An extreme value approximation to the integrated tail distribution

Ants Kaasik

University of Tartu, Estonia, email: ants.kaasik@ut.ee

Abstract

It is well known that the infinite time ruin probability can be expressed as a probability of a random sum of independent summands exceeding the initial capital when the Cramér-Lundberg model is assumed (see e.g. [1, p. 399]). Suppose the claims have distribution function $F(x)$ such that the expected value of a claim, $\mu$, is finite. Then the independent summands have the integrated tail distribution with distribution function $F^I(x)$ for which

$$\bar{F}^I(x) = 1 - F^I(x) = \frac{\bar{F}(y)dy}{\mu}$$

holds.

In articles, typically, it is assumed that $F^I(x)$ is known, however this is rarely the case in practice. The purpose of the research is to study the approximation of the integrated tail distribution which is based solely on the claim data. This kind of problem was tackled in [3], where $F(x)$ was replaced by its empirical counterpart and $\mu$ with the sample mean (an empirical cumulative distribution function (ecdf) approach). While justified theoretically, the ecdf approach is handicapped by its finite support – the right endpoint of which is determined by the largest member of the sample (but the tail region is the crucial part when estimating the ruin probability). To overcome this a different route needs to be studied.

Suppose that the claim distribution is in a domain of attraction of an extreme value distribution $H_{\xi}$ (in our case $\xi \in [0, 1]$). Define

$$F_u(x) = F(u + x) - F(u)\bar{F}(u), \quad u > 0, x > 0$$

the conditional distribution function. A well known result (see e.g. [2, pp. 165–166]) states that the generalized Pareto distribution (GPD) with distribution function $G(x)$ is a good approximation of $F_u$ for large values of $u$ because

$$\lim_{u \to \infty} \sup_{x > 0} |F_u(x) - G_{\xi, \sigma(u)}(x)| = 0.$$  \hspace{1cm} (3)

We show that when mild assumptions on the convergence (3) are satisfied the GPD is usable for integrals. More precisely,

$$\sup_{x > 0} \left| \int_{x}^{\infty} \bar{F}_u(y)dy - \int_{x}^{\infty} \bar{G}_{\xi, \sigma(u)}(y)dy \right| = o(a(u)),$$  \hspace{1cm} (4)

1
where
\[ a(u) = \frac{\int_u^\infty \bar{F}(x)dx}{\bar{F}(u)} \] (5)

is the mean residual life function of the claim distribution and that when
\( \bar{F}_u(x) = G_{\xi,\sigma}(x) \) for all \( x > 0 \), the tail of the integrated tail distribution also
has a GPD form.

When performing simulations, in the role of \( F(x) \) members of the subexponential distributions were used. Results show that the GPD approach outperforms the ecdf approach when the supremum relative error is considered regardless of the fact that the parameters of the GPD were estimated disregarding the properties of a specific sample. It is also noted that when the sample size is insufficient, good approximation of \( F^I(x) \) is not possible when no additional information about \( F(x) \) is available.

**Keywords**
Cramér-Lundberg model, generalized Pareto distribution, integrated tail distribution, M/G/1 queue, subexponential distribution.

**References**

