ELSEVIER

Insurance: Mathematics and Economics 32 (2003) 115–133

INSURANCE
MATHEMATICS & ECONOMICS

www.elsevier.com/locate/econbase

Influence functions of empirical nonparametric estimators of net reinsurance premiums☆

Vytaras Brazauskas*

Department of Mathematical Sciences, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA

Received 1 September 2001; received in revised form 1 October 2002; accepted 6 November 2002

Abstract

In this paper we illustrate the usefulness of influence functions for studying robustness properties of empirical nonparametric estimators of net premiums of various reinsurance treaties. (Proportional, stop-loss, excess-of-loss, largest claims, and excédent du coût moyen relatif (ECOMOR) type reinsurance treaties are considered.) We use the gross error sensitivity (GES) and the upper breakdown point (UBP) as robustness criteria. It is found that empirical nonparametric estimators are not robust with respect to GES and UBP. Alternative methods of estimation are suggested.

MSC: IM10; IM52; IB90

Keywords: Breakdown point; Empirical nonparametric; Influence function; Reinsurance premium; Robustness

1. Introduction

The *net premium principle* (also known as *principle of equivalence*) is a popular choice among premium calculation principles in the actuarial literature. It has been investigated in applications by Beirlant et al. (1996), Embrechts et al. (1997), Hipp (1996), Reiss and Thomas (1997), and others. Also, it is one of a few principles that satisfy all five "desirable properties" discussed by Gerber (1979), Section 5.3: *nonnegative safety loading, no ripoff, consistency, additivity*, and *iterativity*.

Based on the net premium, which is defined as the expected value of the total claim amount for an individual or a portfolio of risks, the reinsurer calculates the total premium that must be paid by the policy holder. (In this case, the policy holder is another insurance company, the *ceding* company, originally covering the risk.) Therefore a reliable estimate of the *cumulative distribution function* (CDF) of the total claim amount is necessary. A brief discussion of typical methodologies for estimation of the CDF of the total claim amount and consequently net premiums of various reinsurance treaties is provided in Section 3.

In this paper the reliability of *empirical nonparametric* estimators of the net premiums is investigated by application of robust statistics theory. An overview of robust statistical methods based on the *influence function*

☆ Paper Presented at the 5th Conference on Insurance: Mathematics and Economics, Penn State, 23–25 July 2001.

* Tel.: +1-414-229-5656; fax: +1-414-229-4907.
E-mail address: vytaras@uwm.edu (V. Brazauskas).

0167-6687/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0167-6687(02)00207-X

approach is given in Section 2. Standard robustness tools, such as *breakdown point*, *gross error sensitivity*, and *change-of-variance sensitivity* are defined. The use of these concepts is illustrated in Fig. 1.

In Section 4, reinsurance contracts are introduced and described. In particular, *proportional*, *stop-loss*, *excess-of-loss*, *largest claims*, and ECOMOR type treaties are considered. The net reinsurance premiums of these treaties are then expressed as statistical functionals of the underlying claims distribution. Finally, corresponding empirical non-parametric estimators of the net premiums are derived and their robustness properties are investigated using methods in Section 2.

In Section 5, behavior of robust and nonrobust estimators of the net proportional reinsurance premium under *contaminated* data scenario is examined. Specifically, for severity of claims we choose a single-parameter Pareto model and mix it with another Pareto model with different parameters. The *trimmed mean* and *generalized median* type estimators are considered for robust estimation of parameters. It is shown that these estimators are significantly less affected by contaminating observations than empirical nonparametric estimators.

Conclusions are drawn and final remarks are made in Section 6.

2. Robustness: the influence function approach

The influence function approach, also known as the "infinitesimal approach", was initiated more than three decades ago by Hampel (1968). And it is based on three fundamental concepts: qualitative robustness, influence function (with a variety of derived quantities), and breakdown point. Loosely speaking, they correspond to continuity and derivative of a function, and to the distance to its nearest singularity, respectively.

The role of qualitative robustness in this setup is as follows. First, it requires that if the level of contamination in the data gradually decreases (or increases) so should behave the bias and variance of the estimator. Second, if the data is perfectly "clean", i.e., the level of contamination is zero, then the estimator must behave like it does under the pre-supposed ideal model. Based on this, one may notice that qualitative robustness is a necessary but rather weak robustness condition. The weakness of this concept is its inability to discriminate between two qualitatively robust estimators. Fortunately, the necessary ranking and comparison of robust procedures can be accomplished using quantitative robustness measures, i.e., the "influence function" (and derived quantities) and the "breakdown point". Let us examine these concepts precisely.

2.1. The influence function

For a parameter $\mathrm{H}(F)$ estimated by $\mathrm{H}(\hat{F}_n)$, where $\hat{F}_n$ is an empirical CDF estimating a CDF F on the basis of a sample of size n, the associated influence function is defined for real-valued x as

$$\mathrm{IF}(x;\mathrm{H}(F)) = \frac{\partial}{\partial s}[\mathrm{H}((1-s)F + s\Delta_x)]_{s=0^+},$$

where Δ_x is the distribution placing all mass at the point x. To interpret this function and understand its role, one uses functional analysis and Taylor expansion of the functional $\mathrm{H}(F)$ in the space of CDFs F, which leads to a first order approximation to the estimation error in terms of the IF:

$$\mathrm{H}(\hat{F}_n) - \mathrm{H}(F) \approx \frac{1}{n}\sum_{i=1}^{n}\mathrm{IF}(X_i;\mathrm{H}(F)),$$

where $X_1, \ldots, X_n$ denote the sample values. The approximate numerical impact or "influence" of the observation X_i on the error of estimation is then given by $n^{-1}\,\mathrm{IF}(X_i;\mathrm{H}(F))$. In particular, the potential impact of outliers and/or contaminating observations becomes so quantified.

Several useful robustness measures based on IFs has been suggested in the statistical literature (Hampel et al. (1986)). The *gross error sensitivity*

$$\mathrm{GES} = \sup_x|\mathrm{IF}(x;\mathrm{H}(F))|,$$

which (divided by n) measures the worst possible effect on the estimator due to contamination of the data by gross errors.

Another, the *change-of-variance sensitivity*

$$\mathrm{CVS} = \sup_x \frac{\mathrm{CVF}(x;\mathrm{H}(F))}{V(\mathrm{H}(F))},$$

with the *change-of-variance function* (CVF), defined by

$$\mathrm{CVF}(x;\mathrm{H}(F)) = \frac{\partial}{\partial s}[V(\mathrm{H}((1-s)F + s\Delta_x))]_{s=0^+},$$

where $V(\mathrm{H}(F)) = \int [\mathrm{IF}(x;\mathrm{H}(F))]^2\,\mathrm{d}F(x)$ is the asymptotic variance. The CVS is a standardized version of the CVF which for negative values points to a decrease in variance, $V(\mathrm{H}(F))$, indicating a higher accuracy and shorter confidence intervals (which is beneficial). Large positive values of the CVF point to nonrobustness.

Remark. An estimator is called *B-robust* (*V-robust*) if its GES (CVS) is finite. Also, we shall mention here that the concept of V-robustness is stronger than the concept of B-robustness. That is, if an estimator is V-robust then it *must* also be B-robust. Further, if an estimator is not B-robust, then it also is not V-robust. (For a rigorous mathematical treatment of B-robustness versus V-robustness, see Hampel et al. (1986, Section 2.5b.)

2.2. *The breakdown point*

The breakdown point (BP) provides a guidance up to what distance from the model a first order approximation provided by the IF can be used. In practice, the BP is loosely characterized as the largest proportion of corrupted sample observations that the estimator can cope with. Typically, there are two types of contamination that may cause corruption of an estimator: lower and upper contamination. The lower contamination is to take observations to some lower limit (for example, 0), while the most extreme form of upper corruption is to take observations to ∞. Thus, separate versions of *lower* and *upper* BP (LBP, UBP) can (and should) be considered. In the reinsurance context, however, contamination of the lower type is of lesser concern because the lower limit of losses is usually pre-defined by a contract. (For example, the lower limit can be represented as a deductible or a retention level.) Thus, in the present treatment we emphasize UBP as a global robustness criterion. As by Brazauskas and Serfling (2000b) we use the following definition of UBP.

