On variational bounds in the compound Poisson approximation of the individual risk model

Bero Roos
On variational bounds in the compound Poisson approximation of the individual risk model

Bero Roos*
University of Hamburg

June 14, 2006
Revised version

Abstract: We present new upper bounds for the total variation distance between the aggregate claims distribution in the individual risk model and a suitable compound Poisson distribution. It turns out that the bounds are generally valid and contain so-called magic factors. Higher-order approximations, including the signed Kornya–Presman measures, are also investigated. In contrast to results of a previous paper by the author, the results do not depend on a joint decomposition of the individual claim amount distributions. Further, we do not need to assume the finiteness of moments.

Keywords: compound Poisson approximation; individual risk model; signed Kornya–Presman measures; magic factors; total variation distance.

JEL Classification: C63.

Subject and Insurance Branch Code: IM20.


*Postal address: Department of Mathematics, SPST, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany. E-mail: roos@math.uni-hamburg.de. Tel.: +49 (0) 40-42838 4925, fax: +49 (0) 40-42838 6955.
1 Introduction

In the classical individual risk model, the aggregate claims distribution is one of the main objectives. Usually, the aggregate claims is understood as the sum of all claims occurring over a certain time period. But it may happen that more information about the claims is needed. For such purposes, it may be convenient to extend the individual model to a multivariate setting of dimension $\ell \in \mathbb{N} = \{1, 2, 3, \ldots \}$, say. Here, we consider a portfolio with $n \in \mathbb{N}$ policies, producing the $\ell$-dimensional individual claims $X_i = (X_{i,1}, \ldots, X_{i,\ell})$ for $i \in \{1, \ldots, n\}$, which are modeled as independent but not necessarily identically distributed random vectors in $\mathbb{R}^\ell$. For $i \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, \ell\}$, the $X_{i,k}$ is the claim of class $k$, corresponding to the $i$th contract. The sum $Y_i = \sum_{k=1}^\ell X_{i,k}$ represents the total claim for the $i$th contract. In risk $i$, a claim occurs with the probability

$$p_i := P(X_i \neq 0)$$

and has the distribution

$$Q_i := P(X_i \in \cdot \mid X_i \neq 0).$$

Without loss of generality, we assume that $p_i > 0$ for all $i$. The aggregate claims vector is given by

$$S_n = \sum_{i=1}^n X_i.$$  

Note that, generally, risks are non-negative random variables, so that one may wonder, why we allow the $X_{i,k}$ to be negative. The results below, however, hold when the $X_{i,k}$ are arbitrary real valued, so that non-negativity would impose an artificial restriction.

Clearly, for $\ell = 1$, we reobtain the classical univariate individual model. Perhaps the simplest non-trivial higher dimensional example is the one, where, for all $i$, $X_i$ is a random vector with at most one non-zero entry, which, in turn, must then be equal to $Y_i$. From the view of the univariate model, this means that, here, each non-zero claim $Y_i$ is assigned to exactly one of the $\ell$ classes.

We may assume, that, for all $i$, the $p_i$ is small. Otherwise, the insurance company would not have accepted this contract. It turns out that, under this assumption, the approximation of the aggregate claims distribution $\mathcal{L}(S_n)$ by a compound Poisson one
On variational bounds

$\text{CPo}(\lambda, Q)$ is good in some sense. This distribution can simply be defined by

$$\text{CPo}(\lambda, Q) = \sum_{m=0}^{\infty} e^{-\lambda} \frac{\lambda^m}{m!} Q^* m,$$

where

$$\lambda = \sum_{i=1}^{n} p_i, \quad Q = \frac{1}{\lambda} \sum_{i=1}^{n} p_i Q_i,$$

$Q^* m$, $(m \in \mathbb{N})$ denotes the $m$-fold convolution of $Q$ with itself, and $Q^* 0 = I_0$ is the Dirac measure at point $0 \in \mathbb{R}^\ell$.

In this paper, we are concerned with upper bounds for the approximation error of $\mathcal{L}(S_n)$ by $\text{CPo}(\lambda, Q)$. It turns out that, due to an additional (magic) factor, our bounds are smaller than previous ones at least in the case when the $Q_i$ are different but $Q_1 \approx \ldots \approx Q_n$ in some sense. However, it may happen that the magic factor is compensated by an additional term which measures how well the $Q_i$ coincide.

