A Method For Fitting A pRARMAX Model: An Application To Financial Data

Marta Ferreira and Luísa Canto e Castro *

Abstract—**Ferreira and Canto e Castro [6] introduces a power max-autoregressive process, in short pARMAX, as an alternative to heavy tailed ARMA. An extension of pARMAX was considered in Ferreira and Canto e Castro [7], by including a random component, and hence called pRARMAX, which makes the model more flexible to applications. It was then developed a methodology settled on minimizing the Bayes risk in classification theory, but only considering standard uniform random components. We now extend this procedure to the more general *Beta* distribution. We illustrate the method with an application to a financial data series. In order to improve estimates of the exceedance probabilities of levels of interest, we use Bortot and Tawn [2] approach and derive a threshold-dependent extremal index which relates with the coefficient of tail dependence of Ledford and Tawn [8] and with the pRARMAX parameter.**

Keywords: Extreme value theory, max-autoregressive models, classification theory, Bayes error

1 Introduction

The Extreme Value Theory (EVT) has been increasingly used in areas such as finance, insurance, engineering, geophysics and telecommunications, due to the growing interest in the possibility of occurrence and impact of extreme events and the need to take them into account in modeling. Initially it was sustained in observations considered independent and identically distributed (i.i.d.), but recently, models for extreme values have been constructed under the more realistic assumption of temporal dependence. Among these, stationary Markov chains are very interesting, in particular the max-autoregressive ones due to a somewhat simple treatment with regard to extremal properties. The MARMA (*max-autoregressive moving average*) processes presented in Davis and Resnick [4], in particular the ARMAX or MARMA$(1,0)$ (Alpuim [1]), and their generalizations have applications in various phenomena, e.g., priority queues [4], accumulation of solar energy [3] and financial series [10].

In modeling dependence, it is important to assess if there is asymptotic tail independence (i.e., a dependence that gradually disappears at more and more extreme levels) or exact dependence. Ledford and Tawn [8] have proposed a model with a new parameter, usually denoted η, that measures the "degree" of tail dependence, known as *coefficient of asymptotic tail dependence*. When computing η for the above mentioned MARMA, another class of max-autoregressive processes arises: the power max-autoregressive pARMAX, which includes a power parameter c $(0 < c < 1)$, that is related with η (Ferreira and Canto e Castro [6]). More precisely, we have $\eta = \max(1/2, c)$ and hence, as $1/2 \le \eta < 1$, the process is asymptotically tail independent with positive association. There are several estimators for η with good properties (see, for instance [8]) and this allows us to obtain good estimates for the model parameter c. In order to make the pARMAX process more applicable to real data, it is considered a generalized version by including a random factor, denoted pRARMAX. More precisely, a sequence $\{X_i\}_{i\in\mathbb{Z}}$ is pRARMAX, if

$$X_i = U_i X_{i-1}^c \vee Z_i \;,\; 0 < c < 1,\; i \in \mathbb{Z}, \tag{1}$$

where $\{Z_i\}_{i\in\mathbb{Z}}$ and $\{U_i\}_{i\in\mathbb{Z}}$ are i.i.d. r.v.'s and independent of each other (if $U = 1$ we obtain pARMAX). For pRARMAX the same connection between the power parameter c and η holds. A sufficient condition for stationarity is to consider innovations $\{Z_i\}_{i\in\mathbb{Z}}$ in the Fréchet max-domain of attraction, which in turn leads to an unit extremal index, i.e., $\theta = 1$. See Ferreira and Canto e Castro [7] for details.

Example: Consider Z such that, $F_Z(x) = \frac{1-x^{-1/\gamma}}{1-(B(\frac{1}{c\gamma}+p,q)/B(p,q))x^{-1/(c\gamma)}}\mathbf{1}_{\{x\ge 1\}}$, where $B(p,q)$ is the Euler Beta function, and $U \frown Beta(p,q)$, $p,q > 0$. Then, $K(x) = \left(1 - x^{-1/\gamma}\right)\mathbf{1}_{\{x\ge 1\}}$, is non-degenerate stationary distribution of X_i.