> The *upper breakdown point* (UBP) is the largest proportion of *upper* sample observations which may be taken to an upper limit without taking the estimator to an uninformative limit not depending on the parameter being estimated.

Clearly, the above notion of BP is constructed around the idea of infinite *bias* of the estimator. In analogy to the bias, a similar concept can be constructed for the variance.

> The *variance breakdown point* is the largest proportion of sample observations which may be corrupted without taking the variance of the estimator to infinity.

See Huber (1981, Sections 1.4 and 6.6) for more detailed discussion of variance BP.

2.3. *An illustration*

In this subsection we illustrate the role of previously presented robustness concepts. Let us start with the description of what might be called an "approximate parametric viewpoint".

Suppose we have a parametric model, which (hopefully) is a good approximation to the true underlying situation. Contrary to the classical parametric viewpoint, we do not assume that the postulated parametric model is *exactly* correct. In other words, we believe in a more realistic approach claiming that our initial parametric guess bears no guarantees that the data will follow it exactly. Such an approximate parametric approach was first introduced in the context of location parameter estimation by Huber (1964), where he suggested the following mathematical formalization of it:

$$F_\varepsilon = (1-\varepsilon)F + \varepsilon G, \tag{1}$$

where G is *any* distribution (or a mixture of distributions) from a neighborhood of possible contaminating distributions, and ε represents the probability that a sample observation comes from the distribution G instead of F. Note that there are no restrictions on G, thus (1) is essentially a nonparametric neighborhood of the distribution F. Eq. (1) is also known as the *gross error model*. (Strictly speaking, in Huber's case G is a symmetric distribution, because he considered symmetric F.)

When applying such an approach, however, one should also be prepared for the worst case scenario. That is, for such contaminating distribution G which produces *largest* bias and variance of the estimator. More importantly, reasonable approximations of these are desirable. It turns out that it is possible to approximate the worst bias and variance by means of the GES and CVS, respectively. Let us find these approximations by following derivations by Hampel et al. (1986, Section 2.7). The worst bias can be approximated by

$$\begin{aligned}\sup_G|\mathrm{H}(F_\varepsilon)-\mathrm{H}(F)| &= \sup_G|\mathrm{H}((1-\varepsilon)F+\varepsilon G)-\mathrm{H}(F)| \approx \sup_G\left|\varepsilon\int \mathrm{IF}(x;\mathrm{H}(F))\,\mathrm{d}G(x)\right| \\ &= \varepsilon\times\sup_x|\mathrm{IF}(x;\mathrm{H}(F))| = \varepsilon\times\mathrm{GES},\end{aligned} \tag{2}$$

and variance by

$$\begin{aligned}\sup_G V(\mathrm{H}(F_\varepsilon)) &= \exp\left\{\sup_G[\log V(\mathrm{H}((1-\varepsilon)F+\varepsilon G))]\right\} \\ &\approx \exp\left\{\sup_G\left[\log V(\mathrm{H}(F))+\varepsilon\int\frac{\mathrm{CVF}(x;\mathrm{H}(F))}{V(\mathrm{H}(F))}\,\mathrm{d}G(x)\right]\right\} \\ &= \exp\left\{\log V(\mathrm{H}(F))+\varepsilon\times\sup_x\frac{\mathrm{CVF}(x;\mathrm{H}(F))}{V(\mathrm{H}(F))}\right\} = V(\mathrm{H}(F))\exp\{\varepsilon\times\mathrm{CVS}\}.\end{aligned} \tag{3}$$

In Fig. 1 we illustrate the most extreme behavior of asymptotic bias and variance of robust estimator under the worst type of contamination.

The plot of Fig. 1 illustrates several important points. First, the maximum asymptotic bias and log(variance) approach 0 and $\ln[V(\mathrm{H}(F))]$, respectively, as the level of contamination tends to 0, emphasizing that "qualitatively robust" estimators are preferred. Second, respective BPs tell us up to what amounts of contamination a robust estimator can be trusted and approximations (2) and (3) can be of value. In particular, these approximations seem to be very accurate up to about 1/3 of the corresponding BP and deteriorate rapidly thereafter. (As a rule of thumb, Hampel et al. (1986, p. 178), suggest to use such approximations up to 1/2 of the corresponding BP.) Finally, we should keep in mind that if the estimator is *nonrobust*, i.e., its BP $=0$ and/or GES $=\infty$, then the whole picture collapses, "approximate parametric" approach is not applicable, and the estimator cannot be trusted at all.

3. Estimation methodologies

In this section we briefly discuss typical methodologies for estimation of claims severity CDF and consequently net premiums of various reinsurance treaties which are defined in Section 4. As will be seen there, all reinsurance premiums under consideration can be represented as statistical functionals of the underlying claims distribution F. Therefore, the quality of estimators resulting from the methods discussed in this section can be investigated by applying the robustness tools in Section 2.

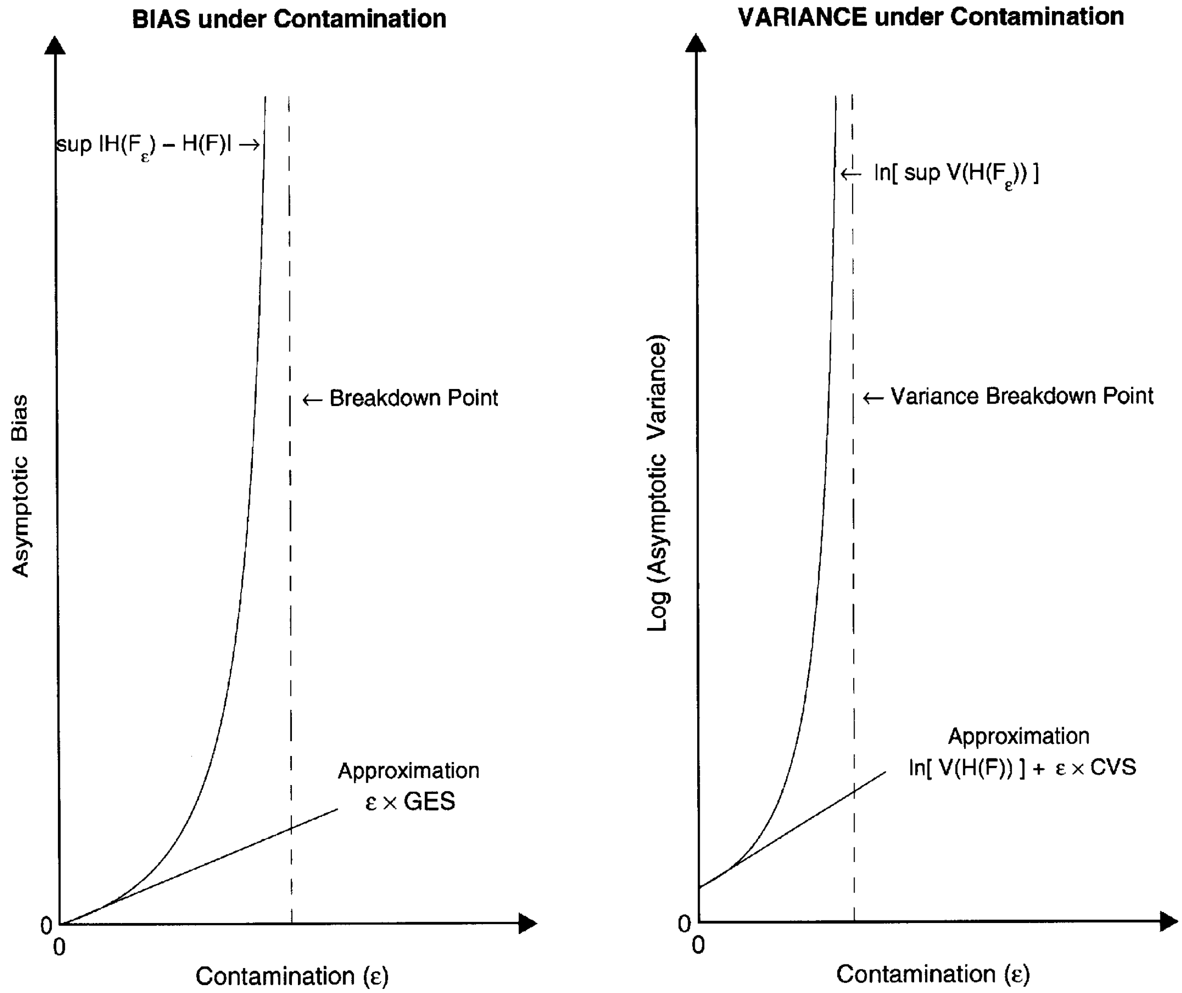

Fig. 1. Illustration of robustness concepts.