As a measure of accuracy, we consider the total variation distance, which is defined by

$$d_{\text{TV}}(R_1, R_2) = \sup_{B \in \mathcal{B}^\ell} |R_1(B) - R_2(B)|,$$

where $\mathcal{B}^\ell$ denotes the Borel $\sigma$-algebra over $\mathbb{R}^\ell$ and $R_1$ and $R_2$ are two finite signed measures on $((\mathbb{R}^\ell, \mathcal{B}^\ell))$. For results concerning other distances, such as the Kolmogorov or the stop-loss metrics, see, for example, Zaïtsev (1983), Gerber (1984), Hipp (1985, 1986), de Pril and Dhaene (1992), Kuon et al. (1993), Čekanavičius (1997), Dhaene and Sundt (1997), and Roos (2005). For a functional approach to approximations of the individual risk model, see Pitts (2004).

2 Facts on compound Poisson approximation

2.1 Basic inequalities and the magic factor

One of the most popular results in compound Poisson approximation is essentially due to Khintchine (1933) and Doeblin (1939) (see also Le Cam, (1960, page 1183)). The result is also contained in Gerber (1984, Theorem 1(a)). It says that

$$d_{\tau} := d_{\text{TV}}(\mathcal{L}(S_n), \text{CPo}(\lambda, Q)) \leq \sum_{i=1}^{n} p_i^2 =: \lambda_2. \quad (1)$$
Note that (1) was initially shown only for the univariate case $\ell = 1$. But the proof for general $\ell \in \mathbb{N}$ is done in the same way. From an observation by Le Cam (1965, page 188) it can be deduced that, for each $n \in \mathbb{N}$, there exist $Q_1, \ldots, Q_n$ such that, for each choice of $p_1, \ldots, p_n$, we have
\[
C_1 \min\{\lambda_2, 1\} \leq d_\tau \leq C_2 \min\{\lambda_2, 1\}.
\]
Here, $C_1$ and $C_2$ denote positive absolute constants. It should be mentioned that, in Zaïtsev (1989, Remark 1.1), Le Cam’s argument has been made more precise under the assumption that
\[
p_0 := \max_{i \in \{1, \ldots, n\}} p_i \leq \frac{1}{2}.
\]
However, as is easily seen, this assumption can be dropped. From this we see that, in general, there is no hope of finding an upper bound independent of the $Q_i$, which is of a better order than $\lambda_2$. But a further result by Le Cam (1960, Theorem 2) tells us that, under the special assumption that $\ell = 1$, $Q_1 = \cdots = Q_n = I_1$ is the Dirac measure at point one and that $p_0 \leq 1/4$, we have
\[
d_\tau \leq 8 \frac{\lambda_2}{\lambda},
\]
which is better than (1), if $\lambda > 8$. From that time on, many papers appeared on Poisson approximation. One of the most important results is due to Barbour and Hall (1984, Theorems 1 and 2), who, by using Stein’s method, showed that, if $\ell = 1$ and $Q_1 = \cdots = Q_n = I_1$, then
\[
\frac{\lambda_2}{32} \min\{\lambda^{-1}, 1\} \leq d_\tau \leq \lambda_2 \min\{\lambda^{-1}, 1\}.
\] (2)
It is easily verified that $\lambda_2 \min\{\lambda^{-1}, 1\} \geq (\lambda_2/\lambda)^2$, which, together with (2), implies that, under the present conditions, $d_\tau$ is small if and only if $\lambda_2/\lambda$ is small. From this, we see that the upper bound $\lambda_2/\lambda$ in (2) is much more important than the $\lambda_2$. In the literature (see, for example, Barbour et. al., 1992, Introduction), the additional factor $\lambda^{-1}$ is sometimes called a magic factor, since, on the one hand, it is highly desirable, but on the other hand, the proof of its existence turns out to be difficult.