Another interesting feature of pRARMAX is that, because of an asymptotically tail independent behavior jointly with an unit extremal index, it presents a thinning of clusters of extremes as the threshold increases, until exceedances occur singly. In such cases, there is an advantage (for inferential purposes) if θ is replaced by a pre-asymptotic extremal index on the approximation,

$$P\left(\bigvee_{i=1}^{n} Y_i \le q\right) \approx P(Y_1 \le q)^{n\theta}, \tag{2}$$

for large n and q. Based on Bortot and Tawn [2] approach, we consider a threshold-dependent extremal index,

$$1-\theta(u) \sim (t(u))^{1-1/\eta} L(t(u)),\ \text{with}\ t(u) = (1-K(u))^{-1}, \tag{3}$$

as $u \to \infty$, where L is a slowly varying function. Observe that $\theta(u)$ is a functional of the coefficient of tail dependence, η,

*31 March 2010; University of Minho, /DMAT-CMAT, Braga - Portugal Email: msferreira@math.uminho.pt & Faculdade de Ciências da Universidade de Lisboa/DEIO-CEAUL, Lisboa - Portugal Email: ldloura@fc.ul.pt

ISBN: 978-988-18210-8-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

given by $\eta = \max(1/2, c)$, hence relates with model parameter c. More precisely, approximation (3) is a consequence of having $\frac{\sum_{j=3}^{r_n} P(X_1 > u_n, \bigvee_{i=2}^{j-1} X_i \leq u_n, X_j > u_n)}{P(X_1 > u_n, X_2 > u_n)} \to 0$, as $n \to \infty$, where $(u_n)_n$ is a real sequence such that $n(1-K(u_n)) \to \tau > 0$, and $(r_n)_{n\geq 1}$ a nondecreasing integer sequence with $r_n = o(n)$, as $n \to \infty$. In addition, we have that, $P(\bigvee_{i=1}^{n} X_i \leq u_n) - K^n(u_n) = O\big(n^{1-1/c}\mathscr{L}(a_{\tau/n}^{1/c})\big)$, where $\mathscr{L}$ is some slowly varying function, and also, $P(\bigvee_{i=1}^{n} X_i \leq u_n) - K^{n\theta(u_n)}(u_n) = o\big(n^{1-1/c}\mathscr{L}(a_{\tau/n}^{1/c})\big)$, with $\theta(u) \equiv \theta(u, r_{[u]})$ given in (3). Therefore, replacing the unit θ by $\theta(u)$ in approximation (2) leads to an improvement of this latter. For details, see Ferreira and Canto e Castro [7].

Making use of pRARMAX flexibility, a methodology for assessing the adjustment of this model to real data was developed in [7]. This procedure is based on minimizing the Bayes risk in classification theory and it had been only considered for $U_i \frown$U(0,1). In this paper, we will show that it easily extends to $U_i \frown$Beta(p,q), $p,q > 0$ (includes U(0,1)). The method is applied to a financial series (*S&P500* log-returns) and we conclude for the goodness-of-fit of the model. We estimate exceeding probabilities of high levels considered as risky amounts.

2 How to fit a pRARMAX model

In the following, we present a summary of the method. For details, see Ferreira and Canto e Castro [7]. The pRARMAX is a suitable model for any given observed time series $\{X_i\}_{i\in\mathbb{Z}}$, if $X_i = \max(U_i X_{i-1}^c, Z_i)$, that is, each X_i either comes from the first component or from the second component of the maximum. Considering $\mathtt{G}_0$ the set of X_i's that come from the second component (Z_i), and $\mathtt{G}_1$ the set of X_i's coming from the first component ($U_i X_{i-1}^c$), we will say that the model fits if, for the observations in $\mathtt{G}_0$ the assumptions considered for Z are not rejected, and for the observations in $\mathtt{G}_1$, when divided by X_{i-1}^c, the hypotheses assumed for U are not rejected, as well. So we need to classify each observation as belonging to $\mathtt{G}_0$ or $\mathtt{G}_1$. If $X_i \geq X_{i-1}^c$, obviously $X_i \in \mathtt{G}_0$, but the case $X_i < X_{i-1}^c$ requires a rule-making, which has four possible outcomes with two of them being misclassifications that will be penalized. We apply *classification theory* based on a Bayesian solution that minimizes the risk of possible wrong decisions and this is conducted in an hypothesis tests context as in Storey [9]. Table 1 summarizes the procedure and lead us to the settlement of the *Bayes error* given by