3.1. *Parametric approach*

In parametric modeling of F (say, F depends on parameters δ_1 and δ_2) one represents functionals $H(F)$ as explicit functions of the parameters δ_1 and δ_2 and obtains estimates $\widehat{H(F)}$ by substitution of $\hat{\delta}_1$ for δ_1 and $\hat{\delta}_2$ for δ_2. Particular choices of parametric families for claims distribution include: the *gamma* distribution and variants, the *Pareto* distribution and variants, the *Benktander-type-I* and *-type-II* distributions, the *lognormal* distribution, the *Weibull* distribution, and others (see Klugman et al. (1998) and Reiss and Thomas (1997) for more examples of loss distributions).

The strength of parametric approach is 2-fold. First, it is in keeping with the principle of parsimony in modeling. Second, it permits inferences to be made beyond the range of the actual observed data. On the other hand, this methodology depends directly upon parametric assumptions, which may be of questionable validity. Thus, in this setting, a key issue is that of *robustness*, i.e., of insensitivity to small deviations from distributional assumptions.

3.2. Semiparametric approach

Instead of dealing with a whole range of claims, semiparametric methodology restricts attention to the tail of distribution (which corresponds to large claims) modeling small and medium claims by some *slowly varying function*. For this type of modeling, a broad and effective assumption is a *Pareto-type* distribution: a distribution F for which the survival function $1-F(c)$ tends to 0 at a polynomial rate $c^{-\alpha}$ as $c \to \infty$, for some index α. In such a case we have

$$\lim_{c\to\infty} \frac{1-F(cx)}{1-F(c)} = x^{-\alpha}, \tag{4}$$

i.e., the conditional distribution of an observation, given that it exceeds a threshold c, becomes for large c approximately a single-parameter Pareto distribution. Therefore for estimation of α one may apply methods designed for standard Pareto distribution.

This approach is actively advocated and extensively used by Beirlant et al. (1996), Embrechts et al. (1997) among others, to model extremal events in (re)insurance problems. When using semiparametric models in applications, however, one should keep in mind that the right tail of the distribution is assumed to be parametric. Thus, besides possessing the strength of a parametric approach this methodology also shares its weaknesses.

3.3. Robust parametric approach

Robust parametric (or "approximate parametric") approach is a solution that we suggest in order to overcome drawbacks of parametric methodology. That is, in the parametric setting (which was discussed above) we recommend to use *robust estimators* of parameters instead of standard ones. Two types of robust estimators for a one-parameter Pareto distribution are presented in Section 5.

Also, for the parametric part of semiparametric models, similar "approximate" approach can be used for robust estimation of index α in (4).

3.4. Empirical nonparametric approach

Empirical nonparametric approach is intuitively appealing and straightforward to handle. Here a parameter of interest $\mathrm{H}(F)$ of the unknown CDF F is estimated by $\mathrm{H}(\hat{F}_n)$, with $\hat{F}_n$ the empirical CDF based on a sample of size n. With this approach one can avoid parametric distributional assumptions, and therefore gain some robustness in return. The empirical nonparametric methodology, however, may turn out to be highly deficient in modeling large claims due to a scarcity of sample data in the tails.

In the actuarial context, this approach is used, for example, by Reiss and Thomas (1997) to estimate the net premium of the excess-of-loss reinsurance treaty. Also, Hipp (1996) uses the so-called Delta Method to derive asymptotic distribution of empirical nonparametric estimators of various actuarial parameters including the net excess-of-loss reinsurance premium. We examine robustness properties of this methodology in detail in the next section.

4. Empirical nonparametric estimators of the net reinsurance premiums

In Sections 4.2–4.6, we first introduce reinsurance treaties and the associated net premiums. Then we reexpress these premiums as statistical functionals of the underlying claims distribution. Finally, we define corresponding empirical nonparametric estimators and evaluate their GESs and UBPs. For description of reinsurance treaties, we closely follow Embrechts et al. (1997, Section 8.7).

4.1. Standard setting

As by Brazauskas (2000) we consider the homogeneous risk model. That is:

The individual claim sizes, $X_1, X_2, \ldots,$ are independent identically distributed nonnegative random variables with common distribution function F, independent of the number N of claims occurring over a specified time period, for example, a year. The total claim amount of an insurance portfolio is then given by

$$S_N = \sum_{i=1}^{N} X_i.$$

Typically, the number of claims N is assumed to be distributed according to Poisson, binomial or negative binomial distribution. In this paper, we focus on estimation of claim sizes treating the quantities associated with N as known parameters. Also, we assume that $\mathbf{E}(X_1)$ exists.

4.2. *Proportional or quota share reinsurance*

This is a common form of reinsurance for claims of "moderate" size. Here simply a fraction $p \in (0, 1)$ of each claim (hence the pth fraction of the whole portfolio) is covered by the reinsurer. Thus the reinsurer pays for the amount

$$R_{\mathrm{P}} = p \sum_{i=1}^{N} X_i,$$

whatever the size of the claims.

The parameter of interest (the net premium) is the expected value of R_{P}, i.e., $\mathbf{E}(R_{\mathrm{P}})$. It follows from Section 4.1 that $\mathbf{E}(R_{\mathrm{P}})$ can be rewritten as

$$\mathbf{E}(R_{\mathrm{P}}) = \mathbf{E}\left\{ p \sum_{i=1}^{N} X_i \right\} = p\mathbf{E}(N)\,\mathbf{E}(X_1) = p\mathbf{E}(N) \int x \,\mathrm{d}F(x).$$

Thus, the statistical functional corresponding to the net proportional premium is

$$H_1(F, p) = p\mathbf{E}(N) \int x \,\mathrm{d}F(x),$$

and its empirical nonparametric estimator is

$$H_1(\hat{F}_n, p) = p\mathbf{E}(N) \int x \,\mathrm{d}\hat{F}_n(x) = p\mathbf{E}(N)\bar{X}, \tag{5}$$

where $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ is the sample mean.

Next, we find the influence function of $H_1(F, p)$. The following steps are easily justified:

$$\begin{aligned} \mathrm{IF}(x; H_1(F, p)) &= \frac{\partial}{\partial s}[H_1((1-s)F + s\Delta_x, p)]_{s=0^+} = \frac{\partial}{\partial s}\left[p\mathbf{E}(N) \int y \,\mathrm{d}\{(1-s)F(y) + s\Delta_x(y)\}\right]_{s=0^+} \\ &= p\mathbf{E}(N)\frac{\partial}{\partial s}\left[(1-s)\int y \,\mathrm{d}F(y) + sx\right]_{s=0^+} = px\,\mathbf{E}(N) - H_1(F, p). \end{aligned} \tag{6}$$

Further, using expression (6) we can evaluate the GES of $H_1(\hat{F}_n, p)$. In particular

$$\mathrm{GES} = \sup_x |\mathrm{IF}(x; H_1(F, p))| = \sup_x |px\,\mathbf{E}(N) - H_1(F, p)| = \infty.$$

Finally, we evaluate the UBP of $H_1(\hat{F}_n, p)$. For any *fixed n*, if even a single X_i is taken to ∞, then $\bar{X} \to \infty$ and consequently $H_1(\hat{F}_n, p) \to \infty$. That is, corruption of a single data value by upper contamination can render the estimator completely uninformative. Thus, the estimator (5) can cope with 0% of "bad" observations in a sample and therefore its UBP is 0. Obviously, GES $= \infty$ in conjunction with UBP $= 0$ implies that $H_1(\hat{F}_n, p)$ is *nonrobust*.