A simple observation made by Le Cam (1965, page 187) and later rediscovered by Michel (1987, page 167) implies that the upper bound in (2) remains valid in the case $\ell \in \mathbb{N}$ and $Q_1 = \cdots = Q_n$. Indeed, more generally, the total variation distance in the case $\ell \in \mathbb{N}$ and $Q_1 = \cdots = Q_n$ is bounded from above by the distance in the case $\ell = 1$ and $Q_1 = \cdots = Q_n = I_1$. Therefore, concerning upper bounds, the preliminary restriction to $\ell = 1$ above was unnecessary.
2.2 Facts under a more general assumption

As explained above, in order to obtain upper bounds for $d_\tau$ of a better order than $\lambda_2$, we have to make suitable assumptions on the $Q_1, \ldots, Q_n$. In Roos (2003), some results are given when the $Q_1, \ldots, Q_n$ can be jointly decomposed in the following form: for all $i \in \{1, \ldots, n\}$,

$$Q_i = \sum_{r=1}^\infty q_{i,r} U_r,$$

for suitable $q_{i,r} \in [0, 1]$ with $\sum_{r=1}^\infty q_{i,r} = 1$ and a sequence of probability measures $U_1, U_2, U_3, \ldots$ on $(\mathbb{R}^\ell, \mathcal{B}^\ell)$, which are not allowed to depend on $i$. Note that it is easily shown that this assumption can be always fulfilled, that is, for given $Q_1, \ldots, Q_n$, there exist $q_{i,r}$’s and $U_r$’s such that (3) is valid. However, a trivial decomposition, based on $q_{i,r} \in \{0, 1\}$ for all $i$ and $r$, should be avoided, since generally, in this case, the order of the respective bounds will not be better than $\lambda_2$. Now, (13) in Roos (2003) states that

$$d_\tau \leq 8.8 \beta,$$

where

$$\beta = \sum_{i=1}^n p_i^2 \min \left\{ \frac{\nu_i}{\lambda}, 1 \right\},$$

$$\nu_i = \sum_{r=1}^\infty \frac{q_{i,r}^2}{q_r}, \quad (i \in \{1, \ldots, n\}), \quad q_r = \frac{1}{\lambda} \sum_{i=1}^n p_i q_{i,r}, \quad (r \in \mathbb{N}).$$

Here, for $r \in \mathbb{N}$, we set $q_{i,r}/q_r = 0$ whenever $q_r = 0$. It is easily verified that, for all $i$, $\nu_i$ is finite. If $\nu_i/\lambda$ is less than one, then looking at (4), the magic factor $\lambda^{-1}$ is in use. However, from Cauchy’s inequality, it follows that, for all $i$, $\nu_i \geq 1$, so that $\nu_i$ itself cannot become small. Note that, in (12) of Roos (2003), it was shown that, if $\alpha(2^{-3/2}) < (2e)^{-1}$, then

$$d_\tau \leq \frac{\alpha(2^{-3/2})}{1 - 2 e \alpha(2^{-3/2})},$$

where, for $x \in [0, \infty)$,

$$\alpha(x) = \sum_{i=1}^n g_1(2p_i) p_i^2 \min \left\{ \frac{x\nu_i}{\lambda}, 1 \right\}, \quad g_1(x) = 2 \frac{e^x}{x^2} (e^{-x} - 1 + x).$$

In practice, due to the constants, (7) is often much better than (4). On the other hand, for discussion of the order, (4) is better suited, because of the absence of a singularity.
In Roos (2003), a further more restrictive decomposition of the individual claim amount distributions was used. However, for the respective results, some moments have to be finite. The main argument of the mentioned paper is a slight modification of an expansion due to Kerstan (1964). It is not clear whether these results can also be proved by using Stein’s method; see, for example, Barbour (2005).

3 Results

3.1 First-order results

Often in applications, the $Q_i$ are absolutely continuous. Here, it may be a problem to derive a non-trivial decomposition (3) of the $Q_1, \ldots, Q_n$. In Theorem 1 below, we present similar bounds as in (4) and (7) but without the assumption of a decomposition like (3). Below, Proposition 1 shows, that, to some extent, it is better to use one of the bounds of Theorem 1 than (4) or (7).

Further notation is needed. Since, for all $i \in \{1, \ldots, n\}$, $Q_i$ is absolutely continuous with respect to $Q$, that is, $Q_i(B) = 0$ for every set $B \in \mathcal{B}$ with $Q(B) = 0$, from the Radon–Nikodym theorem, it follows that $Q_i$ has a $Q$-density

$$f_i : \mathbb{R}^\ell \rightarrow [0, \infty).$$

In other words, $f_i$ is measurable and, for each $B \in \mathcal{B}$, we have $Q_i(B) = \int_B f_i \, dQ$.