$$BE(\Gamma) = (1-\lambda)P(T \in \Gamma, H=0) + \lambda P(T \notin \Gamma, H=1) \quad (4)$$

where Γ is the significance region of the associated hypothesis tests. None of the errors (type I or II) is fixed in advance, they can take any value as long as $BE(\Gamma)$ is minimum. Assuming $X_i|_{X_i<X_{i-1}^c, H_i} \frown (1-H_i)\cdot F_0 + H_i \cdot F_1$, F_0 and F_1 are the d.f.'s of the r.v.'s in $\mathtt{G}_0$ and in $\mathtt{G}_1$, respectively, with densities f_0 and f_1, the significance region

$$\mathscr{B}_\lambda = \Big\{t : \pi_0 f_0(t)/(\pi_0 f_0(t) + \pi_1 f_1(t)) \leq \lambda\Big\}, \quad (5)$$

with $P(H=1) = \pi_1 = 1-\pi_0$, minimizes the *Bayes error* in (4), for each λ ($0 \leq \lambda \leq 1$). Since our decision criterion only respects, $X_i < X_{i-1}^c$, we need to compute,

$$P(X_i \leq x | X_i < X_{i-1}^c) = \frac{P(X_i \leq x, X_i < X_{i-1}^c)}{P(X_i < X_{i-1}^c)} = \frac{F_1(x) + F_0(x)}{P(X_i < X_{i-1}^c)}, \quad (6)$$

with $F_1(x) = P(X_i \leq x, X_i < X_{i-1}^c, U_i X_{i-1}^c > Z_i)$ and $F_0(x) = P(X_i \leq x, X_i < X_{i-1}^c, U_i X_{i-1}^c \leq Z_i)$. For each fixed λ, we determine the significance region, $\mathscr{B}_\lambda$, defined in (5). In the following, we illustrate the calculations for the model pRARMAX of example above, where now the r.v. U is $Beta(p,q)$, $p,q > 0$. We have successively,

$$\begin{aligned} F_1(x) &= \int_0^1\!\!\int_1^x P\Big(\big(\tfrac{z}{u}\big)^{\frac{1}{c}} < X_{n-1} \leq \big(\tfrac{x}{u}\big)^{\frac{1}{c}}\Big) dF_Z(z) dF_U(u) \\ &= E(U^{\frac{1}{\gamma c}})\left[\int_1^x z^{-\frac{1}{\gamma c}} dF_Z(z) - x^{-\frac{1}{c\gamma}} F_Z(x)\right], \end{aligned}$$

$$\begin{aligned} F_0(x) &= \int_0^1 \int_1^x P\Big(z^{1/c} < X_{n-1} < \big(\tfrac{z}{u}\big)^{1/c}\Big) dF_Z(z) dF_U(u) \\ &= \frac{E\big(U^{1/(\gamma c)}\big)}{c\gamma} \int_1^x z^{-1/(c\gamma)} f_Z(z) dz, \end{aligned}$$

$$f_1(x) = \frac{E(U^{1/(\gamma c)})}{c\gamma} x^{-1/(c\gamma)-1} F_Z(x)$$

and

$$f_0(x) = \frac{E(U^{1/(\gamma c)})}{c\gamma} x^{-1/(c\gamma)} f_Z(x).$$

In what concerns, $\pi_1 = P(H=1) = P(X_i \in \mathtt{G}_1 | X_i < X_{i-1}^c)$, we have that,

$$\begin{aligned} \pi_1 &= \frac{P(U_i X_{i-1}^c > Z_i)}{P(X_{i-1}^c > Z_i)} = \frac{\int_0^1 \int_1^\infty (z/u)^{-1/(\gamma c)} dF_Z(z) dF_U(u)}{\int_1^\infty z^{-1/(\gamma c)} dF_Z(z)} \\ &= E\big(U^{1/(\gamma c)}\big)\big) \end{aligned} \quad (7)$$