4.3. Stop-loss reinsurance

The reinsurer covers losses in the portfolio exceeding a well defined limit K, the so-called *ceding company's retention level*. This means that the reinsurer pays for the amount

$$R_{\rm SL} = (S_N - K)^+ = \max\{0, S_N - K\}.$$

This type of reinsurance is useful for protecting the company against insolvency due to excessive claims on the coverage.

The parameter of interest $\mathbf{E}(R_{\rm SL})$ can be rewritten as

$$\mathbf{E}(R_{\rm SL}) = \mathbf{E}(S_N - K)^+ = \mathbf{E}_N\{\mathbf{E}[(S_N - K)^+ | N = n]\} = \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\mathbf{E}\left(\sum_{i=1}^{n} X_i - K\right)^+\right]$$
$$= \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\int (x - K)^+ \,\mathrm{d}F^{*n}(x)\right],$$

where F^{*n} denotes the n-fold convolution of F, i.e., $F^{*n}(x) = \mathbf{P}\{\sum_{i=1}^{n} X_i \le x\}$.

Thus, the statistical functional corresponding to the net stop-loss premium with retention level K is

$$H_2(F, K) = \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\int (x - K)^+ \,\mathrm{d}F^{*n}(x)\right],$$

and its empirical nonparametric estimator is

$$H_2(\hat{F}_n, K) = \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\int (x - K)^+ \,\mathrm{d}\hat{F}^{*n}(x)\right]$$
$$= \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}n^{-n} \sum_{1\le i_1,\dots,i_n\le n} (X_{i_1} + \cdots + X_{i_n} - K)^+\right]. \quad (7)$$

Next, in order to derive the influence function we first have to reexpress functional $H_2(F_s, K) = H_2((1 - s)F + s\Delta_x, K)$ into more tractable form. Using the fact that

$$F_s^{*n} = (1 - s)^n F^{*n} + snF^{*(n-1)} * \Delta_x + o(s) = (1 - sn)F^{*n} + snF^{*(n-1)} * \Delta_x + o(s),$$

we have

$$H_2(F_s, K) = \sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\int (y - K)^+ \,\mathrm{d}F_s^{*n}(y)\right]$$
$$= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N = n\}\left[(1 - sn)\int (y - K)^+ \,\mathrm{d}F^{*n}(y) + sn\int (y - K)^+ \,\mathrm{d}(F^{*(n-1)} * \Delta_x(y)) + o(s)\right]\right\}$$
$$= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N = n\}\left[(1 - sn)\int (y - K)^+ \,\mathrm{d}F^{*n}(y) + sn\int (x + y - K)^+ \,\mathrm{d}F^{*(n-1)}(y) + o(s)\right]\right\}. \quad (8)$$

Now using (8) we find the influence function of $H_2(F, K)$

$$\mathrm{IF}(x;\, H_2(F, K)) = \frac{\partial}{\partial s}[H_2(F_s, K)]_{s=0^+} = \sum_{n=1}^{\infty}\left\{n\mathbf{P}\{N=n\}\left[\int (x+y-K)^+\,\mathrm{d}F^{*(n-1)}(y) - \int (y-K)^+\,\mathrm{d}F^{*n}(y)\right]\right\}. \quad (9)$$

It follows from (9) that

$$\mathrm{GES} = \sup_x |\mathrm{IF}(x;\, H_2(F, K))| = \sup_x \left|\int (x+y-K)^+\,\mathrm{d}F^{*(n-1)}(y)\right| = \infty.$$

Finally, let us examine the behavior of estimator (7) when a single X_i is taken to ∞. In such a case, $\sum_{1\le i_1,\dots,i_n\le n}(X_{i_1}+\dots+X_{i_n}-K)^+ \to \infty$ which implies that this estimator has UBP $= 0$. Thus, the estimator $H_2(\hat{F}_n, K)$ is classified as *nonrobust*.

4.4. Excess-of-loss reinsurance

The reinsurance company pays for all individual losses in excess of some limit D, the so-called *priority*. This means that the reinsurer covers the amount

$$R_{\mathrm{EL}} = \sum_{i=1}^{N}(X_i - D)^+.$$

The parameter of interest $\mathbf{E}(R_{\mathrm{EL}})$ can be rewritten as

$$\mathbf{E}(R_{\mathrm{EL}}) = \mathbf{E}\left\{\sum_{i=1}^{N}(X_i - D)^+\right\} = \mathbf{E}(N)\,\mathbf{E}(X_1 - D)^+ = \mathbf{E}(N)\int (x-D)^+\,\mathrm{d}F(x).$$

Thus, the statistical functional corresponding to the net excess-of-loss premium with priority D is

$$H_3(F, D) = \mathbf{E}(N)\int (x-D)^+\,\mathrm{d}F(x),$$

and its empirical nonparametric estimator is

$$H_3(\hat{F}_n, D) = \mathbf{E}(N)\int (x-D)^+\,\mathrm{d}\hat{F}_n(x) = \mathbf{E}(N)\overline{(X-D)^+}, \quad (10)$$

where $\overline{(X-D)^+} = n^{-1}\sum_{i=1}^{n}(X_i - D)^+$.

Next, we derive the influence function of $H_3(F, D)$

$$\mathrm{IF}(x;\, H_3(F, D)) = \frac{\partial}{\partial s}[H_3(F_s, D)]_{s=0^+} = \frac{\partial}{\partial s}\left[\mathbf{E}(N)\int (y-D)^+\,\mathrm{d}\{(1-s)F(y) + s\Delta_x(y)\}\right]_{s=0^+}$$
$$= \mathbf{E}(N)\frac{\partial}{\partial s}\left[(1-s)\int (y-D)^+\,\mathrm{d}F(y) + sx\right]_{s=0^+} = \mathbf{E}(N)(x-D)^+ - H_3(F, D). \quad (11)$$

Further, it follows from (11) that

$$\mathrm{GES} = \sup_x |\mathrm{IF}(x;\, H_3(F, D))| = \sup_x |\mathbf{E}(N)(x-D)^+ - H_3(F, D)| = \infty.$$

Finally, we evaluate the UBP of $H_3(\hat{F}_n, D)$. Using the same arguments as in derivation of UBP for $H_1(\hat{F}_n, p)$ and $H_2(\hat{F}_n, K)$, we arrive at similar conclusion for the empirical nonparametric estimator of the excess-of-loss net premium: UBP $= 0$ and hence the estimator is *nonrobust*.

4.5. Largest claims reinsurance

At the time when the contract is underwritten the reinsurance company guarantees that the k largest claims in a fixed time frame, for example, a year, will be covered. (The number k is either fixed or it grows sufficiently slowly with time.) This means that one has to study the quantity

$$R_{\mathrm{LC}} = \sum_{i=1}^{k} X_{N,N-i+1},$$

where $X_{n,m}$ denotes the mth order statistic in a sample of n ordered observations $X_{n,1} \leq \cdots \leq X_{n,m} \leq \cdots \leq X_{n,n}$.