**Theorem 1** Generally, we have

$$d_\tau \leq 8.8 \tilde{\beta},$$

$$d_\tau \leq \frac{\tilde{\alpha}(2^{-3/2})}{1 - 2e\tilde{\alpha}(2^{-3/2})},$$

where

$$\tilde{\beta} = \sum_{i=1}^n p_i^2 \min \left\{ \frac{1}{\lambda} \int f_i^2 \, dQ, 1 \right\},$$

$$\tilde{\alpha}(x) = \sum_{i=1}^n g_1(2p_i) p_i^2 \min \left\{ \frac{x}{\lambda} \int f_i^2 \, dQ, 1 \right\}, \quad (x \in [0, \infty)),$$

and, for (11), we assume that $\tilde{\alpha}(2^{-3/2}) < (2e)^{-1}$. Here, $g_1$ and $f_i$ are defined as in (8) and (9), respectively.
Note that, if \( Q_1 = \cdots = Q_n \), then, for all \( i \in \{1, \ldots, n\} \),
\[
\int f_i^2 \, dQ = 1,
\]
so that \( \tilde{\beta} = \lambda_2 \min\{\lambda^{-1}, 1\} \) and, similarly as in (2), we obtain the magic factor \( \lambda^{-1} \).

If, in contrast to the above assumption, \( Q_1 \approx \cdots \approx Q_n \) in some sense, then we expect that, for all \( i \), \( \int f_i^2 \, dQ \approx 1 \), which again gives a magic factor. Often the integrals \( \int f_i^2 \, dQ \) can be evaluated as follows. Suppose that, for \( i \in \{1, \ldots, n\} \), \( Q_i \) has a density \( h_i \) with respect to a \( \sigma \)-finite measure \( \mu \) on \((\mathbb{R}^\ell, \mathbb{B}^\ell)\). Then \( Q \) has the \( \mu \)-density
\[
h := \frac{1}{\lambda} \sum_{i=1}^{n} p_i h_i
\]
and it easily follows that, for all \( i \in \{1, \ldots, n\} \),
\[
\int f_i^2 \, dQ = \int_{\{h > 0\}} \frac{h^2}{h} \, d\mu.
\]
From this, we see once more that, if the \( Q_1, \ldots, Q_n \) and, in turn, the densities \( h_1, \ldots, h_n \) are approximately equal, then \( \int f_i^2 \, dQ \approx 1 \) for all \( i \).

One may ask whether, for given \( Q_1, \ldots, Q_n \), a decomposition (3) exists such that (4) is a smaller bound than (10). The following proposition shows that the answer is no, if we concentrate on the magic factor, i.e. if, in (5) and (12), we consider the first entry in the min-term. In this respect, using (4) or (7), there is no hope of obtaining much better bounds than the ones of Theorem 1.

**Proposition 1** Let (3) be valid. For \( i \in \{1, \ldots, n\} \), let \( \nu_i \) and \( f_i \) be defined as in (6) and (9), respectively. Then
\[
\sum_{i=1}^{n} p_i^2 \int f_i^2 \, dQ \leq \sum_{i=1}^{n} p_i^2 \nu_i.
\]

### 3.2 Second-order result

We now carry over the above idea to the second order result in Roos (2003, Theorem 3), which says that, in comparison with CPo\((\lambda, Q)\), the finite signed measure
\[
\text{CPo}_2(\lambda, Q) = \left( I_0 - \frac{1}{2} \sum_{i=1}^{n} p_i^2 (Q_i - I_0)^2 \right) * \text{CPo}(\lambda, Q),
\]
may be a better approximation of \( \mathcal{L}(S_n) \). Note that \( \text{CPo}_2(\lambda, Q) \) can be derived from an expansion of \( \mathcal{L}(S_n) \) due to Kerstan (1964).
Theorem 2 If \( \tilde{\alpha}(1) < 2^{1/2}e^{-1} \), then we have

\[
d_{TV}(L(S_n), \text{CPo}_2(\lambda, Q)) \leq \frac{4}{3} \tilde{\gamma} + (\tilde{\alpha}(1))^2 \left(1 + \frac{0.82 \tilde{\alpha}(1)}{1 - 2^{-1/2}e \tilde{\alpha}(1)}\right),
\]

where

\[
\tilde{\gamma} = \sum_{i=1}^{n} g_2(2p_i)p_i^3 \min\left\{0.46 \left(\frac{1}{\lambda} \int f_i^2 \text{d}Q\right)^{3/2}, 1\right\},
\]

\[
g_2(x) = \frac{3(g_1(x) - 1)}{2x}, \quad (x \in [0, \infty)),
\]

and \( g_1, f_i, \) and \( \tilde{\alpha}(1) \) are defined as in (8), (9), and (13), respectively.