Therefore, with $\pi_0 = 1-\pi_1$, the significance region (5) is given by,

$$\mathscr{B}_\lambda = \left\{t : \frac{\pi_0 t^{-1/(c\gamma)} f_Z(t)}{(\pi_0 t^{-1/(c\gamma)} f_Z(t) + (1-\pi_0) t^{-1/(c\gamma)-1} F_Z(t))} \leq \lambda\right\}. \quad (8)$$

Note that r.v. U in this context has distribution *Beta* conditional on $X_i = U_i X_{i-1}^c$ and on the criterion $U_i X_{i-1}^c > t_\lambda$, where t_λ is the critical value obtained in (8). Thus being,

$$\begin{aligned} & P(U_i \leq u | U_i X_{i-1}^c > t_\lambda, X_i = U_i X_{i-1}^c) \\ &= \frac{\int_1^\infty \int_1^\infty P\big(\frac{t_\lambda}{x^c} \vee \frac{z}{x^c} < U_i \leq u\big) dF_Z(z) dK(x)}{\int_1^\infty \int_1^\infty P\big(U_i > \frac{t_\lambda}{x^c} \vee \frac{z}{x^c}\big) dF_Z(z) dK(x)}, \end{aligned} \quad (9)$$

where, for the numerator, and taking $A = (t_\lambda/u)^{1/c}$, we obtain,

$$\int_A^\infty F_U(u) F_Z(ux^c) dK(x) - \int_A^\infty \int_1^{t_\lambda} F_U\big(\tfrac{t_\lambda}{x^c}\big) dF_Z(z) dK(x) - \int_A^\infty \int_{t_\lambda}^{ux^c} F_U\big(\tfrac{z}{x^c}\big) dF_Z(z) dK(x).$$

Since the denominator in (9) is derived from the previous expression by taking $u = 1$, after some algebra we have,

ISBN: 978-988-18210-8-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$P\big(U_n \leq u | U_n X_{n-1}^c > t_\lambda, X_n = U_n X_{n-1}^c\big)$$
$$= \frac{1}{I}\int_A^\infty x^{-c}\int_{t_\lambda}^{ux^c} f_U(z/x^c)F_Z(z)dzdK(x)$$

where, $I = \int_{t_\lambda^{1/c}}^\infty \big[\int_{t_\lambda}^{x^c} f_U(z/x^c)F_Z(z)dz\big]dK(x)$. Hence the density function is given by,

$$\frac{1}{I}\int_A^\infty f_U(u)F_Z(ux^c)dK(x)$$
$$= \frac{u^{1/(c\gamma)+p-1}(1-u)^{q-1}}{B(p,q)c\gamma I}\int_{t_\lambda}^\infty F_Z(y)y^{-1/(c\gamma)-1}dy. \qquad (10)$$

Therefore, the r.v.'s U_i captured by the criterion are $Beta\big(p+\frac{1}{c\gamma},q\big)$ distributed. Here is a summary of the steps to fit a pRARMAX model to a time series data:

1. Test if the given sample, $\mathbf{X} = (X_1, X_2, ..., X_n)$, is in the Fréchet max-domain of attraction and estimate the tail index, here denoted γ_X (e.g., Hill estimator);

2. Estimate parameter c of model pRARMAX through the estimation of η which is the tail index (γ_T) of $\mathbf{T^{(n)}} = (T_1^{(n)}, ..., T_{n-1}^{(n)})$, with $T_i^{(n)} = \min\big(\frac{n+1}{n+1-R_i}, \frac{n+1}{n+1-R_{i+1}}\big)$, $i = 1, ..., n$, where R_i is the rank of X_i among $(X_1, ..., X_n)$;

3. Based on the criterion: "if $X_i > X_{i-1}^{\widehat{c}}$ ($\widehat{c} = \widehat{\gamma}_T$, obtained in step 2.) then $X_i = Z_i$", separate the innovations, $\mathbf{Z}$, and test if this sample is also in the Fréchet max-domain of attraction;