In this case, the parameter of interest $\mathbf{E}(R_{\mathrm{LC}})$ can be rewritten as

$$\begin{aligned}\mathbf{E}(R_{\mathrm{LC}}) &= \mathbf{E}\left\{\sum_{i=1}^{k} X_{N,N-i+1}\right\} = \mathbf{E}_N\left\{\mathbf{E}\left[\sum_{i=1}^{k} X_{N,N-i+1} | N=n\right]\right\} \\ &= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k}\mathbf{E}(X_{n,n-i+1})\right\} = \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k}\int x\,\mathrm{d}F_{n,n-i+1}(x)\right\},\end{aligned} \quad (12)$$

where $F_{n,m}(\cdot)$ denotes the CDF of mth order statistic, that is,

$$F_{n,m}(x) = \mathbf{P}\{X_{n,m} \leq x\} = \sum_{j=m}^{n}\binom{n}{j}[F(x)]^j[1-F(x)]^{n-j}.$$

Further, the differential of $F_{n,m}$ is given by

$$\mathrm{d}F_{n,m}(x) = n\binom{n-1}{n-m}[F(x)]^{m-1}[1-F(x)]^{n-m}\,\mathrm{d}F(x).$$

Combining the last expression with (12) we arrive at the following expression for the parameter of interest $\mathbf{E}(R_{\mathrm{LC}})$ (and equivalently, for the corresponding statistical functional $H_4(F,k)$):

$$\sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k} n\binom{n-1}{i-1}\int x[F(x)]^{n-i}[1-F(x)]^{i-1}\,\mathrm{d}F(x)\right\}. \quad (13)$$

Thus, the empirical nonparametric estimator of the net k largest claims premium is given by the following formula

$$\begin{aligned}H_4(\hat{F}_n,k) &= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k} n\binom{n-1}{i-1}\int x[\hat{F}_n(x)]^{n-i}[1-\hat{F}_n(x)]^{i-1}\,\mathrm{d}\hat{F}_n(x)\right\} \\ &= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k}\binom{n-1}{i-1}\sum_{j=1}^{n} X_{n,j}\left(\frac{j}{n}\right)^{n-i}\left(1-\frac{j}{n}\right)^{i-1}\right\}.\end{aligned} \quad (14)$$

In order to derive the influence function of $H_4(F, k)$, it will be beneficial to first find the influence function of $\mathrm{E}(X_{n,m}) = H(F; n, m)$. We proceed as follows

$$\begin{aligned}
&\mathrm{IF}(x; H(F; n, m)) \\
&= \frac{\partial}{\partial s}[H(F_s; n, m)]_{s=0^+} = \frac{\partial}{\partial s}\left[n\binom{n-1}{n-m}\int y[F_s(y)]^{m-1}[1-F_s(y)]^{n-m}\,\mathrm{d}\{(1-s)F(y)+s\Delta_x(y)\}\right]_{s=0^+} \\
&= n\binom{n-1}{n-m}\frac{\partial}{\partial s}\left[(1-s)\int y[F_s(y)]^{m-1}[1-F_s(y)]^{n-m}\,\mathrm{d}F(y)\right. \\
&\qquad \left. + sx[1+s(1-F(x))]^{m-1}[(1-s)(1-F(x))]^{n-m}\right]_{s=0^+} \\
&= n\binom{n-1}{n-m}\left[-\int y[F(y)]^{m-1}[1-F(y)]^{n-m}\,\mathrm{d}F(y) + \int y[F(y)]^{m-1}[1-F(y)]^{n-m-1}[(n-m)F(y)\right. \\
&\qquad \left. -(m-1)(1-F(y))]\,\mathrm{d}F(y) + x[F(x)]^{m-1}[1-F(x)]^{n-m}\right] \\
&= -mH(F; n, m) + (n-m)H(F; n, m+1) + n\binom{n-1}{n-m}x[F(x)]^{m-1}[1-F(x)]^{n-m}. \qquad (15)
\end{aligned}$$

Next, we combine expression (15) with (12) and (13), and find that the influence function of $H_4(F, k)$ is given by

$$\begin{aligned}
&\mathrm{IF}(x; H_4(F, k)) \\
&= \frac{\partial}{\partial s}[H_4((1-s)F + s\Delta_x, k)]_{s=0^+} \\
&= \sum_{n=1}^{\infty}\left\{\mathbf{P}\{N=n\}\sum_{i=1}^{k}\left[-(n-i+1)H(F; n, n-i+1) + (i-1)H(F; n, n-i+2)\right.\right. \\
&\qquad \left.\left. + n\binom{n-1}{i-1}x[F(x)]^{n-i}[1-F(x)]^{i-1}\right]\right\}. \qquad (16)
\end{aligned}$$

Further, we note that only the last term in (16) depends on x. Thus, for calculations of the GES, we have

$$\begin{aligned}
\mathrm{GES} &= \sup_x|\mathrm{IF}(H_4(F, k))| \\
&= \text{constant (with respect to } x) + \sup_x\left|\sum_{n=1}^{\infty}\mathbf{P}\{N=n\}\sum_{i=1}^{k}n\binom{n-1}{i-1}x[F(x)]^{n-i}[1-F(x)]^{i-1}\right| = \infty.
\end{aligned}$$

The supremum is infinite since summation $\sum_{i=1}^{k}(\cdots)$ can be reexpressed as

$$x(F(x))^{n-1} + \sum_{i=2}^{k}\binom{n-1}{i-1}x[F(x)]^{n-i}[1-F(x)]^{i-1}, \qquad (17)$$

where the term $x(F(x))^{n-1}$ is unbounded.

Finally, examination of (14) shows that UBP $= 0$ (since for $i = 1$, $H_4(\hat{F}_n, k) \to \infty$ as $X_{n,n} \to \infty$) and hence the estimator $H_4(\hat{F}_n, k)$ is *nonrobust*.

Remarks.

(i) Note that all other terms in (17) are bounded for distributions with a finite first moment. Thus, for example, if a reinsurance treaty was defined as "k largest claims *except* the largest claim", then the GES of the net premium for such a treaty would be finite and, consequently, its empirical nonparametric estimator would be robust.

(ii) Besides being nonrobust the empirical nonparametric estimator given by (14) is also significantly *biased*. This can be easily seen by choosing, for instance, $k = 1$ in formula (14):

$$\mathbf{E}(\widehat{\mathbf{E}(X_{n,n})}) = \mathbf{E}\left(\sum_{j=1}^{n} X_{n,j}\left(\frac{j}{n}\right)^{n-1}\right) = \mathbf{E}(X_{n,n}) + \sum_{j=1}^{n-1}\mathbf{E}(X_{n,j})\left(\frac{j}{n}\right)^{n-1} > \mathbf{E}(X_{n,n}), \quad n > 1.$$

An unbiased (and perhaps better known) nonparametric estimator of $\mathbf{E}(X_{n,j})$ is the jth largest sample observation $X_{n,j}$. And a nonparametric estimator of $\mathbf{E}(R_{\mathrm{LC}})$ based on observations $X_{n,n-j+1}$, $j = 1, \ldots, k$, has an advantage of being unbiased. Therefore, we will use this estimator (instead of empirical nonparametric) in a numerical example of Section 5.

4.6. *Excédent du coût moyen relatif (ECOMOR) reinsurance*

This form of a treaty can be considered as an excess-of-loss reinsurance with a random deductible which is determined by the kth largest claim in the portfolio. This means that the reinsurer covers the claim amount

$$R_{\mathrm{ECO}} = \sum_{i=1}^{N}(X_i - X_{N,N-k+1})^+,$$

for a fixed number $k \geq 2$.