Observe that, by continuity, \( g_2(x) \) is equal to one at \( x = 0 \) and increases to 2.3958... at \( x = 2 \). Therefore, if constants do not play a great rôle, in (15), the \( g_2(2p_i) \) can be replaced by 2.396. Further, note that, in the present context, Čekanavičius (1998, proof of Corollary 3.1) has shown that,

\[
d_{TV}(L(S_n), \text{CPo}_2(\lambda, Q)) \leq 2\lambda_2^2 + \frac{8}{3} \sum_{i=1}^{n} p_i^3,
\]

which, in contrast to (14), does not contain a magic factor.

### 3.3 Approximations by signed Kornya–Presman measures

In what follows, we present a consequence of Theorem 1 in Roos (2002) concerning the approximation by signed Kornya–Presman measures, which are defined by

\[
\text{KP}(s) = \exp \left(\sum_{i=1}^{n} \sum_{k=1}^{s} \frac{(-1)^{k+1}}{k} p_i^k (Q_i - I_0)^*^k\right),
\]

where \( s \in \mathbb{N} \) is fixed. It seems that such approximations were first considered by Kornya (1983) and Presman (1983), as a result of which we speak of Kornya–Presman signed measures. It should be mentioned, however, that the signed measures used by Kornya and Presman are slightly different (see also Hipp, 1986). Let

\[
A_1 = 1 - A_2 \in [0, 1],
\]

be arbitrary and and \( f_i \) be defined as in (9). Recall that \( p_0 = \max_{i \in \{1, \ldots, n\}} p_i \). For \( x \in [0, \infty) \), set

\[
\tilde{\beta}_s(x) = \sum_{i=1}^{n} p_i^{s+1} \min\left\{\left(\frac{x}{\lambda} \int f_i^2 \text{d}Q\right)^{(s+1)/2}, 1\right\},
\]

\[
V_s(x) = \left(1 - (1 - x) \exp \left(\sum_{m=1}^{s} \frac{x^m}{m}\right)\right) \frac{s + 1}{x^{s+1}}.
\]
Let \([x]\) denote the smallest integer greater or equal to \(x\). Observe that \(\beta_1(1)\) coincides with \(\beta\) from (12).

**Theorem 3** Let

\[
\begin{align*}
c_1(s) &= \begin{cases} 
(s + 1) \cdot 2^{-5/2}, & \text{for odd } s, \\
(s + 1) \cdot 2^{1/(2(s+1))-5/2}, & \text{for even } s,
\end{cases} \\
c_2(s, p_0) &= \frac{e^{2s} \left[s/2 - 1\right]!}{\sqrt{2\pi} (s + 1)} V_s(2p_0), \\
c_3(s, p_0) &= \frac{e^{2s+1}}{s + 1} V_s(2p_0), \\
c_4(s, p_0) &= 4e \sum_{m=2}^{s} \frac{(2p_0)^{m-2}}{m}.
\end{align*}
\]

If \(c_3(s, p_0) p_0^{s-1} \beta_1(2^{-3/2}A_2^{-1}) < 1\) and \(c_4(s, p_0) \beta_1(2^{-3/2}A_1^{-1}) < 1\), then

\[
d_{TV}(\mathcal{L}(S_n), KP(s)) \leq \eta \beta_s(c_1(s)A_2^{-1}), \tag{17}
\]

where

\[
\eta = \frac{c_2(s, p_0)}{(1 - c_3(s, p_0) p_0^{s-1} \beta_1(2^{-3/2}A_2^{-1}))/|s/2| (1 - c_4(s, p_0) \beta_1(2^{-3/2}A_1^{-1}))}.
\]

It should be mentioned that the left-hand side of (17) is independent of \(A_1\). Therefore, in applications, one can minimize the upper bound over all possible \(A_1 \in [0, 1]\). Further note that Hipp (1986, formula (6)) has shown that, if \(p_0 < 1/2\), then

\[
d_{TV}(\mathcal{L}(S_n), KP(s)) \leq \exp \left( \sum_{i=1}^{n} \frac{(2p_i)^{s+1}}{(s + 1)(1 - 2p_i)} \right) - 1. \tag{18}
\]