4. Capture the observations corresponding to U, through the criterion: "if $X_i < X_{i-1}^{\widehat{c}}$ and $X_i \in \mathscr{B}_\lambda$ given in (8), where $\widehat{\gamma}_X$ and $\widehat{c}$ are the estimates obtained in steps 1. and 2., respectively, then, $U_i = X_i / X_{i-1}^{\widehat{c}}$"; λ must be chosen in order to minimize the Bayes error in (4) but also allowing to capture a reasonable number of "true values" of U ($\gtrsim 30$);

5. Test whether the sample of r.v.'s U captured in the previous step has distribution $Beta(1/(\widehat{\gamma}_X\widehat{c})+p,q)$ (use, for instance, the Kolmogorov-Smirnov test).

Table 1: Possible outcomes of a classification criterion along with an interpretation under an hypothesis test procedure with misclassification penalties λ.

	X_i classified in $\mathtt{G}_0$	X_i classified in $\mathtt{G}_1$
	H classified as 0	H classified as 1
X_i belongs to $\mathtt{G}_0$		
$H=0$	correct decision	type I error $(1-\lambda)$
X_i belongs to $\mathtt{G}_1$		
$H=1$	type II error (λ)	correct decision

2.1 An application to financial data

In financial markets often one has to decide on a big risky investment while cannot afford to have a loss larger than a certain amount. Hence, it may be of interest to know the probability that the maximum volatility exceeds that amount of risk. We will see that a pRARMAX process with Pareto marginals and random components, $U \frown$U(0,1), performs quite well in modeling the volatility of *S&P500* stock market index, through the implementation of the procedure outlined above. More precisely, we examine the square of the log-returns $R_i = \log P_{i+1}/P_i$, $1 \leq i \leq n-1$ ("volatility" can be measured through $|R_i|$ or, equivalently, R_i^2), where P_i denotes the index calculated at the end of the ith trading day in the years 1957 to 1987, amounting a sample size $n = 7733$. The series $\{R_i\}_i$ and $\{R_i^2\}_i$ are plotted in Figure 1, in which the large peak corresponds to Monday stock market crash on the 19th October 1987, known as "Black Monday". According to step

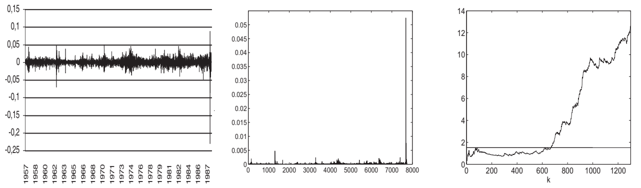

Figure 1: Daily log-returns (left), respective squares (middle) of *S&P500* stock market index and sample path of the extreme value condition test applied to the series $\{R_i^2\}_i$, (horizontal line: critical value above which reject $X \in \mathscr{D}(G_\gamma)_{\gamma \geq 0}$).

1., we test if the data present a heavy tail. In Figure 2 (top left) the horizontal line corresponds to the critical value above which we reject the extreme value condition. Hence it is not rejected for $k \lesssim 700$. The sample path of Hill estimator (Figure 2 top right) shows an upward trend, whereas the moment and maximum likelihood estimators in (bottom left and right respectively) are much more stable: from these latter we advance the estimate, $\widehat{\gamma}_X \approx 0.5$ (value where the paths yield an almost flat line; the argument for this value will be further strengthened ahead). In order to obtain a data series with standard Pareto marginals, a robust regression was implemented leading to a scale estimate, $a = 13618.3$, and a shift estimate, $b = 1.1$. Thus, our analysis focuses on the transformed data, $X_i = aR_i^2 + b$, ($1 \leq i \leq n$ with $n = 7733$). From now on we will refer to this data set as "$\mathbf{X}$".

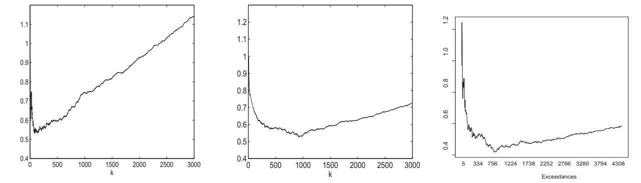

Figure 2: Sample paths of Hill (left), moment (middle) and maximum likelihood estimators (right) for the squared log-returns R_i^2.