It is straightforward to show that in this case the parameter of interest $\mathbf{E}(R_{\mathrm{ECO}})$ can be rewritten as a function of $\mathbf{E}(X_{N,N-k+1})$ and $\mathbf{E}(R_{\mathrm{LC}})$ (based on $k-1$ largest claims). In particular

$$\begin{aligned}\mathbf{E}(R_{\mathrm{ECO}}) &= \mathbf{E}\left\{\sum_{i=1}^{N}(X_i - X_{N,N-k+1})^+\right\} = \mathbf{E}\left\{\sum_{i=1}^{k-1}(X_{N,N-i+1} - X_{N,N-k+1})\right\}\\ &= \mathbf{E}(R_{\mathrm{LC}}|\text{based on } k-1 \text{ largest claims}) - (k-1)\mathbf{E}(X_{N,N-k+1})\\ &= H_4(F, k-1) - (k-1)\sum_{n=1}^{\infty}\left[\mathbf{P}\{N = n\}\int x\,\mathrm{d}F_{n,n-k+1}(x)\right].\end{aligned}$$

Thus, the associated statistical functional is given by

$$H_5(F, k) = H_4(F, k-1) - (k-1)\sum_{n=1}^{\infty}[\mathbf{P}\{N = n\}H(F; n, n-k+1)],$$

and its empirical nonparametric estimator is

$$H_5(\hat{F}_n, k) = H_4(\hat{F}_n, k-1) - (k-1)\sum_{n=1}^{\infty}[\mathbf{P}\{N = n\}H(\hat{F}_n; n, n-k+1)].$$

Since $H_5(\hat{F}_n, k)$ is a function of $H_4(\hat{F}_n, k-1)$ and $H(\hat{F}_n; n, n-k+1)$, it shares the same robustness properties as these two estimators, i.e., GES $= \infty$ and UBP $= 0$. Thus, it is also *nonrobust*.

Remark. As was the case with the estimator $H_4(\hat{F}_n, k)$, the empirical nonparametric estimator of $\mathbf{E}(R_{\mathrm{ECO}})$ is also biased. Moreover, it can even produce negative estimates of the net premium. Therefore, in Section 5 we will use an unbiased nonparametric estimator based on sample order statistics $X_{n,n-j+1}$, $j = 1, \dots, k$.

4.7. Discussion

Clearly, the empirical nonparametric estimators of the net premiums of the above considered reinsurance treaties are nonrobust with respect to UBP and GES criteria. At first look, this conclusion does not appear to be surprising because the net premium is an expected value-type functional. However, similar conclusions not necessarily can be obtained for all functionals of this type. In fact, as derivations in Sections 4.2–4.6 suggest, that depends on the underlying function the expected value of which is computed. If the function itself is *unbounded* the functional will be *nonrobust*, and if it is *bounded* the functional (most likely) will be *robust*. Since reinsurance treaties (that are currently available in the actuarial literature) are unbounded functions, the net premium functionals corresponding to those treaties are found to be nonrobust. Also, we note here that this conclusion does bear a certain element of surprise. That is because it contradicts a quite popular belief that empirical nonparametric estimators must be robust simply because they are not based on assumptions, which may be violated. We emphasize again that robustness of such methods depends on a function (or a parameter) one tries to estimate.

Further, in a very recent paper by Marceau and Rioux (2001), similar conclusions were reached for the excess-of-loss and stop-loss premiums, and the probability of ruin. Their approach was based on the influence function and its empirical equivalent the sensitivity function. In order to obtain robust estimators, these authors proposed to use the minimum distance methods, in particular, the minimum Cramér–von Mises estimator. The robustness of this estimator was then illustrated for the case when individual claims are coming from an exponential distribution.

Finally, although the work by Marceau and Rioux and this paper have some overlap, these papers also very nicely complement each other by presenting different robustness tools and investigating complimentary risk measures. Thus, we hope that the ideas of statistical robustness presented and illustrated here and by Marceau and Rioux (2001) will continue to attract attention of other researchers as well as stimulate interest of practicing actuaries.

5. Example: robust estimators versus nonrobust

In this section we present a numerical example which illustrates what may happen in practice if one uses nonrobust estimators instead of robust ones. In order to do this, we investigate behavior of robust and nonrobust estimators of the net premiums for all treaties under contaminated data scenario. For severity of claims, we consider the contamination model of form

$$F_\varepsilon = (1-\varepsilon)\mathrm{Pa}(\sigma, \alpha) + \varepsilon\, \mathrm{Pa}(\sigma^*, \alpha^*),$$

where ε represents the level of contamination and $\mathrm{Pa}(\sigma, \alpha)$ denotes the Pareto distribution having CDF

$$F(x) = 1 - \left(\frac{\sigma}{x}\right)^\alpha, \quad x \geq \sigma, \tag{18}$$

defined for $\alpha > 0$ and $\sigma > 0$.

Besides insurance claims, the Pareto distribution (18) is suitable for many situations including models in telecommunications, medicine, finance, and in applications of extreme value theory. Robust estimators of parameter α (with σ known or unknown) are developed and investigated by Brazauskas and Serfling (2000a,b, 2001a,b). Let us briefly review robust estimators of α in the Pareto model, which will be relevant for further discussion.

5.1. Robust estimation of α in $Pa(\sigma, \alpha)$ with σ known

Consider a sample $X_1, \dots, X_n$ from the model $\mathrm{Pa}(\sigma, \alpha)$ as described by (18). Denote the ordered sample values by $X_{n,1} \le \cdots \le X_{n,n}$. For brevity, results are presented only in the case of σ known.

5.1.1. Trimmed mean estimators

For specified β_1 and β_2 satisfying $0 \le \beta_1, \beta_2 < 1/2$, a trimmed mean (TM) is formed by discarding the proportion β_1 lowermost observations and the proportion β_2 uppermost observations and averaging the remaining ones in some sense. In particular

$$\hat{\alpha}_{\mathrm{TM}} = \left(\sum_{i=1}^{n} c_{ni} (\log X_{n,i} - \log \sigma) \right)^{-1},$$

with

$$c_{ni} = \begin{cases} 0, & 1 \le i \le [n\beta_1], \\ \dfrac{1}{d}, & [n\beta_1] + 1 \le i \le n - [n\beta_2], \\ 0, & n - [n\beta_2] + 1 \le i \le n, \end{cases}$$

where $[\cdot]$ denotes "greatest integer part", and

$$d = d(\beta_1, \beta_2, n) = \sum_{j=[n\beta_1]+1}^{n-[n\beta_2]} \sum_{i=0}^{j-1} (n-i)^{-1}.$$

This choice of c_{ni}'s makes $\hat{\alpha}_{\mathrm{TM}}^{-1}$ mean unbiased.

It follows from the definition of c_{ni}'s that the TM estimator $\hat{\alpha}_{\mathrm{TM}}$ is completely unaffected by proportion β_2 of uppermost observations. This implies that $\hat{\alpha}_{\mathrm{TM}}$ has UBP $= \beta_2$ and therefore it is globally *robust*. For details see Brazauskas and Serfling (2000b).

The GES for these estimators is derived and evaluated by Brazauskas and Serfling (2001b) where it was found that the TM estimator $\hat{\alpha}_{\mathrm{TM}}$ is *B-robust* (also, see Kimber (1983a,b)). Table 1 provides numerical values for GES/α and UBP for the TM estimators.

5.1.2. Generalized median estimators

Generalized median (GM) statistics are defined by taking the median of the $\binom{n}{k}$ evaluations of a given kernel $h(x_1, \dots, x_k)$ over all k-sets of the data. Such an estimator was considered by Brazauskas and Serfling (2000b) for

Table 1
GES/α and UBP for selected β_2

Robustness measure	β_2			
	0.05	0.10	0.15	0.20
GES/α, $\beta_1 = \beta_2$	2.56	2.10	1.87	1.72
GES/α, $\beta_1 = 0$	2.56	2.09	1.85	1.69
UBP	0.05	0.10	0.15	0.20

Table 2
Values of C_k for k = 1–7

$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$
1.44	1.19	1.12	1.09	1.07	1.06	1.05

the parameter α in the case of σ known:

$$\hat{\alpha}_{\mathrm{GM}} = \mathrm{Median}\{h(X_{i_1}, \dots, X_{i_k}; \sigma)\},$$

with a particular choice of kernel h:

$$h(x_1, \dots, x_k; \sigma) = \frac{1}{C_k} \frac{1}{k^{-1} \sum_{j=1}^{k} \log x_j - \log \sigma},$$

where C_k is a multiplicative median-unbiasing factor, i.e., chosen so that in each case the distribution of $h(X_{i_1}, \dots, X_{i_k}; \sigma)$ has *median* α. Values of C_k are provided in Table 2. (For $k > 7$, C_k is given by a very accurate approximation, $C_k \approx k/(k - 1/3)$.)

