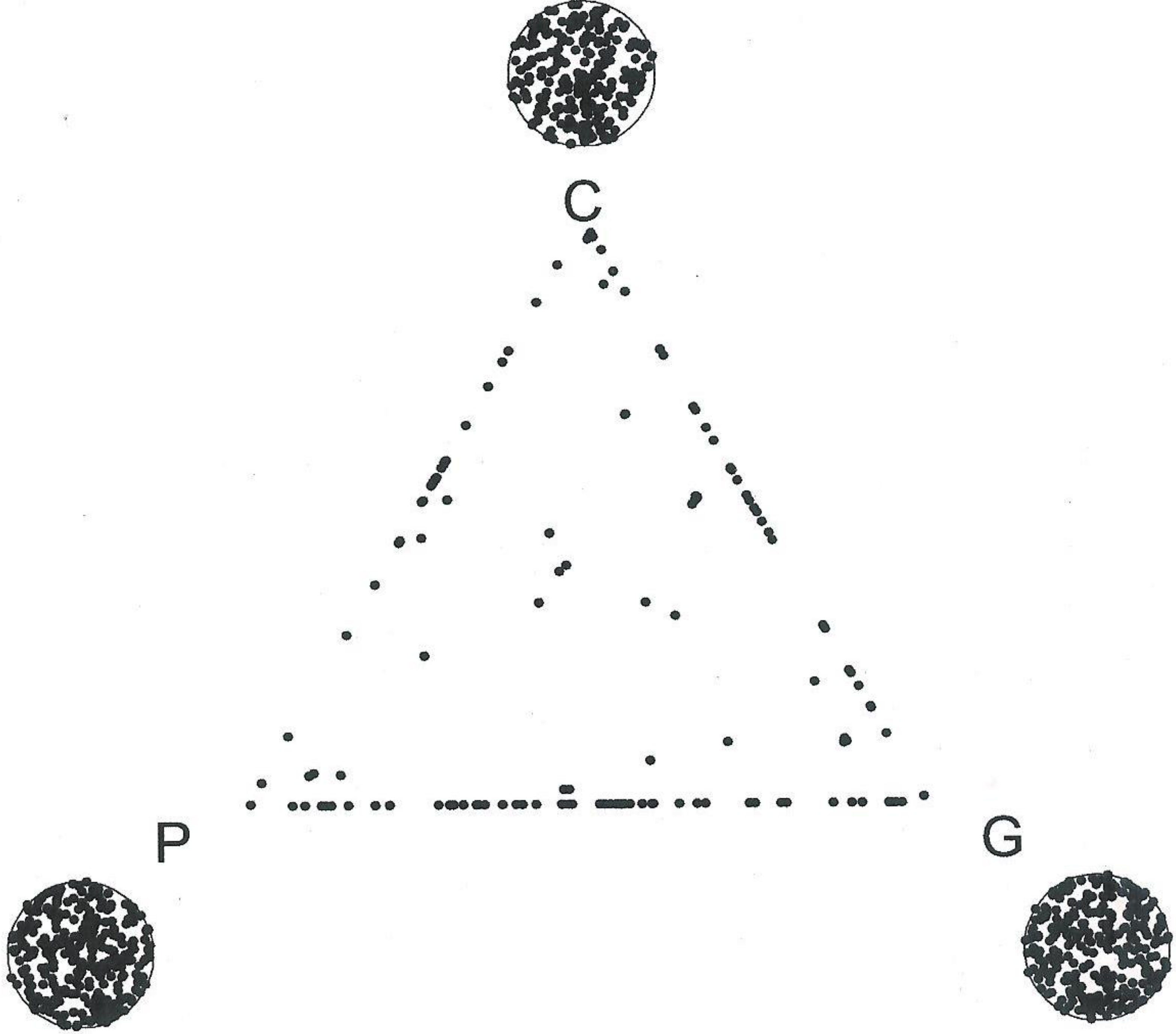

C
P
G

After fitting an M4 model, a new data set was simulated using the fitted model. The following figures correspond to the earlier ones, but using simulated data. This is a check on the realism of the fitted model.