We test again the extreme value condition which is not rejected for $165 \lesssim k \lesssim 900$ as shown in Figure 3 (top left). Considering the sample paths of Hill, moment and maximum likelihood estimators in Figure 3, the previous estimate, $\widehat{\gamma}_X \approx 0.5$, seems even more plausible (note that the behavior of the Hill estimator has changed completely yielding an almost flat line). To evaluate the effect on the tail of the large peak, we have considered the data until the day immediately before and obtained

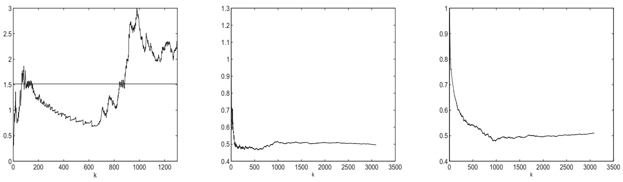

Figure 3: Left: sample path of the extreme value condition test applied to the transformed data $\mathbf{X}$, (horizontal line: critical value above which reject $X \in \mathscr{D}(G_\gamma)_{\gamma\geq 0}$); sample paths of Hill (middle) and moment (right) for $\mathbf{X}$.

$0.4 \lesssim \gamma_X \lesssim 0.45$.

According to step 2., we transform $\mathbf{X}$ into $\mathbf{T}^{(\mathbf{n})}$ and then estimate parameter c through the tail index of $\mathbf{T}^{(\mathbf{n})}$. Observing Figure 4, the estimate is about 0.85. However, due to some stability around 0.75, we consider, $\widehat{c} = 0.85$, $\widehat{c} = 0.8$ and $\widehat{c} = 0.75$. The innovations Z, captured on step 3., seem also to

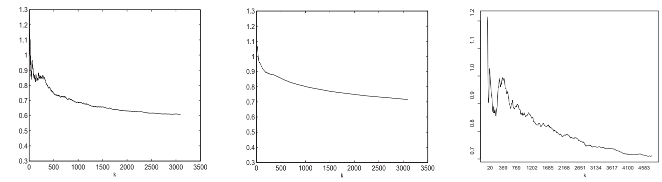

Figure 4: Sample paths of Hill (left), moment (center) and maximum likelihood (right) estimators, of the transformed sample $\mathbf{T}^{(\mathbf{n})}$ from $\mathbf{X}$.

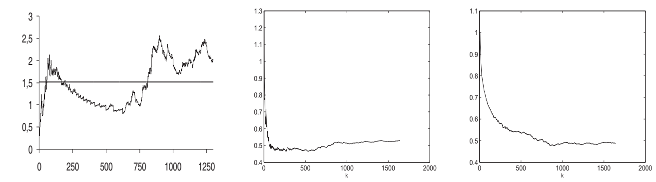

Figure 5: Left: sample path of the extreme value condition test for the innovations Z, captured from $\mathbf{X}$ on step 3.; sample paths of Hill (center) and moment (right) estimators, for Z.

confirm a Fréchet max-domain of attraction (Figure 5). Carrying out step 4., we capture the observations corresponding to U, for $\lambda = 0.05, 0.1, ..., 0.5$ and for the three scenarios ($\widehat{c} = 0.85, 0.8, 0.75$). On step 5., we apply the Kolmogorov-Smirnov test for the distribution $Beta(1/(0.5 * \widehat{c}) + 1, 1)$. In the case, $\widehat{c} = 0.85$, rejection is obtained for $\lambda \geq 0.20$ (see Figure 6) and the choice $\lambda = 0.15$ matches with the simulation study in Ferreira and Canto e Castro [7] (with 29 observations captured). Taking $\widehat{c} = 0.8$ (less catches) then $Beta(1/(0.5 * 0.8) + 1, 1)$ is rejected for $\lambda \geq 0.3$ with the best fit occurring for $\lambda = 0.2$, and with $\widehat{c} = 0.75$ (even less catches) only rejects $Beta(1/(0.5 * 0.75) + 1, 1)$ for $\lambda = 0.5$, both matching once again the simulation results in Ferreira and Canto e Castro [7].