A detailed study of robustness of the GM estimators, including the case of σ unknown and small sample performance, is available by Brazauskas and Serfling (2000a, 2001a,b). It was found that these estimators are globally and *B-robust*. The UBP of $\hat{\alpha}_{\mathrm{GM}}$ is given by the following formula: UBP $= 1 - 2^{-1/k}$. In Table 3, we provide numerical values for GES/α and UBP for the GM estimators.

5.2. *Underlying situation*

Suppose we have a data set of observed claims which is contaminated (due to unknown reasons) with only 4% of "bad" observations. In reality this 4% contamination model can occur as follows. There are 24 claims, $X_1, \dots, X_{24}$, distributed according to model (18) with parameters $\sigma = 1$ and $\alpha = 1.50$ and 1 observation, X_{25}, from $\mathrm{Pa}(\sigma^* = 4, \alpha^* = 1.05)$, which is not representative of the model being estimated. That is, the underlying probabilistic rule which generates claims is

$$F_{0.04} = (1 - 0.04)\mathrm{Pa}(\sigma = 1, \alpha = 1.50) + 0.04\,\mathrm{Pa}(\sigma^* = 4, \alpha^* = 1.05). \quad (19)$$

Clearly, it is impossible (for an actuary) to know that observation X_{25} comes from $\mathrm{Pa}(\sigma^* = 4, \alpha^* = 1.05)$ and not $\mathrm{Pa}(\sigma = 1, \alpha = 1.50)$, because X_{25} will be a number which is ≥ 4 (since $\sigma^* = 4$) and that is a quite *probable* realization for the $\mathrm{Pa}(\sigma = 1, \alpha = 1.50)$ model. Indeed, the probability of such an event

$$\mathbf{P}_{\mathrm{Pa}(1,1.50)}(X \geq 4) = (\tfrac{1}{4})^{1.50} = 0.125,$$

is relatively high. Thus, our hypothetical actuary has no other choice but to assume that *all* 25 observations came from a single model, in particular, the $\mathrm{Pa}(\sigma = 1, \alpha = 1.50)$ model.

For estimation of premiums $\mathbf{E}(R_{\mathrm{P}})$, $\mathbf{E}(R_{\mathrm{SL}})$, $\mathbf{E}(R_{\mathrm{EL}})$, $\mathbf{E}(R_{\mathrm{LC}})$, and $\mathbf{E}(R_{\mathrm{ECO}})$, if the true underlying model is indeed $\mathrm{Pa}(1, 1.50)$, the GM and TM estimators are *consistent*, *asymptotically normal*, and *asymptotically unbiased*

Table 3
GES/α and UBP for k = 1–4 and 7

Robustness measure	k				
	1	2	3	4	7
GES/α	1.44	1.90	2.27	2.60	3.38
UBP	0.50	0.29	0.21	0.16	0.09

estimators of the corresponding claim severity quantity. The nonparametric estimators that we consider here are all unbiased but nonrobust and also have infinite variances. Let us examine behavior of robust and nonparametric estimators when the underlying model *slightly* deviates from the $\mathrm{Pa}(\sigma = 1, \alpha = 1.50)$ model.

5.3. Behavior of estimators

First, we find the target values that we wish to estimate for each treaty. (Here all data points $X_1, \ldots, X_{25}$ are coming from the target model $\mathrm{Pa}(\sigma = 1, \alpha = 1.50)$.)

- *Proportional reinsurance*. The net proportional premium is given by $\mathbf{E}(R_{\mathrm{P}}) = p\mathbf{E}(N)\,\mathbf{E}(X)$. Thus, the target value is:

$$\mathbf{E}(X) = \frac{\sigma\alpha}{\alpha - 1} = 3.$$

- *Stop-loss reinsurance* ($K = 75$). The net stop-loss premium is given by $\mathbf{E}(R_{\mathrm{SL}}) = \sum_{n=1}^{\infty} \mathbf{P}\{N = n\}\mathbf{E}(S_n - K)^+$. Thus, the target quantity is $\mathbf{E}(S_n - K)^+ = \int (x - K)^+ \,\mathrm{d}F^{*n}(x)$. Typically, the distribution function F^{*n} is found using the fast Fourier transform (FFT) method which is described by Klugman et al. (1998, pp. 317–320). In our case, however, this approach is inapplicable since it requires an explicit formula of the characteristic function which is quite intractable for the Pareto distribution (see Arnold (1983, pp. 52–56)). Nevertheless, a reliable approximation can be found using methods of extreme value theory. In particular, for very large x, we have $1 - F^{*n}(x) \approx n(1 - F(x))$ (see Beirlant et al. (1996, p. 117)). Hence, the approximated target value is:

$$\mathbf{E}(S_n - K)^+ \approx n\mathbf{E}(X - K)^+ = n\frac{K^{-\alpha+1}}{\alpha - 1} = 5.77.$$

- *Excess-of-loss reinsurance* ($D = 4$). The net excess-of-loss premium is given by $\mathbf{E}(R_{\mathrm{EL}}) = \mathbf{E}(N)\,\mathbf{E}(X - D)^+$. Thus, the target value is:

$$\mathbf{E}(X - D)^+ = \frac{D^{-\alpha+1}}{\alpha - 1} = 1.$$

- *Largest claims reinsurance* ($k = 1$). The net largest claims premium is given by $\mathbf{E}(R_{\mathrm{LC}}) = \sum_{n=1}^{\infty} \mathbf{P}\{N = n\} \sum_{j=1}^{k} \mathbf{E}(X_{n,n-j+1})$. Thus, the target value is:

$$\sum_{j=1}^{k} \mathbf{E}(X_{n,n-j+1}) = \mathbf{E}(X_{n,n}) = \prod_{j=1}^{n} \left(\frac{j\alpha}{j\alpha - 1}\right) = 23.01.$$

- ECOMOR *reinsurance* ($k = 2$). The net ECOMOR premium is given by $\mathbf{E}(R_{\mathrm{ECO}}) = \sum_{n=1}^{\infty} \mathbf{P}\{N = n\} \sum_{j=1}^{k-1} (\mathbf{E}(X_{n,n-j+1}) - \mathbf{E}(X_{n,n-k+1}))$. Thus, the target value is:

$$\sum_{j=1}^{k-1} \mathbf{E}(X_{n,n-j+1} - X_{n,n-k+1}) = \mathbf{E}(X_{n,n} - X_{n,n-1}) = \frac{1}{\alpha} \prod_{j=1}^{n} \left(\frac{j\alpha}{j\alpha - 1}\right) = 15.34.$$

Further, robust estimators for the above presented quantities are found by replacing α with the values of TM and GM estimates $\hat{\alpha}$. Corresponding nonparametric estimators are: $\bar{X}$ (proportional), $n^{-n} \sum_{1 \le i_1, \ldots, i_n \le n} (X_{i_1} + \cdots + X_{i_n} - 75)^+$ (stop-loss), $\overline{(X - 4)^+}$ (excess-of-loss), $X_{n,n}$ (largest claims), and $X_{n,n} - X_{n,n-1}$ (ECOMOR). Behavior of nonparametric estimators when the underlying model is (19) instead of $\mathrm{Pa}(1, 1.5)$, is investigated using simulations, and behavior of robust estimators under contamination is evaluated via simulations as well as by the theoretical approximation (2). These investigations are summarized in Table 4.

