International Journal of Theoretical and Applied Finance
Vol. 13, No. 6 (2010) 1–20

DOI: 10.1142/S0219024910006066

ON THE IMPACT OF HIDDEN TRENDS FOR A COMPOUND POISSON MODEL WITH PARETO-TYPE CLAIMS

PETER GRANDITS* and REINHOLD KAINHOFER†

Financial and Actuarial Mathematics
Institute for Mathematical Methods in Economics
Vienna University of Technology
Wiedner Hauptstraße 8-10/105-1
A-1040 Vienna, Austria
*pgrand@fam.tuwien.ac.at
†reinhold@kainhofer.com

GRIGORY TEMNOV

Edgeworth Centre for Financial Mathematics
School of Mathematical Sciences
University College Cork, Ireland
g.temnov@ucc.ie

Received 9 May 2009
Accepted 13 April 2010

We consider a compound-Poisson model with Pareto-type claims. In contrast to the classical case, where the claims are assumed to be iid., we assume that scaling and location parameters of the Pareto distribution follow a certain trend. We investigate the impact of this trend on parameter estimation and on the VaR (Value-at-Risk), if one mis-specifies (or even neglects) this trend. In the first part we show a consistency result for the mis-specified model, in the second part the deviations of the true parameters from the ones obtained by applying an iid. procedure is measured. Finally we study the impact of the mis-specification on a typical risk measure like the VaR.

Keywords: ML estimation; consistency; pareto distribution; mis-specified models; trends.

1. Introduction

In this article we look at a series of insurance or credit risk claims with recorded claim times T_i and sizes Y_i, modelled as a compound Poisson process. In the present paper, the claim times are assumed to follow an inhomogeneous process. The typical approach in risk theory is to assume the claim amounts Y_i to be independent and identically distributed according to a claim size distribution F, e.g. the Pareto or generalized Pareto distribution. In reality, however, a typical claim may be time-dependent and the claim size distribution will change in time, so that the variables Y_i are subject to some certain trend. Detrending to iid. random variables is in general not (or only approximately) possible in practice, because the values of the trend parameters are only estimated and not exactly known. So, even after the transformation the Y_i are not exactly iid., but subject to a (smaller) trend, parameterized by r in this paper. The aim of this article is to investigate the error in parameter estimation and the estimation of the VaR that is made by applying methods derived for iid. claim amounts to data with trends of the general form $Y_i = X_i q(r, T_i) + d(r, T_i)$. Here r is some (small) real parameter. In our setting, we consider the process on a limited time interval (the length of which is normalized to 1, so that $T_i \in [0, 1]$). In order to consider the limit to infinitely many claims in a proper way, i.e. with respect to single chosen measure, we make a certain artificial assumption. Specifically, the claim times are assumed to follow a Poisson process with periodic jump intensity, so that the jump counting process is a cadlag process. Our results are obtained on the time interval $[0, 1]$ without loss of generality, as a simple time-scaling will provide results on other intervals.

We first show a consistency result for the mis-specified model, i.e. we show that if we calculate the maximizer of the likelihood function for the iid. model, but using data with trends, we still observe a.s. convergence of the maximizer to a certain parameter value, say $\theta(r)$. In the next step we provide a Taylor expansion of the maximum likelihood estimators $\theta(r)$ with respect to r around $r = 0$. In a final section we investigate the impact of the mis-specification on the VaR. For the claim size distribution we will assume a Pareto distribution. For other claim size distributions a similar path might be possible, but our calculations depend explicitly on the form of the Pareto density.

In [3], Grandits and Temnov investigated a similar problem, only that the claim times were assumed to be equally spaced in time, while here we assume them to follow an inhomogeneous Poisson process. Moreover, they considered pure inflation only, i.e. $Y_i e^{-rT_i}$ being iid. The special case of an homogeneous Poisson process has the same limits as the equally-spaced case, and the results of [3] can be reproduced here, too. Let us finally note that one could, of course, estimate also the parameter r with the given data set. But in practice this is often not done, since it involves an increase in the dimensionality of the problem.

The impact of hidden trends on the estimates of severity parameters has been discussed in the literature, e.g. in application to the risk analysis and insurance problems. To mention one of the works, [5] deals with the analysis of different scenarios for the parameters of the generalized Pareto distribution changing in time. In [5], the authors detect some hidden trends in the insurance data using the graphical diagnostics of the data. Then different scenarios of the location, scale and shape parameters changing in time, each following a linear or an exponential law, are considered. Various combinations of the trends are then analyzed in order to find out which of such scenarios correspond to a better fit of the data.

For the claim size distribution F of the discounted claims we will assume the Pareto distribution with shape parameter α^* and scale parameter σ^*. In particular, the CDF and the PDF are explicitly given by

$$F_{\alpha^*,\sigma^*}(x) = 1 - \left(1 + \frac{x}{\sigma^*}\right)^{-\alpha^*} \qquad f_{\alpha^*,\sigma^*}(x) = \frac{\alpha^*}{\sigma^*}\left(1 + \frac{x}{\sigma^*}\right)^{-\alpha^*-1}.$$

The observed claims Y_i will be dependent on the claim time T_i and thus not be iid. on their own. In particular, we assume an affine time-dependence of the form

$$Y_i = X_i q(r, T_i) + d(r, T_i)$$

for a small positive parameter $r \in [0, r_0]$ with the additional conditions that q and d are smooth functions, which fulfill the limiting requirements in r

$$q(r,t) \xrightarrow{r\to 0} 1 \quad d(r,t) \xrightarrow{r\to 0} 0$$

uniformly on $[0, 1]$, as well as uniform boundedness conditions in t

$$0 < q(r,t) \leq Cq(r,1)$$
$$0 < d(r,t) \leq Cd(r,1)$$

for some positive constant C. We restrict ourselves to positive additive trends $d(r,t)$, in order to avoid "negative claims", which would cause additional technical difficulties. We postpone the (of course important) case of negative values of d to future research. The random variables X_i are assumed to be iid. Pareto-distributed, the claim times T_i are modelled as the jump times of an inhomogeneous Poisson process.

This form of time dependence includes constant inflation rate r of the form $Y_i = X_i e^{rT_i}$ for the claims. For time-dependent, but still deterministic inflation rates $\delta(r,t)$, parameterized by a base inflation value r and the time t, the particular forms of the Y_i are obvious.

2. Preliminaries on Poisson Processes

As we model the claim arrival process as an inhomogeneous Poisson process, we first need to recall some important properties of inhomogeneous Poisson processes, which will greatly simplify our investigations as compared to the equally-spaced case.

Let $(N_t)_{t\geq 0}$ be a (possibly inhomogeneous) Poisson process on $[0, 1]$ with jump intensity $\lambda(t)$ and mean value function: $\mu(t) = \int_0^t \lambda(\tau)d\tau$. Assume that the function λ is normalized to 1 so that $\int_0^1 \lambda(\tau)d\tau = 1$ and $\mathbf{E}[N_1] = \mu(1) = 1$.

On the base of the process N_t, we now construct an inhomogeneous process $\widetilde{N}_t$ on $[0, M]$, where $M \in \mathbb{N}$, in the following way: Let the jump intensity $\widetilde{\lambda}(t)$ of $\widetilde{N}_t$ be periodic with period 1 and coinciding with the intensity $\lambda(t)$ on each of its periods $[0, 1], (1, 2] \ldots (M-1, M]$. We denote $\widetilde{\mu}(t) = \int_0^t \widetilde{\lambda}(\tau)d\tau$ and assume that $\int_0^M \widetilde{\lambda}(\tau)d\tau = M$, so that $\mathbf{E}[\widetilde{N}_M] = \widetilde{\mu}(M) = M$.

The intention behind the parameterization set above is to provide a simple way to take the limit to infinitely many claims ($M \to \infty$) while the limiting model is defined via a single random measure and, as we shall see below, is expressed through the intensity $\lambda(t)$.

Correspondingly, we extend the functions q and d defined above by defining the periodic functions $\widetilde{q}$ and $\widetilde{d}$ on the time interval $[0, M]$, coinciding correspondingly with q and d on each of the periods. The functions $\widetilde{\lambda}$, $\widetilde{q}$ and $\widetilde{d}$ are assumed to be piecewise smooth functions.