(a)
(b)
(c)
1000
500
100
50
10
5
1
0
1000
3000
5000
Day


(39,93)
GE day 0
Pfizer day 0
(32,62)
Citibank day 0
Pfizer day 0
(32,91)
Citibank day 0
GE day 0
(39,65)
GE day 1
Pfizer day 0
(32,47)
Citibank day 1
Pfizer day 0
(32,46)
Citibank day 1
GE day 0
(39,71)
GE day -1
Pfizer day 0
(32,46)
Citibank day -1
Pfizer day 0
(32,49)
Citibank day -1
GE day 0
(38,61)
Pfizer day 1
Pfizer day 0
(40,68)
GE day 1
GE day 0
(26,42)
Citibank day 1
Citibank day 0

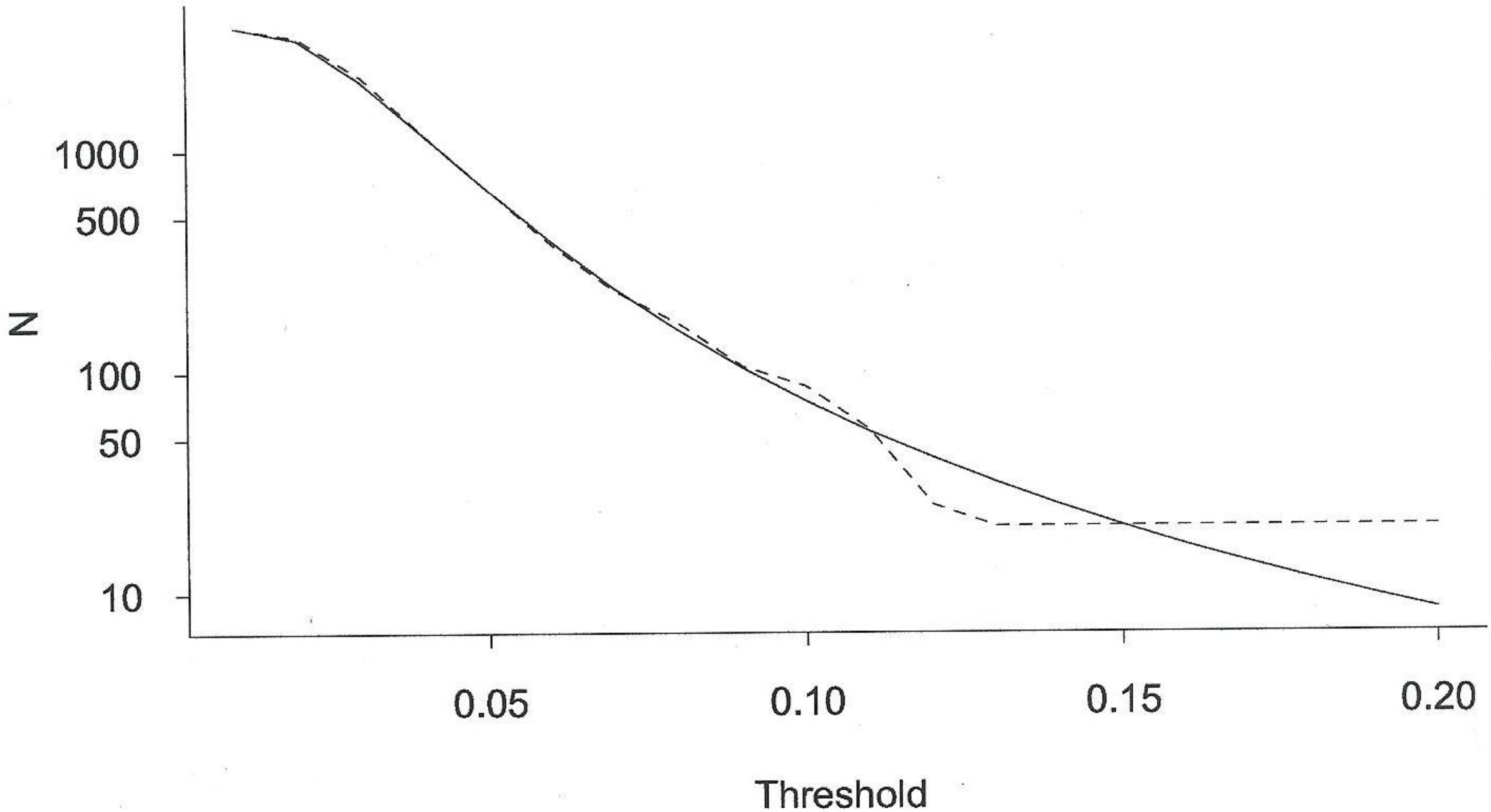

C
P
G

Finally, we attempt to validate the model by calibrating observed *vs.* expected probabilities of extreme events under the model.