Table 4
Behavior of robust and nonparametric estimators under contamination

Estimator	Approximated values[a]					Simulated values[b]														
	P	SL	EL	LC	ECO	Run no. 1					Run no. 2					Run no. 3				
						P	SL	EL	LC	ECO	P	SL	EL	LC	ECO	P	SL	EL	LC	ECO
TM, $\beta_1 = 0, \beta_2 = 0.20$	3.51	11.22	1.44	31.67	22.64	3.51	11.18	1.44	31.61	22.60	3.52	11.30	1.45	31.78	22.74	3.51	11.29	1.45	31.77	22.73
TM, $\beta_1 = 0, \beta_2 = 0.15$	3.57	11.98	1.50	32.77	23.60	3.58	12.13	1.51	32.99	23.78	3.59	12.29	1.52	33.21	23.97	3.60	12.30	1.52	33.23	23.99
TM, $\beta_1 = 0, \beta_2 = 0.10$	3.67	13.24	1.59	34.56	25.14	3.62	12.66	1.55	33.73	24.42	3.64	12.86	1.56	34.02	24.67	3.63	12.79	1.56	33.92	24.59
TM, $\beta_1 = 0, \beta_2 = 0.05$	3.89	16.17	1.79	38.56	28.64	3.67	13.20	1.59	34.49	25.08	3.68	13.44	1.60	34.83	25.38	3.68	13.35	1.60	34.71	25.28
GM, $k = 1$	3.42	10.14	1.36	30.06	21.27	3.40	9.97	1.35	29.81	21.05	3.41	10.08	1.36	29.98	21.20	3.41	10.04	1.36	29.92	21.15
GM, $k = 2$	3.59	12.23	1.52	33.13	23.91	3.58	12.11	1.51	32.95	23.75	3.59	12.24	1.52	33.14	23.92	3.59	12.20	1.51	33.08	23.86
GM, $k = 3$	3.75	14.29	1.66	36.01	26.40	3.65	12.98	1.57	34.19	24.82	3.66	13.14	1.58	34.42	25.02	3.66	13.10	1.58	34.35	24.96
GM, $k = 4$	3.91	16.46	1.80	38.94	28.97	3.67	13.22	1.59	34.53	25.12	3.69	13.49	1.61	34.90	25.44	3.68	13.42	1.60	34.81	25.36
GM, $k = 7$	4.36	23.31	2.23	47.72	36.78	3.68	13.43	1.60	34.82	25.37	3.70	13.71	1.62	35.20	25.70	3.70	13.62	1.61	35.08	25.60
NONPARAMETRIC	NA	NA	NA	NA	NA	3.96	32.93	1.92	41.30	31.65	4.84	57.96	2.73	61.60	50.74	4.06	40.68	1.93	42.15	31.58
Target values	3.00	5.77	1.00	23.01	15.34	3.00	5.77	1.00	23.01	15.34	3.00	5.77	1.00	23.01	15.34	3.00	5.77	1.00	23.01	15.34

[a] Theoretical approximations are based on formula (2).
[b] Simulations are based on 10,000 samples of size $n = 25$ from model (19).

Three runs of simulations clearly demonstrate that behavior of nonparametric estimators with (theoretically) infinite variance is completely unpredictable (Table 4). Sometimes they may converge to values that are relatively close to each other (e.g., Run nos. 1 and 3) while in the case of Run no. 2 they approach absolutely different values. Moreover, even though all of them are unbiased estimators when data follow Pa(1, 1.5), due to the nonrobustness these nonparametric estimators approach values that are way off target when data follow a 4% contaminated model.

As expected, the TM and GM estimators are stable for all three simulation runs and just moderately deviate from the target values. Also, note that simulated and approximated values are very close for estimators with high UBP and diverge for the ones with low UBP. A better performance (i.e., values closer to the target value) of simulated values for the low UBP estimators means that Pa(4, 1.05) is not the worst contaminating model in this case.

6. Final remarks

The example of Section 5 demonstrates that while robust estimators of the net premiums *do react* to contaminating observations, corresponding nonparametric estimators tend to *over-react*. Moreover, even when the data set is "clean", the nonparametric estimators may be biased or have infinite variance which leads to inconsistency. Therefore, for estimation of reinsurance premiums, the empirical nonparametric approach should be used with *extreme caution*.

A more reliable alternative is the robust parametric approach which we propose to use for fitting the distribution of claims severity which is then applied to the estimation of net premiums. This approach, however, has to be developed on a case-by-case basis. Since the one-parameter Pareto model is a quite restrictive model (although most parsimonious and of paramount importance for insurance applications), the use of robust estimators for more flexible distributions is of interest. We conclude with a short list of papers that already addressed the problem of robust estimation for other models which are appropriate for reinsurance portfolios. For robust estimation in gamma samples see Kimber (1983a,b); robust estimation of the mean, variance, and limited expected value of the lognormal distribution is treated by Serfling (2002) and Marazzi and Ruffieux (1999) developed robust estimators for the mean of the lognormal, gamma, and Weibull distributions; and Davison and Smith (1990), Peng and Welsh (2001) proposed robust estimators for parameters of the generalized Pareto distribution.

Acknowledgements

The author is very appreciative of valuable comments provided by Professor R.J. Serfling and constructive remarks by an anonymous referee. These have led to many improvements in the paper. Also, support by a grant from the Actuarial Education and Research Fund is gratefully acknowledged.

References

Arnold, B.C., 1983. Pareto Distributions. International Cooperative Publishing House, Fairyland, MD.

Beirlant, J., Teugels, J.L., Vynckier, P., 1996. Practical Analysis of Extreme Values. Leuven University Press, Leuven, Belgium.

Brazauskas, V., 2000. Robust parametric modeling of the proportional reinsurance premium when claims are approximately Pareto-distributed. In: Proceedings of the American Statistical Association, Business and Economic Statistics Section, 2000, pp. 144–149.

Brazauskas, V., Serfling, R., 2000a. Robust estimation of tail parameters for two-parameter Pareto and exponential models via generalized quantile statistics. Extremes 3 (3), 231–249.

Brazauskas, V., Serfling, R., 2000b. Robust and efficient estimation of the tail index of a single-parameter Pareto distribution. North American Actuarial Journal 4 (4), 12–27.

Brazauskas, V., Serfling, R., 2001a. Small sample performance of robust estimators of tail parameters for Pareto and exponential models. Journal of Statistical Computation and Simulation 70, 1–19.

Brazauskas, V., Serfling, R., 2001b. Reply to discussion: robust and efficient estimation of the tail index of a single-parameter Pareto distribution. North American Actuarial Journal 5 (3), 126–128.

Davison, A.C., Smith, R.L., 1990. Models for exceedances over high thresholds (with discussion). Journal of the Royal Statistical Society Series B 52, 393–442.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

Gerber, H.U., 1979. An Introduction to Mathematical Risk Theory. S.S. Huebner Foundation for Insurance Education, University of Pennsylvania.

Hampel, F.R., 1968. Contributions to the theory of robust estimation. Ph.D. Dissertation. University of California, Berkeley.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics: The Approach Based on Influence Functions. Wiley, New York.

Hipp, C., 1996. The delta-method for actuarial statistics. Scandinavian Actuarial Journal 1, 79–94.

Huber, P.J., 1964. Robust estimation of a location parameter. Annals of Mathematical Statistics 35, 73–101.

Huber, P.J., 1981. Robust Statistics. Wiley, New York.

Kimber, A.C., 1983a. Trimming in gamma samples. Applied Statistics 32, 7–14.

Kimber, A.C., 1983b. Comparison of some robust estimators of scale in gamma samples with known shape. Journal of Statistical Computation and Simulation 18, 273–286.

Klugman, S.A., Panjer, H.H., Willmot, G.E., 1998. Loss Models: From Data to Decisions. Wiley, New York.

Marazzi, A., Ruffieux, C., 1999. The truncated mean of an asymmetric distribution. Computational Statistics and Data Analysis 32, 79–100.

Marceau, E., Rioux, J., 2001. On robustness in risk theory. Insurance: Mathematics and Economics 29, 167–185.

Peng, L., Welsh, A.H., 2001. Robust estimation of the generalized Pareto distribution. Extremes 4 (1), 53–65.

Reiss, R.-D., Thomas, M., 1997. Statistical Analysis of Extreme Values from Insurance, Finance, Hydrology and Other Fields. Birkhauser, Basel.

Serfling, R., 2002. Efficient and robust fitting of lognormal distributions. North American Actuarial Journal 6 (4), 95–109.