Let T_i denote jump times of the jump count process N_t defined above (for the homogeneous PP the increments are independent and identically distributed with exponential distribution; for the inhomogeneous PP this is no longer the case). The jump sizes $X_1, X_2, \ldots, X_{N_t}$ are iid. F-distributed, independent of the jump times T_i. The notation $\widetilde{T}_i$ will correspond to the process $\widetilde{N}_t$.

From the theory of the inhomogeneous Poisson process (see e.g. [4]) we know the following results. In particular, a useful property of Poisson processes is that, conditional on the number $N_t = n$ of claims, the claim times have the same distribution as the order statistics of an iid. sample.

Proposition 2.1. *Let the Poisson process N_t have a continuous, a.e. positive intensity function $\lambda(t)$ and arrival times $0 < T_1 < T_2 < \cdots$ a.s., and let $\mathbf{\Upsilon} = (\Upsilon_1, \ldots, \Upsilon_n)$ be a sample from r.v. Υ with PDF $g_t(x) = \lambda(x)/\mu(t)$ (so that $g_1(x) = \lambda(x)$) and G being its CDF, and $\overline{\mathbf{\Upsilon}}$ be the corresponding order statistics vector $\overline{\mathbf{\Upsilon}} = (\Upsilon_{(1)}, \ldots, \Upsilon_{(n)})$ with the joint PDF $g_{\overline{\mathbf{\Upsilon}}}$. Then, if f is the joint PDF of the random vector $\mathbf{T} = (T_1, \ldots, T_n)$ then for the corresponding conditional PDF $f_t(x_1, \ldots, x_n) \equiv f(x_1, \ldots, x_n \,|\, N_t = n)$ we have the following identity*

$$f_t(x_1, \ldots, x_n) = g_{\overline{\mathbf{\Upsilon}}}(x_1, \ldots, x_n) = g_{G^{-1}(\overline{\mathbf{U}})}(x_1, \ldots, x_n),$$

where $\overline{\mathbf{U}} = (U_{(1)}, \ldots, U_{(n)})$ is the order statistics vector of the sample $U_i \sim U(0,1)$ of iid. uniformly distributed random variables on $[0,1]$, and $g_{G^{-1}(\overline{\mathbf{U}})}$ is the PDF of the corresponding random vector $(G^{-1}(U_{(1)}), \ldots, G^{-1}(U_{(n)}))$.

Clearly, for the process $\widetilde{N}_t$ the same property for $(\widetilde{T}_i, X_i)$ holds with df $\widetilde{G}$ whose pdf is $\widetilde{g}(x) = \frac{\widetilde{\lambda}(x)}{\widetilde{\mu}(t)}$, and considering it at times $t = M \in \mathbb{N}$ we have $\widetilde{g}(x) = \frac{\widetilde{\lambda}(x)}{\widetilde{\mu}(M)} = \frac{\widetilde{\lambda}(x)}{M}$.

Proposition 2.2. *Let (X_i) be an iid. sequence, independent of the sequence $(\widetilde{T}_i)$ of the arrival times of an inhomogeneous Poisson process $\widetilde{N}_t$ with intensity $\widetilde{\lambda}(t)$. Then for any measurable function $h : \mathbb{R}^2 \to \mathbb{R}$ the identity*

$$\sum_{i=1}^{N_t} h(\widetilde{T}_i, X_i) \overset{d}{=} \sum_{i=1}^{N_t} h(\widetilde{G}^{-1}(U_i), X_i)$$

holds, where (U_i) is an iid. $U(0,1)$ sequence, independent of (X_i) and $(\widetilde{T}_i)$, and where $\widetilde{G}^{-1}$ is understood in the sense of the generalized inversion.

3. The Maximum Likelihood Equation for Claims Without and With Trend

In the absence of inflation and other time-dependent effects for the claim sizes X_i, the log-likelihood function for the Pareto distribution is

$$l(\alpha, \sigma) = n \log \alpha - n \log \sigma - (\alpha + 1) \sum_{i=1}^{n} \log \left(1 + \frac{X_i}{\sigma}\right) \tag{3.1}$$

with the maximum likelihood estimators α^* and σ^* fulfilling the equations

$$\left.\frac{\partial l(\alpha, \sigma)}{\partial \alpha}\right|_{\alpha^*, \sigma^*} = \frac{1}{\alpha^*} - \frac{1}{n} \sum_{i=1}^{n} \log \left(1 + \frac{X_i}{\sigma^*}\right) = 0$$

$$\left.\frac{\partial l(\alpha, \sigma)}{\partial \sigma}\right|_{\alpha^*, \sigma^*} = -\frac{(\alpha^* + 1)}{\sigma^*} \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{(\sigma^*)^2 + X_i \sigma^*} = 0$$

As mentioned in the introduction, the r.v.'s Y_i are not iid. due to a (remaining) trend. However, since we will always consider in the sequel expressions (conditioned on $N_t = n$), which are symmetric with respect to the permutation of the indices $i = 1, \ldots, n$ (due to the iid. property of the sequence $\{X_i\}$), we assume in the following that the sequence $\{\widetilde{T}_i\}_{i=1}^n$ is an iid. sample with the CDF $\widetilde{G}$. Inserting the values $Y_i = X_i q(r, \widetilde{T}_i) + d(r, \widetilde{T}_i)$ instead of X_i into the log-likelihood function (3.1), the maximizer $(\alpha(r), \sigma(r))$ will be different from (α^*, σ^*) without trend. Then we have the equations

$$\left.\frac{\partial l(\alpha, \sigma)}{\partial \alpha}\right|_{\alpha(r), \sigma(r)} = \frac{1}{\alpha(r)} - \frac{1}{n} \sum_{i=1}^{n} \log \left(1 + \frac{Y_i}{\sigma(r)}\right) = 0,$$

$$\left.\frac{\partial l(\alpha, \sigma)}{\partial \sigma}\right|_{\alpha(r), \sigma(r)} = -\frac{(\alpha(r) + 1)}{\sigma(r)} \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i}{(\sigma(r))^2 + Y_i \sigma(r)} = 0.$$

The summands in both equations are iid. random variables for some distribution that can — theoretically — be calculated from the distributions of X_i and $\widetilde{T}_i$. To be able to apply several well known theorems from probability theory, moments of sufficient order of the summands need to exist.

- For $Z_j = \log(1 + \frac{X_j \widetilde{q}(r, \widetilde{T}_j) + \widetilde{d}(r, \widetilde{T}_j)}{\sigma(r)})$ with $Z_1, \ldots, Z_n$ iid., all moments of these random variables are finite due to slower asymptotic growth of the logarithm than any polynomial and if we assume a lower bound on σ, i.e. $0 < \eta \leq \sigma$, then:

$$\mathbf{E}[Z_j^n] = \int_0^\infty \int_0^\infty \log^n \left(1 + \frac{x \widetilde{q}(r, t) + \widetilde{d}(r, t)}{\sigma}\right) dF(x) dG(t)$$

$$\leq \int_0^\infty \log^n \left(1 + C \frac{x \widetilde{q}(r, M) + \widetilde{d}(r, M)}{\sigma}\right) dF(x)$$

$$
\begin{aligned}
&\leq \int_0^\infty \log^n \left(1 + \frac{xC}{\eta}\widetilde{q}(r,M) + \frac{\widetilde{d}(r,M)C}{\eta}\right) dF(x) \\
&\leq D_1(\epsilon) + D_2(\epsilon)C^\epsilon \int_{M(\epsilon)}^\infty \left(\frac{x}{\eta}\widetilde{q}(r,M) + \frac{\widetilde{d}(r,M)}{\eta}\right)^\epsilon \\
&\qquad \times dF(x) < \infty \quad \text{for any } \epsilon < \alpha^*.
\end{aligned} \tag{$*$}
$$

- For $\widehat{Z}_j = \frac{X_i\widetilde{q}(r,T_i)+\widetilde{d}(r,T_i)}{\sigma(r)+X_i\widetilde{q}(r,T_i)+\widetilde{d}(r,T_i)}$ with $\widehat{Z}_1, \ldots, \widehat{Z}_n$ iid., we have

$$
0 \leq \widehat{Z}_j = \frac{X_i\widetilde{q}(r,T_i) + \widetilde{d}(r,T_i)}{\sigma(r) + X_i\widetilde{q}(r,T_i) + \widetilde{d}(r,T_i)} = \left(\frac{\sigma(r)}{X_i\widetilde{q}(r,T_i) + \widetilde{d}(r,T_i)} + 1\right)^{-1} \leq 1
$$

  and thus also $0 \leq \mathbf{E}[\widehat{Z}_j^n] \leq 1$ for any $n \geq 0$.