Due to the magic factor \(\lambda^{-(s+1)/2}\), the bound in (17) can be much more precise than the one in (18). Indeed, one of the reasons is that, if \(\lambda \to \infty\) and if \(\max_{i \in \{1, \ldots, n\}} \int f_i^2 \, dQ\) is bounded by an absolute constant, then \(\beta_s(1) \to 0\). However, an error bound derived from (18), which is too large for a given order of approximation can easily be reduced by increasing the order of approximation, which is usually possible with a small increase of computation time.

### 3.4 Comparison of the results

Let us give a comparison of the order of the bounds in Theorems 1–3. For the sake of simplicity, we consider the univariate case \(\ell = 1\) and assume that \(\max_{i \in \{1, \ldots, n\}} \int f_i^2 \, dQ\)
is bounded by some absolute constant. In order to get rid of the singularities in the upper bounds in (14) and (17), we assume that $\lambda_2 \min\{\lambda^{-1}, 1\}$ is bounded by some suitable small absolute constant.

Table 1 collects the main terms of the bounds without consideration of constants. The bounds with (resp. without) magic factors $\lambda^{-\kappa}$ for $\kappa > 0$ are derived by taking the first (resp. second) entry in the min-terms of the results. The terms in the last line of Table 1 coincide with the order of the bounds in (1), (16), and (18). For (18), we have to assume that $p_0$ is bounded by some absolute constant $c < 1/2$. As is easily shown, we have $\sum_{i=1}^n p_i^3 \leq \lambda_2^{3/2}$. Therefore, (14) yields a bound of a better order than (10). Similarly, if $s \geq 2$, (17) is better than (14). Further, we have $\sum_{i=1}^n p_i^{s+1} \leq \lambda$, which implies that, if $s \geq 2$, the bound in (17), unlike the other ones, is small when $\lambda$ is large.

Table 1: Comparison of the order of the bounds in Theorems 1–3

<table>
<thead>
<tr>
<th>number of formula</th>
<th>(10)</th>
<th>(14)</th>
<th>(17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>order of the upper bound with magic factor</td>
<td>$\frac{\lambda_2}{\lambda}$</td>
<td>$\frac{1}{\lambda^{3/2}} \sum_{i=1}^n p_i^3 + \left(\frac{\lambda_2}{\lambda}\right)^2$</td>
<td>$\frac{1}{\lambda^{(s+1)/2}} \sum_{i=1}^n p_i^{s+1}$</td>
</tr>
<tr>
<td>order of the upper bound without magic factor</td>
<td>$\lambda_2$</td>
<td>$\sum_{i=1}^n p_i^3 + \lambda_2^2$</td>
<td>$\sum_{i=1}^n p_i^{s+1}$</td>
</tr>
</tbody>
</table>

3.5 A numerical example

In what follows, we consider the univariate case $\ell = 1$ and assume that we have $n = 93$ contracts with

$$p_i = \begin{cases} 
0.03, & \text{if } i = 1, \ldots, 24, \\
0.04, & \text{if } i = 25, \ldots, 42, \\
0.05, & \text{if } i = 43, \ldots, 72, \\
0.06, & \text{if } i = 73, \ldots, 93.
\end{cases}$$

Here, we have $\lambda = 4.2$ and $\lambda_2 = 0.201$. Our portfolio is three times larger than Gerber’s (1979, page 53) portfolio. However, in contrast to Gerber’s assumptions, our aim is to discuss an example, where the individual claim amount distributions $Q_i$, $(i \in \{1, \ldots, 93\})$ are non-identical and absolutely continuous with finite mean and
infinite variance. Therefore, we suppose that $Q_i$ has the Pareto-type Lebesgue density

$$h_i(x) = \frac{2}{i(1 + x/i)^3}, \quad (x \in (0, \infty)).$$

Note that the mean of $Q_i$ is equal to $i$, so that, loosely speaking, we cannot say that the $Q_i$ coincide well. But, on the other hand, Table 2 shows that, even in this example, the upper bounds with magic factors are considerably smaller than the comparable ones without magic factors. Further, we see that, as we expect, (11) is much better than (10). Note that, in the case $s = 1$, (17) has been used with $A_1 = 0$. This is not problematic since, as usual, we set $1/0 = \infty$, so that, for $A_1 = 0$, we get $\tilde{\beta}_1(2^{-3/2}A_1^{-1}) = \lambda_2$.