The "extreme event" considered is that there is at least one exceedance of a specific threshold u by one of the three series in one of the next 10 days after a given day.

To make the comparison honest, the period of study is divided into four periods each of length just uder 5 years. The univariate and multivariate EV model is fitted to each of the first three 5-year period, and used to predict extreme events in the following period.

The final plot shows observed (dashed lines) and expected (solid lines) counts for a sequence of thresholds u. There is excellent agreement between the two curves.



1000
500
N
100
50
10
0.05
0.10
0.15
0.20
Threshold

REFERENCES

Chamú Morales, F. (2005), Estimation of max-stable processes using Monte Carlo methods with applications to financial risk assessment. Ph.D. dissertation, Department of Statistics and Operations Research, University of North Carolina, Chapel Hill. http://www.stat.unc.edu/postscript/rs/ChamuDissFinal.pdf

Coles, S.G. (2001), *An Introduction to Statistical Modeling of Extreme Values.* Springer Verlag, New York.

Coles, S.G. and Tawn, J.A. (1991), Modelling extreme multivariate events. *J.R. Statist. Soc. B* **53**, 377–392.

Davis, R.A. and Resnick, S.I. (1989), Basic properties and prediction of max-ARMA processes. *Advances in Applied Probability* **21**, 781–803.

Davis, R.A. and Resnick, S.I. (1993), Prediction of stationary max-stable processes. *Annals of Applied Probability* **3**, 497–525.

Finkenstadt, B. and Rootzén, H. (editors) (2003), *Extreme Values in Finance, Telecommunications and the Environment.* Chapman and Hall/CRC Press, London.

Gumbel, E.J. and Goldstein, N. (1964), Analysis of empirical bivariate extremal distributions. *J. Amer. Statist. Soc.* **59**, 794–816.

Hall, P., Peng, L. and Yao, Q. (2002), Moving-maximum models for extrema of time series. *Journal of Statistical Planning and Inference* **103**, 51–63.

Heffernan, J.E. and Resnick, S. (2005), Hidden regular variation and the rank transform. *Advances in Applied Probabilitty* **37-2**, 393–414.

Joe, H. (1990), Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statist. and Probab. Letters* **9**, 75–81.

Ledford, A.W. and Tawn, J.A. (1996), Modelling dependence within joint tail regions, *J.R. Statist. Soc. B* **59**, 475–499

Ledford, A.W. and Tawn, J.A. (1997), Statistics for near independence in multivariate extreme values. *Biometrika* **83**, 169–187.

Maulik, K. and Resnick, S. (2005), Characterizations and examples of hidden regular variation. *Extremes* **7**, 31–67

Ramos, A. and Ledford, A. (2007a), A new class of models for bivariate joint tails. Preprint.

Ramos, A. and Ledford, A. (2007b), An alternative point process framework for modeling multivariate extreme values. Preprint.

Resnick, S. (2002), Second order regular variation and asymptotic independence. *Extremes* **5**, 303–336

Resnick, S. (2007), *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.

Shi, D. (1995), Fisher information for a multivariate extreme value distribution. *Biometrika* **82**, 644–649.

Smith, R.L. (2003), Statistics of extremes, with applications in environment, insurance and finance. Chapter 1 of *Extreme Values in Finance, Telecommunications and the Environment*, edited by B. Finkenstädt and H. Rootzén, Chapman and Hall/CRC Press, London, pp. 1–78.

Smith, R.L. and Goodman, D. (2000), Bayesian risk analysis. Chapter 17 of *Extremes and Integrated Risk Management*, edited by P. Embrechts. Risk Books, London, 235-251.

Smith, R.L. and Weissman, I. (1996), Characterization and estimation of the multivariate extremal index. Unpublished paper, available from http://www.stat.unc.edu/postscript/rs/extremal.pdf.

Tawn, J.A. (1990), Modelling multivariate extreme value distributions. *Biometrika* **77**, 245–253.

Zhang, Z. (2002), *Multivariate Extremes, Max-Stable Process Estimation and Dynamic Financial Modeling*. Ph.D. dissertation, Department of Statistics, University of North Carolina, Chapel Hill. http://www.stat.unc.edu/postscript/rs/zjz3.pdf

Zhang, Z. and Smith, R.L. (2007), On the estimation and application of max-stable processes. (Re-)Submitted for publication. http://www.stat.unc.edu/postscript/rs/zhangsmith07.pdf