Notice that in both cases the bounds for the moments depend only on η, but not on the particular value of σ. Thus all statements depending only on upper bounds for the moments of Z_j and $\widehat{Z}_j$ hold uniformly for $\sigma \in [\eta, \infty)$.

4. Strong Law of Large Numbers

Proposition 4.1. *Let $(\widetilde{T}_i, X_i)$ be a (homogeneous or inhomogeneous) compound Poisson process with jumps X_i at times $\widetilde{T}_i$. We set $\mathbf{E}[\widetilde{N}_t] = \widetilde{\mu}(t) = \int_0^t \widetilde{\lambda}(\tau)d\tau$ with $\int_0^M \widetilde{\lambda}(\tau)d\tau = M$ for the mean value function of the jump count process $\widetilde{N}_t$, where $\widetilde{\lambda}$ is periodic with period 1 (so that $\widetilde{\lambda}$, on each of its periods, coincides with a continuous function λ).*

Let the jump size distribution be $X_i \sim F = Pareto(\alpha^*, \sigma^*)$ with probability density function $f_X(x) = \frac{\alpha^*}{\sigma^*}(1 + \frac{x}{\sigma^*})^{-\alpha^*-1}$ for constants α^*, σ^*. Let furthermore r be constant.

Then there exists a function $K : \mathbb{N} \mapsto \mathbb{R}^+$ such that uniformly for $\sigma \in [\eta, \infty)$

$$
\begin{aligned}
&\mathbf{P}\left[\left|\frac{1}{\widetilde{N}_M}\sum_{1=1}^{\widetilde{N}_M} \log\left(1 + \frac{X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)}{\sigma}\right)\right.\right. \\
&\qquad \left.\left. - \int_0^M \int_0^\infty \log\left(1 + \frac{x\widetilde{q}(r,z) + \widetilde{d}(r,z)}{\sigma}\right) dF(x)d\widetilde{G}(z)\right| > \varepsilon\right] \\
&\quad \leq K(M) \xrightarrow{M\to\infty} 0
\end{aligned} \tag{4.1a}
$$

$$
\mathbf{P}\left[\left|\frac{1}{\widetilde{N}_M}\sum_{i=1}^{\widetilde{N}_M} \frac{X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)}{\sigma + X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)}\right.\right.
$$

$$- \int_0^M \int_0^\infty \frac{x\widetilde{q}(r,z) + \widetilde{d}(r,z)}{\sigma + x\widetilde{q}(r,z) + \widetilde{d}(r,z)} dF(x) d\widetilde{G}(z) \Bigg| > \varepsilon \Bigg]$$

$$\leq K(M) \xrightarrow{M \to \infty} 0 \tag{4.1b}$$

and $\sum_{M=1}^{\infty} K(M) < \infty$.

Proof. For brevity, define

$$Z_j = \log\left(1 + \frac{X_j \widetilde{q}(r,\widetilde{T}_j) + \widetilde{d}(r,\widetilde{T}_j)}{\sigma}\right) \quad \text{and} \quad \widehat{Z}_j = \frac{X_j \widetilde{q}(r,\widetilde{T}_j) + \widetilde{d}(r,\widetilde{T}_j)}{\sigma + X_j \widetilde{q}(r,\widetilde{T}_j) + \widetilde{d}(r,\widetilde{T}_j)} \tag{4.2}$$

with

$$\mathbf{E}[Z_j] = \int_0^M \int_0^\infty \log\left(1 + \frac{x\widetilde{q}(r,z) + \widetilde{d}(r,z)}{\sigma}\right) dF(x) d\widetilde{G}(z) \tag{4.3}$$

$$\mathbf{E}[\widehat{Z}_j] = \int_0^M \int_0^\infty \frac{x\widetilde{q}(r,z) + \widetilde{d}(r,z)}{\sigma + x\widetilde{q}(r,z) + \widetilde{d}(r,z)} dF(x) d\widetilde{G}(z). \tag{4.4}$$

Conditioning on the total number of claims $\widetilde{N}_M = n$, which is Poisson-distributed with mean value M, and using the conditional independence property of the jump times of the Poisson process, we can rewrite the probabilities (4.1a) and (4.1b) of the theorem as

$$\mathbf{P}\left[\left|\frac{1}{\widetilde{N}_M} \sum_{j=1}^{\widetilde{N}_M} Z_j - \mathbf{E}[Z]\right| > \varepsilon\right] = \mathbf{E}\left[\mathbf{P}\left(\left|\frac{1}{\widetilde{N}_M} \sum_{j=1}^{\widetilde{N}_M} Z_j - \mathbf{E}[Z]\right| > \varepsilon \;\Bigg|\; \widetilde{N}_M = n\right)\right] \tag{4.5}$$

with the similar form for $\widehat{Z}_j$.

Due to the properties of the inhomogeneous Poisson process, the Z_j and $\widehat{Z}_j$ are independent and identically distributed random variables, conditional on $N_M = n$ fixed, and all their moments are finite as already shown. Thus we can apply Lemma 2.1 of [3] for any fixed $n = 1, 2, \ldots$ to obtain the bound

$$\mathbf{P}\left(\left|\frac{1}{\widetilde{N}_M} \sum_{j=1}^{\widetilde{N}_M} Z_j - \mathbf{E}[Z]\right| > \varepsilon \;\Bigg|\; \widetilde{N}_M = n\right) \leq \frac{C}{\epsilon^4 n^2} \tag{4.6}$$

for the inner term in (4.5) and the same bound for $\widehat{Z}$. The constant C depends only on the moments of Z_j and $\widehat{Z}_j$ up to order 4 and for a fixed $\eta > 0$ can be chosen uniformly $C = C(\eta)$ for all $\sigma \in [\eta, \infty)$. The whole probability in (4.6) in both cases (for Z_j and $\widehat{Z}_j$) can then be bounded by

$$\mathbf{P}(N_M = 0)\underbrace{\mathbf{P}(|\mathbf{E}[Z]| > \varepsilon)}_{\leq 1} + \sum_{n=1}^{\infty} \mathbf{P}(\widetilde{N}_M = n)\frac{C}{\varepsilon^4 n^2}$$

$$\leq e^{-M}\left\{1 + \frac{C}{\varepsilon^4}\sum_{n=1}^{\infty}\frac{M^n}{n^2 n!}\right\} =: K(M), \tag{4.7}$$

where we used that $\mathbf{P}(|\frac{1}{\widetilde{N}_M}\sum_{j=1}^{\widetilde{N}_M} Z_j - \mathbf{E}[Z]| > \varepsilon \,|\, \widetilde{N}_M = 0) = \mathbf{P}(|\mathbf{E}[Z]| > \varepsilon)$.

Clearly, the series $\sum_{n=1}^{\infty}\frac{M^n}{n^2 n!}$ is convergent.

It remains to show the assertion that the $K(M)$ form an absolutely convergent series. The first term e^{-M} leads to the absolutely convergent series $\sum_{n=1}^{\infty} e^{-n} = \frac{1}{1-e}$. The second term $\sum_{n=1}^{\infty}\frac{M^n}{n^2 n!}$ leads to a double series and is a bit trickier:

$$\sum_{M=0}^{\infty}\sum_{n=1}^{\infty} e^{-M}\frac{M^n}{n^2 n!} = \sum_{n=1}^{\infty}\frac{1}{n^2 n!}\sum_{M=0}^{\infty} M^n e^{-M}, \tag{4.8}$$

where the exchange of the two sums will later be justified by obtaining an absolutely convergent series dominating the whole expression. As all terms are non-negative, any reordering is then allowed and leads to an absolutely convergent series, too.