Though pRARMAX fits to all data set, it is actually profiled for the modeling of large values, also known as *rare events*. As already mentioned, we are interested in estimating the probability that the maximum volatility exceeds a risky amount, for which we use the approximation in (2) considering both, $\theta = 1$ (the true value) and θ replaced by the pre-asymptotic version,

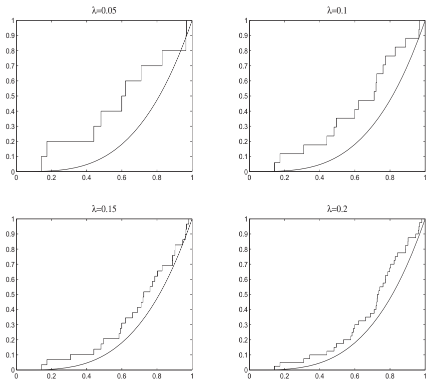

Figure 6: Empirical and theoretical d.f.'s of the random coefficients, U, captured from $\mathbf{X}$ through step 4. (with $\widehat{c} = 0.85$), for significance regions with, $\lambda = 0.05$, ..., $\lambda = 0.20$.

$\theta(u) = 1 - u^{\frac{1}{\gamma}(1-1/c)}$, derived from (3). Beside estimate 0.5, we also take $\gamma = 0.45$ and $\gamma = 0.4$, so we can see the effect of the very large peak. From Table 2, where we are considering the risky level 0.2, the probability estimates decrease significantly with the decrease of γ. Yet, in what respects c, a very small decrease takes place. Hence γ is a crucial parameter. Furthermore, we can also see that the higher the γ and the c the greater the differences in estimates.

Table 2: Estimates of the probability that the maximum volatility exceeds 0.2, based on (2), with $\theta = \theta(u) = 1 - u^{\frac{1}{\gamma}(1-1/c)}$ (first 3 lines) and with $\theta = 1$ (last line), considering $n = 10000$.

	$\gamma = 0.5$	$\gamma = 0.45$	$\gamma = 0.4$
$c = 0.85$	0.053222	0.010474	0.002881
$c = 0.8$	0.057372	0.015703	0.003028
$c = 0.75$	0.059219	0.01609	0.00308
$\theta = 1$	0.060295	0.016281	0.003101

References

[1] Alpuim, M.T., "An extremal markovian sequence," *J. Appl. Probab.*, V26, pp. 219-232, 1989.

[2] Bortot, P., Tawn, J., "Models for the extremes of Markov Chains," *Biometrika* V85, pp. 851-867, 1998.

[3] Daley D., Haslet J., "A thermal energy storage with controlled input," *Adv. in Appl. Probab.*, V14, N2, pp. 257-271, 1982.

[4] Davis, R., Resnick, S, "Basic properties and prediction of max-ARMA processes," *Adv. Appl. Probab.*, V21, pp. 781-803, 1989.

[5] Draisma, G., Drees, H., Ferreira, A., de Haan, L., "Bivariate tail estimation: dependence in asymptotic independence", *Bernoulli*, V10, pp. 251-280, 2004.

ISBN: 978-988-18210-8-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

[6] Ferreira, M., Canto e Castro, L., “Tail and dependence behaviour of levels that persist for a fixed period of time,” *Extremes,* V11, pp. 113-133, 2008.

[7] Ferreira M., Canto e Castro L, “Modeling rare events through a *p*RARMAX process,” *Notas e Comunicações do CEAUL*, 03/09.

[8] Ledford, A., Tawn, J.A., “Statistics for near independence in multivariate extreme values,” *Biometrika*, V83, pp. 169-187, 1996.

[9] Storey, J.D., “The Positive False Discovery Rate: A Bayesian Interpretation and the *q*-Value,” *Ann. Statist.*, V31, pp. 2013-2035, 2003.

[10] Zhang, Z., Smith, R.L., “Modelling Financial Time Series Data as Moving Maxima Processes”, *Tech. Rep. Dept. Stat. (Univ. North Carolina, Chapel Hill, NC, 2001)*; http://www.stat.unc.edu/faculty/rs/ papers/RLS_Papers.html.