To bound the expression (4.8) one has to notice that the inner sum $\sum_{M=0}^{\infty} M^n e^{-M}$ is just the discretized version of the integral representation of the

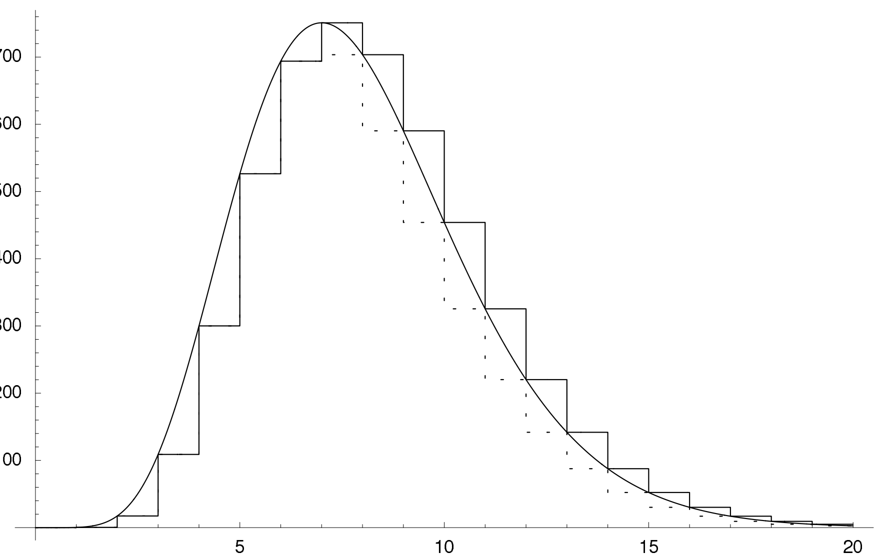

Fig. 1. Comparing the Gamma function to the discretized sum.

Gamma function

$$\Gamma(n+1) = \int_0^\infty x^n e^{-x} dx$$

with $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$. Thus we will bound the sum by the gamma function and some additional correction terms. It is easily seen that the maximum of $x^n e^{-x}$ lies at $x = n \in \mathbb{N}$, so that we have the following bounds

$$\begin{aligned} M^n e^{-M} &= \int_M^{M+1} M^n e^{-M} dx \le \int_M^{M+1} x^n e^{-x} dx \quad \text{for } M \le n-1, \quad M \in \mathbb{R}, \\ M^n e^{-M} &= \int_{M-1}^{M} M^n e^{-M} dx \le \int_{M-1}^{M} x^n e^{-x} dx \quad \text{for } M \ge n+1, \quad M \in \mathbb{R}. \end{aligned}$$

Plugging these estimates into the sum, we arrive at

$$\begin{aligned} \sum_{M=0}^{\infty} M^n e^{-M} &\le \sum_{M=0}^{n-1} \int_M^{M+1} x^n e^{-x} dx + n^n e^{-n} + \sum_{M=n+1}^{\infty} \int_{M-1}^{M} x^n e^{-x} dx \\ &= \underbrace{\int_0^\infty x^n e^{-x} dx}_{=\Gamma(n+1)=n!} + n^n e^{-n} = n! + n^n e^{-n}. \end{aligned}$$

According to Sterling's formula, $n! = (\frac{n}{e})^n \sqrt{2\pi n} \exp(\theta(n)/12n)$ with $0 < \theta(n) < 1$. Altogether, we get

$$\begin{aligned} \sum_{M=0}^{\infty} e^{-M} \sum_{n=1}^{\infty} \frac{M^n}{n^2 n!} &\le \sum_{n=1}^{\infty} \frac{n! + n^n e^{-n}}{n^2 n!} \\ &= \underbrace{\sum_{n=1}^{\infty} \frac{1}{n^2}}_{=\zeta(2)} + \sum_{n=1}^{\infty} \frac{n^{n-2} e^{-n}}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n} \exp(\theta(n)/12n)} \\ &= \zeta(2) + \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{\infty} \frac{1}{n^{2+\frac{1}{2}}} \underbrace{e^{-\theta(n)/12n}}_{\le 1} \le \zeta(2) + \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{\infty} \frac{1}{n^{2.5}} \\ &= \zeta(2) + \frac{1}{\sqrt{2\pi}} \zeta(2.5) < \infty \end{aligned}$$

where $\zeta(\cdot)$ denotes the Riemann zeta function and the reordering of the terms of the series is again justified by the non-negativity of the terms and the resulting absolutely convergent series.

Since both terms of (4.7) lead to absolutely convergent series, their sum is also absolutely convergent and the terms can be reordered in any way. This completes the proof of the proposition for both equations. □

The previous proposition shows that the prerequisites for the Borel-Cantelli lemma are satisfied, so it immediately follows that the mean of the discretized observations converge almost surely to the expectation as we let $M \in \mathbb{N}$ tend towards infinity.

Theorem 4.1. *Let $\widetilde{N}_t$ be a Poisson process with mean value function $\widetilde{\mu}(t)$ such that $\mathbf{E}[\widetilde{N}_M] = M \in \mathbb{N}$. Then we have the limits*

$$\frac{1}{\widetilde{N}_M}\sum_{i=1}^{\widetilde{N}_M}\log\left(1+\frac{X_i\widetilde{q}(r,\widetilde{T}_i)+\widetilde{d}(r,\widetilde{T}_i)}{\sigma}\right) \xrightarrow[M\to\infty]{a.s.} \int_0^1\int_0^\infty \log\left(1+\frac{xq(r,z)+d(r,z)}{\sigma}\right)dF(x)dG(z)$$

$$\frac{1}{\widetilde{N}_M}\sum_{i=1}^{\widetilde{N}_M}\frac{X_i\widetilde{q}(r,\widetilde{T}_i)+\widetilde{d}(r,\widetilde{T}_i)}{\sigma+X_i\widetilde{q}(r,\widetilde{T}_i)+\widetilde{d}(r,\widetilde{T}_i)} \xrightarrow[M\to\infty]{a.s.} \int_0^1\int_0^\infty \frac{xq(r,z)+d(r,z)}{\sigma+xq(r,z)+d(r,z)}dF(x)dG(z)$$

and the convergence is uniform for $\sigma \in [\eta,\infty[$, $\eta > 0$.

Remark 4.1. The above limit $M \to \infty$ is proved only for $M \in \mathbb{N}$ using the Borel-Cantelli lemma, which is sufficient for our case. To show almost sure convergence for $\mathbb{R} \ni M \to \infty$, one might apply continuity and monotonicity arguments to extend the above result to the general case $M \in \mathbb{R}$.

5. Convergence of The Maximum Likelihood Estimator

We will first show in this section that the maximum likelihood function for claims with trend has a maximizer in a large enough compact parameter set, if the intensity M is large enough. Since we have proved in the last section that the sum in the loglikelihood function can be approximated by an expression involving a double integral, it will turn out that it is crucial for our proof to get an idea of the asymptotic behavior of this double integral as $\sigma \to \infty$. This asymptotic behavior is provided in the following lemma.

Lemma 5.1. *Let $a = \max\{n \in \mathbb{N} \,|\, n < \alpha^* + 1\}(\to a \geq 1)$.*

Then we have for $\sigma \to \infty$ in the case $a = 1$

$$\int_0^1\int_0^\infty \ln\left(1+\frac{xq(r,z)+d(r,z)}{\sigma}\right)dF(x)dG(z) \sim \begin{cases}\nu_1^{(1)}\sigma^{-\alpha^*} & \text{for } \alpha^* < 1\\ \nu_2^{(1)}\dfrac{\ln\sigma}{\sigma} & \text{for } \alpha^* = 1,\end{cases}$$

for some positive constants $\nu_1^{(1)}, \nu_2^{(1)}$. In the case $a = 2$ we have

$$\int_0^1\int_0^\infty \ln\left(1+\frac{xq(r,z)+d(r,z)}{\sigma}\right)dF(x)dG(z) \sim \frac{\nu_1^{(2)}}{\sigma} - \begin{cases}\dfrac{\nu_{21}^{(2)}}{\sigma^{\alpha^*}} & \text{for } \alpha^* \in]1,2[\\ \nu_{22}^{(2)}\dfrac{\ln\sigma}{\sigma^2} & \text{for } \alpha^* = 2,\end{cases}$$

for some positive constants $\nu_1^{(2)}, \nu_{21}^{(2)}, \nu_{22}^{(2)}$. *Finally for* $a > 2$ *we have*

$$\int_0^1 \int_0^\infty \ln\left(1 + \frac{xq(r,z) + d(r,z)}{\sigma}\right) dF(x)dG(z) \sim \frac{1}{\sigma}\left(\rho_{01} + \rho_{10}\frac{\sigma^*}{\alpha^* - 1}\right)$$
$$+ \frac{1}{\sigma^2}\left(-\frac{\rho_{02}}{2} - \rho_{11}\frac{\sigma^*}{\alpha^* - 1} - \rho_{20}\frac{(\sigma^*)^2}{(\alpha^* - 1)(\alpha^* - 2)}\right),$$

with $\rho_{ij} = \int_0^\infty (q(r,z))^i (d(r,z))^j dG(z)$.

Proof. The proof works analogously (repeated integration by parts for the inner x-integral) to the proof of Lemma 3.1 in [3] □

We define now the normalized log-likelihood function as

$$l(\alpha,\sigma) := \ln\alpha - \ln\sigma - (\alpha+1)\frac{1}{\widetilde{N}_M}\sum_{i=1}^{\widetilde{N}_M} \ln\left(1 + \frac{X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)}{\sigma}\right) \tag{5.1}$$

and show that it attains its maximum over $(\mathbb{R}_+)^2$ in a large compact set A, defined by

$$A = \{(\alpha,\sigma) | \alpha \geq \eta^2, \sigma \geq \eta, \sqrt{\alpha^2 + \sigma^2} \leq W\} \quad \text{for } \eta > 0 \text{ small and } W > 0 \text{ large.}$$

More precisely, we have

Proposition 5.1. *For* W *large enough and* η *and* r *small enough, the following is valid*:

$$\mathbf{P}[\exists M(\omega,\eta,W) \text{ s.t. } \forall M' \geq M, \quad \text{we have } \max_A l > \sup_{(\mathbb{R}_+)^2 \setminus A} l] = 1.$$

Proof. Like in [3] we cover $\{(\mathbb{R}_+)^2 \setminus A\}$ by five different sets, as shown in Fig. 2, and sequentially show, that l does not attain its maximum inside the sets $S_1 - S_5$ a.s. for M' large enough. The sets $S_1 - S_5$ are defined as follows.

$$S_1 := \{(\alpha,\sigma) | \alpha \in]0,W], \sigma \in]0,\eta]\}$$
$$S_2 := \{(\alpha,\sigma) | \alpha \in]0,\eta^2], \sigma \in [\eta,W]\}$$
$$S_3 := \left\{(\alpha,\sigma) | \alpha \geq \sqrt{W^2 - \sigma^2}, \sigma \in \left]0, G^* := \frac{e^{D^*}}{2}\right]\right\}$$
$$S_4 := \{(\alpha,\sigma) | \alpha \geq \sqrt{W^2 - \sigma^2}, \sigma \in [G^*, \sqrt{W}]\}$$
$$S_5 := \left\{(\alpha,\sigma), \alpha \begin{cases} > 0, \text{ if } \sigma \geq W \\ \geq \sqrt{W^2 - \sigma^2}, \text{ if } W > \sigma \geq \sqrt{W} \end{cases}, \sigma \in [\sqrt{W}, \infty[\right\}$$

The only difference to the definition in [3] is that now the constants G^*, D^* may also depend on the function $q(r,t)$ and not only on the r.v. X_i.

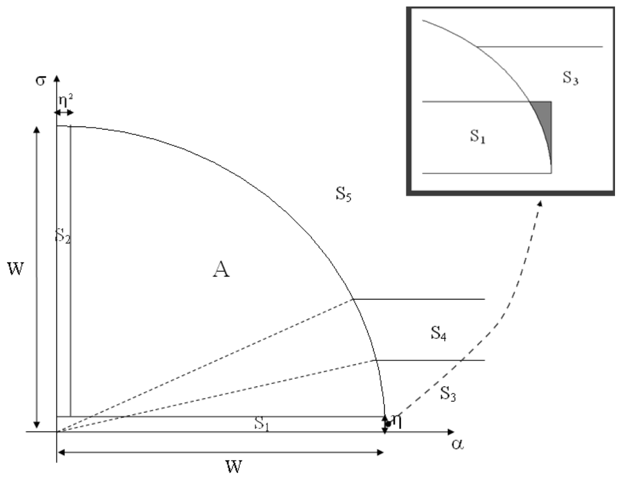

Fig. 2. Covering of $(\mathbb{R}_+)^2$ by sets for the proof of Proposition 5.1.

We condition on $\{\widetilde{N}_M = n\}$ and get in complete analogy to the proof of Proposition 3.3 in [3] estimates of the form

$$\mathbf{P}(l(\alpha,\sigma) \leq f_1(\eta) \text{ on } S_i) \geq 1 - \frac{C}{n^2}, \quad \text{or} \quad \mathbf{P}(l(\alpha,\sigma) \leq f_2(W) \text{ on } S_j) \geq 1 - \frac{C}{n^2}$$

where $i = 1, 2$, $j = 3, 4$, $f_1(\eta)$ and $f_2(W)$ tend to $-\infty$ for $\eta \to 0$, respectively for $W \to \infty$. Arguing as in the proof of Proposition 4.1, we get via the Borel-Cantelli lemma, for W large enough (resp. σ small enough)

$$\mathbf{P}\left[\exists M(\omega, W(\text{resp. } \sigma)\ s.t.\ \forall M' \geq M : l(1,1) > \sup_{S_i} l\right] = 1, \tag{5.2}$$

for $i = 1, 2, 3, 4$.

In order to get a similar result for the set S_5 we will use the following Lemma (for the proof we refer to the Appendix).

Lemma 5.2. *Let*

$$L(\alpha,\sigma) := -(\alpha+1)\int_0^1\int_0^\infty \ln\left(1 + \frac{xq(r,z) + d(r,z)}{\sigma}\right) dF(x)dG(z) + \ln\frac{\alpha}{\sigma}. \tag{5.3}$$

Then, for W large enough and r small enough, there exists a point $P \in A$, s.t. $L(P) > \sup_{S_5} L$.

Assuming Lemma 5.2, we can find for all W large enough and r small enough a $\delta(W,r)$, s.t. there exists a point $P \in A$ with $L(P) - \delta(W,r) \geq \sup_{S_5} L(\alpha,\sigma)$. By Proposition 4.1 we have

$$\mathbf{P}(|l - L| > \varepsilon \text{ on } \{\sigma \geq \eta\}) \leq K(M,\epsilon),$$

where $\sum_{M=1}^{\infty} K(M,\epsilon) < \infty$ holds for all $\epsilon > 0$. Choosing $\varepsilon = \frac{\delta}{3}$, one gets

$$\mathbf{P}\left(\sup_A l > \sup_{S_5} l\right) \geq 1 - K(M,\delta/3),$$

which clearly proves the analogue of equation (5.2) for the set S_5. This concludes the proof of Proposition 5.1 □

We can now formulate our main theorem, namely that the maximum-likelihood estimators in the misspecified model converge almost surely to certain parameter values $\alpha(r), \sigma(r)$.

Theorem 5.1. *For $r > 0$ small enough, the ML estimator $\widehat{\theta}_M = (\widehat{\alpha}_M, \widehat{\sigma}_M)$, which exists ultimately a.s. by Proposition* 5.1, *fulfills*

$$\widehat{\theta}_M \stackrel{a.s.}{\to} \theta(r) = (\alpha(r), \sigma(r)), \tag{5.4}$$

for certain parameters $\alpha(r), \sigma(r)$.

After our preparations the proof is exactly the same as the proof of Theorem 3.11 in [3].

6. Expansion of The ML Estimator for Small r

Clearly, by plugging the observed values $X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)$ into the maximum likelihood equation for the X_i, the solutions $(\alpha(r),\sigma(r))$ will deviate from the solutions (α^*,σ^*) without trend. In this section we will investigate the order of these deviations for small r by Taylor-expanding the solutions.

The Maximum likelihood function (see (5.1)) is

$$l(\alpha,\sigma) = \log\alpha - \log\sigma - (\alpha+1)\frac{1}{\widetilde{N}_M}\sum_{i=1}^{\widetilde{N}_M}\log\left(1 + \frac{X_i\widetilde{q}(r,\widetilde{T}_i) + \widetilde{d}(r,\widetilde{T}_i)}{\sigma}\right)$$

with the maximizer $(\widehat{\alpha}_M(r), \widehat{\sigma}_M(r))$, which exists by Proposition 5.1, fulfilling $\frac{\partial l}{\partial\alpha} = 0$ and $\frac{\partial l}{\partial\sigma} = 0$. Taking limits $M \to \infty$ and going to the continuous case using Theorem 4.1, the maximizer $(\alpha(r),\sigma(r))$ fulfill the maximum likelihood equations

$$\frac{1}{\alpha(r)} - \int_0^1\int_0^\infty \log\left(1 + \frac{xq(r,z)+d(r,z)}{\sigma(r)}\right) dF_{(\alpha^*,\sigma^*)}(x)\,dG(z) = 0 \tag{6.1a}$$

$$-1 + (\alpha(r)+1)\int_0^1\int_0^\infty \frac{x}{\sigma(r)q(r,z)+d(r,z)+x} dF_{(\alpha^*,\sigma^*)}(x)\,dG(z) = 0. \tag{6.1b}$$

Remark 6.1. For $r = 0$ the solution reduces to $(\alpha(0),\sigma(0)) = (\alpha^*,\sigma^*)$, since we have assumed $q(0,t) = 1$ and $d(0,t) = 0$.

Remark 6.2. For $G(z) = z$, i.e. the arrival times are uniformly distributed in the time interval $[0,1]$, the arrival process is given by a homogeneous Poisson process with the same limiting distribution as the equally spaced arrivals treated by Grandits and Temnov [3].

Applying the implicit function theorem to get the derivatives of equations (6.1) with respect to r at $r = 0$ leads to the first-order terms of the Taylor expansion fulfilling

$$-\frac{(\alpha^*+1)}{(\alpha^*)^2}\alpha_r(0) + \frac{1}{\sigma^*}\sigma_r(0) = \mathbf{E}[q'(0,z)] + \mathbf{E}[d'(0,z)]\frac{\alpha^*}{\sigma^*}$$

$$\frac{(\alpha^*+2)}{(\alpha^*+1)\alpha^*}\alpha_r(0) - \frac{1}{\sigma^*}\sigma_r(0) = -\mathbf{E}[q'(0,z)] - \mathbf{E}[d'(0,z)]\frac{(\alpha^*+1)}{\sigma^*},$$

where the expectation is meant w.r.t. $G(z)$.

The unique solution of this system is

$$\alpha_r(0) = 0 + \frac{(\alpha^*+1)(\alpha^*)^2}{\sigma^*}\mathbf{E}[d'(0,z)] \tag{6.3a}$$

$$\sigma_r(0) = \sigma^*\mathbf{E}[q'(0,z)] + (\alpha^* + (1+\alpha^*)^2)\mathbf{E}[d'(0,z)]. \tag{6.3b}$$

As long as $\mathbf{E}[d'(0,z)] \neq 0$, both α and σ are in general affected by first-order terms. Clearly, any first-order effect in α comes purely from the shift term $d(r,z)$. In [3], where the claim sizes are assumed to be of the form $Y_i = X_i e^{riT/n}$, no such first-order term is obtained for the expansion of $\alpha(r)$.

In these cases with $\alpha_r(0) = 0$, we are also interested in the size and form of the second-order corrections. For simplicity, we only investigate two special cases. The first is the pure inflation case $Y_i = X_i e^{rT_i}$ with constant rate r, where all Y_i are iid. random variables with claim times T_i following the inhomogeneous Poisson process described above. The general case with $q(r,t) \neq 0$ but $\mathbf{E}[q'(0,z)] = 0$ works similar, but includes several cross-terms and cross-moments of the derivatives of d and q. As the general form does not provide any new insight, we refrain from reproducing it here and only look at the inflation case and also at the case of inflation plus constant shift, which will be described further.

(1) Let us turn to the case of inflation $Y_i = X_i e^{rT_i}$. As $d(r,z) = 0$, the first-order effects (6.3) simplify to

$$\alpha_r(0) = 0 \quad \text{and} \quad \sigma_r(0) = \sigma^*\mathbf{E}[z],$$

which is the same form as in the case with equally-spaced claim times.

To obtain the second-order terms in the Taylor expansions of $\alpha(r)$ and $\sigma(r)$ at $r = 0$, the implicit function theorem has to be applied once more to arrive at the equations

$$-\frac{1}{(\alpha^*)^2}\alpha_{rr}(0) + \frac{1}{(\sigma^*)(\alpha^*+1)}\sigma_{rr}(0) = \frac{1}{(\alpha^*+1)}\mathbf{E}[z]^2 + \frac{\alpha^*}{(\alpha^*+2)(\alpha^*+1)}\mathrm{Var}(z)$$

$$-\frac{1}{(\alpha^*+1)\alpha^*}\alpha_{rr}(0) + \frac{1}{(\alpha^*+2)\sigma^*}\sigma_{rr}(0) = \frac{1}{(\alpha^*+2)}\mathbf{E}[z]^2 + \frac{(\alpha^*-1)}{(\alpha^*+2)(\alpha^*+3)}\mathrm{Var}(z)$$

with the unique solution

$$\alpha_{rr}(0) = -2\mathrm{Var}(z)\frac{(\alpha^*+1)^2(\alpha^*)^2}{(\alpha^*+3)(\alpha^*+2)} \tag{6.5a}$$

$$\sigma_{rr}(0) = \sigma^*\left(\mathbf{E}[z]^2 - \frac{2(\alpha^*)^2+\alpha^*+1}{(\alpha^*+3)}\mathrm{Var}(z)\right). \tag{6.5b}$$

For T_i uniformly distributed in $[0,1]$, i.e. $G(z) = z$, we again obtain the results of Grandits and Temnov.

(2) Another useful case worthy to be mentioned is the combination $Y_i = X_i e^{rT_i} + ArT_i$ of inflation and linear shift with some positive constant A (a realistic range of A is $A \in [0,1]$). For simplicity, we keep the assumption of T_i uniformly distributed in $[0,1]$ for this case. For the first order effect in α in σ, we have the following solutions

$$\alpha_r(0) = \frac{(\alpha^*+1)(\alpha^*)^2}{\sigma^*}\frac{A}{2}$$
$$\sigma_r(0) = \frac{\sigma^*}{2} + (\alpha^* + (1+\alpha^*)^2)\frac{A}{2}.$$

For this case, we are interested in finding the second order corrections as well. Like in the pure inflation case, we get back to equations (6.1), apply the implicit function theorem twice and calculate the second derivatives with respect to r at $r = 0$. Expressions become significantly more complicated than in the pure inflation case, but one can check that the unique solution of the corresponding system of equations is

$$\begin{cases} \alpha_{rr}(0) = -\dfrac{1}{6}\dfrac{1+\alpha^*}{(2+\alpha^*)(3+\alpha^*)}\dfrac{(\alpha^*)^2}{(\sigma^*)^2}[(\sigma^*)^2(1+\alpha^*) + 2A\sigma^*((\alpha^*)^2+4\alpha^*+5) \\ \qquad + A(25+53\alpha^*+18(\alpha^*)^2 - 7(\alpha^*)^3 - 3(\alpha^*)^4)] \\ \sigma_{rr}(0) = \dfrac{1}{6\sigma^*(\alpha^*+3)}[(\sigma^*)^2(4+\alpha^*-(\alpha^*)^2) + A\sigma^*(4+14\alpha^* \\ \qquad + 9(\alpha^*)^2 + (\alpha^*)^3) + A^2(3(\alpha^*)^5 + 10(\alpha^*)^4 - 30(\alpha^*)^2 - 31\alpha^* - 8)] \end{cases} \tag{6.7}$$

7. Numerical Examples

We present some examples to illustrate the effect caused by ignoring the trends in the data or misestimating its parameters. For the examples presented below we used realistic values of the Pareto parameters, as well as the ranges of trend parameters which one may meet in insurance practice. Analysis of loss models with parameters changing in time is addressed quite often in the literature. For example, in [5] some simple (mostly linear) trends of the parameters of generalized Pareto distribution are considered in an insurance context. In [1], trends of the parameters of loss distributions are analyzed on such real data examples as sea-level data or rainfall data. Those examples deal mostly with trends in the location and scale parameters.

In our notation, $q(r,t)$ corresponds to the trend in the scale parameter, while $d(r,t)$ describes the trend in the location. This connection allows us to determine the possible range of the rate r from the literature.

For two examples below, we add a new parameter T to the model, corresponding to the number of years of observation. In fact, this means extending the unit time interval of observations to $[0,T]$, $T \in \mathbb{N}$. One can easily check that the results for the first- and second-order corrections obtained above keep the same form, one just has to multiply the values of the effects by T. For instance, in the inflation case we have $\sigma_r(0) = \sigma^* T\mathbf{E}[z]$ (still with 0 for $\alpha_r(0)$), and likewise for the representations (6.3), (6.5) and (6.7).

Example 1. Inflation case. We first consider the constant inflation case $Y_i = X_i e^{rT_i}$. We assume that arrival times T_i follow a homogeneous Poisson process. Let $\alpha^* = 1$, $\sigma^* = 0.2$ and $T = 5$. We are interested in how the bias of the parameters changes with the change of the inflation rate from 0 to $r = 0.04$ (the maximum of this range corresponds to the inflation rate of 4% over a time interval of 5 years).

Note that for the homogeneous Poisson process, arrival times in the interval $[0,T]$ constitute the points of a uniform ordered sample in $[0,T]$ (see e.g. [4]). That gives us again simple representations $\mathbf{E}[z] = \frac{T}{2}, \mathrm{Var}(z) = \frac{T^2}{12}$. Using (6.3) and (6.5), we calculate $(\alpha_r(0), \sigma_r(0))$ and $(\alpha_{rr}(0), \sigma_{rr}(0))$.

To get an idea how precise our obtained expansion is, we compare it with results of Monte Carlo simulations. Specifically, the sequence Y_i was simulated for given parameters and plugged into the likelihood expression (3.1). Thus, the result of its maximization is the biased estimator $(\overline{\alpha}, \overline{\sigma})$, and the empirical bias can be compared with the theoretical one provided by our expansion. The theoretical bias up to second order is the absolute change of the parameters: $\Delta\sigma = \sigma_r(0) \cdot r + \frac{\sigma_{rr}(0)}{2} r^2$, $\Delta\alpha = \alpha_{rr}(0) \cdot \frac{r^2}{2}$.

For each point, 10 simulations of 5000 observations with a time horizon of 5 years was made. We get the results indicated in Table 1, where $(\overline{\alpha}, \overline{\sigma})$ stand for the observed values of parameters, while $R(\overline{\alpha})$ and $R(\overline{\sigma})$ are the values of the observed bias after corrections, i.e. $R(\overline{\sigma}) = (\overline{\alpha} - (\alpha^* + \Delta\alpha))$ and similar for $R(\overline{\sigma})$.

Example 2. Inflation plus linear trend case. We deal now with the model $Y_i = X_i e^{rT_i} + ArT_i$ with arrival times T_i following a homogeneous Poisson process. We

Table 1.

	α			σ		
r	$\overline{\alpha}$	$\Delta\alpha$	$R(\overline{\alpha})$	$\overline{\sigma}$	$\Delta\sigma$	$R(\overline{\sigma})$
0.01	0.999952	−0.00004	$8 \cdot 10^{-6}$	0.202	0.002	10^{-8}
0.02	0.99971	−0.00028	10^{-5}	0.21017	0.01017	10^{-7}
0.03	0.9993	−0.00071	-10^{-5}	0.21642	0.01642	10^{-7}
0.04	0.99888	−0.00113	-10^{-5}	0.22065	0.02065	$2 \cdot 10^{-7}$

Table 2.

	α				σ			
A	$\overline{\alpha}$	$\Delta_1\alpha$	$\Delta_2\alpha$	$R(\overline{\alpha})$	$\overline{\sigma}$	$\Delta_1\sigma$	$\Delta_2\sigma$	$R(\overline{\sigma})$
0.1	1.0826	0.1	−0.0186	−0.002	0.270797	0.07	0.0007	0.0001
0.15	1.115	0.15	−0.0363	0.0012	0.2943	0.095	−0.001	0.0003
0.2	1.1495	0.20	−0.06	0.0095	0.31796	0.12	−0.004	0.002
0.25	1.1854	0.25	−0.09	0.0253	0.33905	0.145	−0.008	0.002

are interested in the change of parameters for some rate r, say $r = 0.04$, and for some range of A. Using (6.3) and (6.7), one can calculate the first and second order effects of the trend. We use Monte Carlo simulations again to check the correspondence between observed the predicted effects of the trend. Denoting the first order changes as $(\Delta_1\alpha, \Delta_1\sigma)$, the second order as $(\Delta_2\alpha, \Delta_2\sigma)$, the observed parameters as $(\overline{\alpha}, \overline{\sigma})$ and the observed bias after the first and second order corrections as $(R(\overline{\alpha}), R(\overline{\sigma}))$ we get results indicated in Table 2.

We observe that the role of the second order corrections become more significant with the growth of A, especially for the parameter α.

Example 3. The change of upper quantiles. Consider again the model $Y_i = X_i e^{rT_i} + ArT_i$ and the parameters as in Example 2. In risk management, it is common to trace the effect on the upper quantile of the severity distribution. With values of $(1.0, 0.2)$ for the true parameters, we consider 0.99-VaR, i.e. 0.99-quantiles of the single-loss distribution, as well as for the aggregate loss distribution. Clearly, if $X \overset{d}{=} \text{Pareto}(\alpha^*, \sigma^*)$ then the VaR of the detrended observations satisfies the following balance equation

$$\mathbf{P}(X \le VaR) = 0.99.$$

One is interested, however, in the VaR inflated to the endpoint of the observations, and hence the true $\widetilde{VaR}$ satisfies

$$\mathbf{P}(Xe^{rT} + ArT \le \widetilde{VaR}) = 0.99.$$

The shifted (observed) VaR_1 is defined from the distribution having the observed parameters $(\overline{\alpha}, \overline{\sigma})$ (which we obtain in this example via Monte Carlo simulations), so that if $Z \overset{d}{=} \text{Pareto}(\overline{\alpha}, \overline{\sigma})$ then

$$\mathbf{P}(Z \le VaR_1) = 0.99.$$

Figures 3 and 4 illustrate the growth of the difference between the true and the observed Values-at-Risk with the increase of the yearly inflation rate for single-loss and for the aggregate loss distributions with Poisson parameter 12 correspondingly. The aggregate loss distributions were calculated via the approach based on fast Fourier transform using aliasing reduction techniques.

Both figures illustrate the similar effect. Particularly, one observes the underestimation of the true VaR for the single loss distribution, as well as for the aggregate

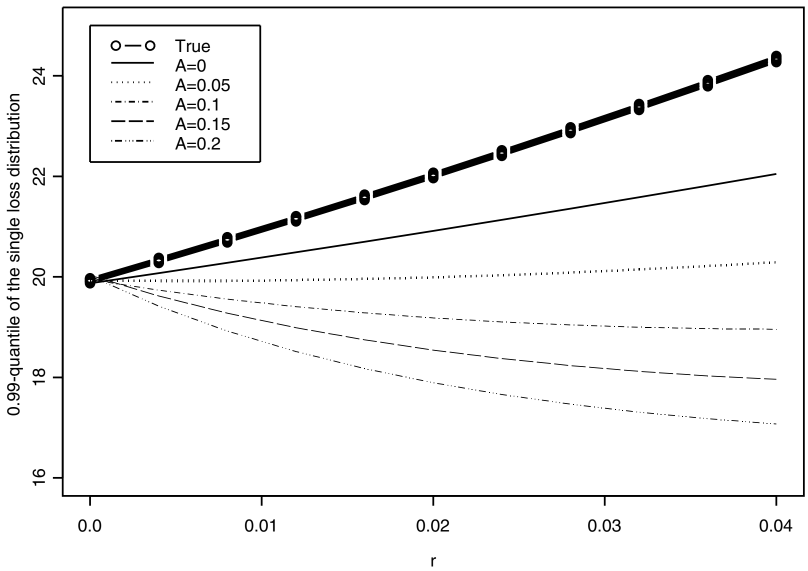

Fig. 3. Comparison of the quantities of VaR_1 and $\widetilde{VaR}$ for the single loss distribution. Biased quantiles VaR_1 are presented by solid or dotted lines for the range of parameters A. Values of true quantiles $\widetilde{VaR}$ are marked as lines with circles on the plot and indicated for the same range of A.

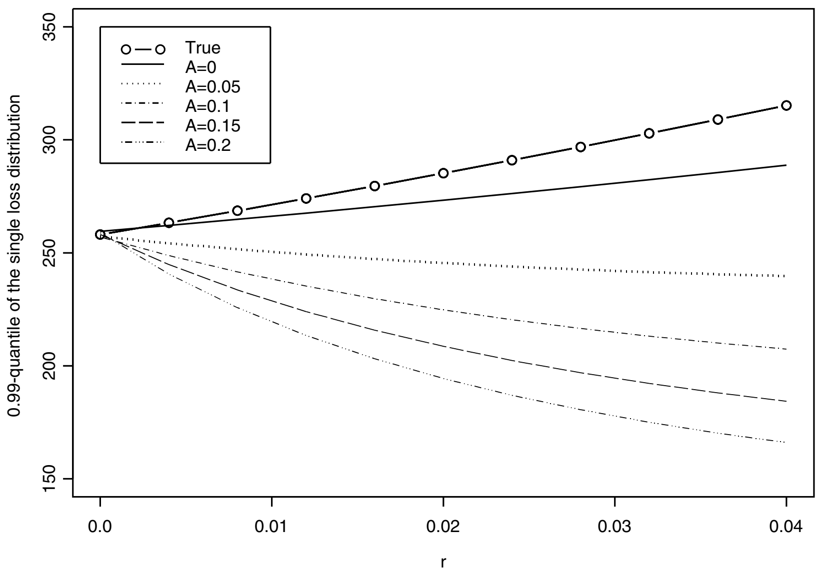

Fig. 4. Comparison of the quantities of VaR_1 and $\widetilde{VaR}$ for the aggregate loss distribution with Poisson parameter 12. Biased quantiles VaR_1 are presented by solid or dotted lines for the range of parameters A. Values of true quantiles $\widetilde{VaR}$ are marked as lines with circles on the plot and indicated for the same range of A (the differences between the true quantiles are almost negligible on this picture).

loss distribution (especially for larger values of r), which becomes more significant with the growth of A. This is in correspondence with theoretical results obtained via our expansion.

We postpone some more numerical examples based on empirical data to future work.

8. Conclusion and Outlook

As mentioned in the introduction, this article develops and generalizes basic results of [3]. Though models described in the present work can be applicable in practice, some further generalization is still to be done. Specifically, in practical applications the Pareto distribution is often used as a model for the severity of losses exceeding some certain threshold μ (see e.g. [2]). Clearly, in this case the affine transformation we have used, namely $X \to X_i q_i + d_i$, has also to change the threshold μ with time: $\mu \to \mu(t)$. However, usually (e.g. in practical risk management) the loss data is collected as exceedances over some *fixed* threshold (say, the same magnitude μ). Due to this, a change of the frequency of observed losses appears. Thus, one has to handle a complex problem arising from the change of both severity and frequency of observed events in time.

Examples presented in this article illustrate that certain trends in the data may bias the estimators of the severity parameters significantly. De-trending the data would be the only right way to avoid a misestimation of the severity parameters, but the true inflation rates are often not known exactly. Besides, one may be interested in forecasting the loss distribution for the next time periods, when it is hard to predict, whether certain trends of the data will be also applicable in the future. Those are the cases, where our results can help in quantifying the effect of the possible trend parameter misprediction on the estimators of the severity parameters and, consequently, on risk measures. In our future work we plan to analyze some real data examples and illustrate the corresponding effects of the trends.

Acknowledgments

Support by the "Christian Doppler Gesellschaft", project "PRisMa Lab", is gratefully acknowledged. For the third author, a significant part of the work was made within "PRisMa Lab" in TU Vienna. From November 2008, the author works under SFI research grant 07/MI/008 associated with Edgeworth Centre for Financial Mathematics.

Appendix A

Proof of Lemma 5.2

The proof is almost completely analogous to the proof of Lemma 3.10 in [3]. The only small modification appears in the case $\alpha^* \in]2, \infty[$. There one has to show the positivity of a certain coefficient in the asymptotic expansion of the limiting log-likelihood L. Note that the case $\alpha^* \in]2, \infty[$ is the only one, where the exact values of the coefficients play a rôle. Therefore only in this case the functions $q(r, z)$ and $d(r, z)$ are relevant. The inequality to be shown in [3]

$$\frac{(\alpha^* - 1)}{\rho_1 \sigma^*} \cdot \frac{\rho_2 (\sigma^*)^2}{(\alpha^* - 1)(\alpha^* - 2)} - \frac{\rho_1 \sigma^*}{(\alpha^* - 1)} > 0,$$

has to be replaced by

$$\frac{(\alpha^* - 1)\rho_{02}}{2\rho_{10}\sigma^* + 2\rho_{01}(\alpha^* - 1)} + \frac{\rho_{11}\sigma^*}{\rho_{10}\sigma^* + \rho_{01}(\alpha^* - 1)} + \frac{\rho_{20}(\sigma^*)^2}{(\rho_{10}\sigma^* + \rho_{01}(\alpha^* - 1))(\alpha^* - 2)} - \rho_{01} - \rho_{10}\frac{\sigma^*}{\alpha^* - 1} > 0,$$

where the ρ's are defined in Lemma 5.1. Since all the denominators are positive by assumption, this is equivalent to

$$(\alpha^* - 1)\rho_{02} + 2\rho_{11}\sigma^* + \frac{2\rho_{20}(\sigma^*)^2}{(\alpha^* - 2)} - 4\rho_{01}\rho_{10}\sigma^* - 2\rho_{01}^2(\alpha^* - 1) - \frac{2\rho_{10}^2(\sigma^*)^2}{(\alpha^* - 1)} > 0.$$

Now let $D := \frac{2\rho_{20}(\sigma^*)^2}{(\alpha^* - 2)} - \frac{2\rho_{10}^2(\sigma^*)^2}{(\alpha^* - 1)}$. Except for D, the absolute value of all other terms on the left hand side of the inequality above can be made arbitrarily small by choosing r small enough. We finally show that $D >$ const. > 0, where the constant can be chosen independently of r. We have by Cauchy-Schwarz

$$D = \frac{2\rho_{20}(\sigma^*)^2}{(\alpha^* - 2)} - \frac{2\rho_{10}^2(\sigma^*)^2}{(\alpha^* - 1)} \geq 2\rho_{20}(\sigma^*)^2 \left(\frac{1}{\alpha^* - 2} - \frac{1}{\alpha * -1} \right) = \text{const.} > 0.$$

The constant on the r.h.s. can be chosen independently of r, since $q(r, z)$ tends uniformly to 1, for $r \to 0$.

References

[1] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in. Statistics (Springer-Verlag, London Ltd., London, 2001).

[2] P. Embrechts, C. Klüppelberg and T. Mikosch, *Modelling Extremal Events*, Volume 33 of Applications of Mathematics, New York (Springer-Verlag, Berlin, 1997).

[3] P. Grandits and G. Temnov, A consistency result for the two-dimensional Pareto distribution in the presence of mis-specified inflation. To appear in *Finance and Stochastics*. Available at http://euclid.ucc.ie/pages/staff/Gregory/gregory.htm (2009).

[4] T. Mikosch, *Non-Life Insurance Mathematics. An Introduction with Stochastic Processes* (Springer, 2004).

[5] A. J. McNeil and T. Saladin, Developing scenarios for future extreme losses using the POT method, in *Extremes and Integrated Risk Management* (Risk books, London, 2000).