Table 2: Numerical comparison of the bounds

<table>
<thead>
<tr>
<th></th>
<th>bounds with magic factors</th>
<th>bounds without magic factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of formula of $s$ of $A_1$</td>
<td>number of formula of $s$</td>
<td>bound</td>
</tr>
<tr>
<td>(10)</td>
<td>– –</td>
<td>0.506408</td>
</tr>
<tr>
<td>(11)</td>
<td>– –</td>
<td>0.025529</td>
</tr>
<tr>
<td>(14)</td>
<td>– –</td>
<td>0.004989</td>
</tr>
<tr>
<td>(17)</td>
<td>1 0</td>
<td>0.028195</td>
</tr>
<tr>
<td>(17)</td>
<td>2 0.5</td>
<td>0.004066</td>
</tr>
<tr>
<td>(17)</td>
<td>3 0.5</td>
<td>0.000254</td>
</tr>
<tr>
<td>(17)</td>
<td>4 0.5</td>
<td>0.000028</td>
</tr>
</tbody>
</table>

4 Proofs

Proof of Theorem 1. Let $\varepsilon \in (0, 1/2]$ be fixed. Then $a_{i,r,\varepsilon} \in [0, \infty), (i \in \{1, \ldots, n\}, r \in \mathbb{N})$, and pairwise disjoint sets $B_{1,\varepsilon}, B_{2,\varepsilon}, \ldots \in \mathbb{R}^\ell$ exist such that, letting

$$h_{i,\varepsilon} := \sum_{r=1}^{\infty} a_{i,r,\varepsilon} 1(B_{r,\varepsilon}) \quad \text{for } i \in \{1, \ldots, n\},$$

we have, for all $i \in \{1, \ldots, n\}$ and $\mathbf{x} \in \mathbb{R}^\ell$,

$$0 \leq f_i(\mathbf{x}) - h_{i,\varepsilon}(\mathbf{x}) \leq \varepsilon.$$
Here, \(1(B)\) denotes the indicator function of a set \(B \subseteq \mathbb{R}^\ell\). For all \(i\), let
\[
m_{i,\varepsilon} = \int h_{i,\varepsilon} \, dQ.
\]
Then \(m_{i,\varepsilon} = 1 - \int (f_i - h_{i,\varepsilon}) \, dQ \in [1 - \varepsilon, 1]\), that is \(1/2 \leq m_{i,\varepsilon} \leq 1\). For all \(i\), let \(Q_{i,\varepsilon}\) be the probability measure on \((\mathbb{R}^\ell, \mathcal{B}^\ell)\) with \(Q\)-density
\[
f_{i,\varepsilon} = \frac{h_{i,\varepsilon}}{m_{i,\varepsilon}}.
\]
For \(i \in \{1, \ldots, n\}\) and \(r \in \mathbb{N}\), let
\[
q_{i,r,\varepsilon} = \frac{a_{i,r,\varepsilon}}{m_{i,\varepsilon}} Q(B_{r,\varepsilon})
\]
and let the probability measure \(U_{r,\varepsilon}\) be defined by
\[
U_{r,\varepsilon}(\cdot) = \begin{cases} \frac{Q(B_{r,\varepsilon} \cap \cdot)}{Q(B_{r,\varepsilon})}, & \text{if } Q(B_{r,\varepsilon}) > 0, \\ I_0, & \text{otherwise.} \end{cases}
\]
Then, for all \(r\) and all \(i\), \(q_{i,r,\varepsilon} \geq 0\) and \(\sum_{r=1}^{\infty} q_{i,r,\varepsilon} = 1\). Further, for a set \(B \in \mathcal{B}^\ell\) and all \(i \in \{1, \ldots, n\}\),
\[
Q_{i,\varepsilon}(B) = \int_B \sum_{r=1}^{\infty} a_{i,r,\varepsilon} \frac{1(B_{r,\varepsilon})}{m_{i,\varepsilon}} \, dQ = \sum_{r=1}^{\infty} q_{i,r,\varepsilon} U_{r,\varepsilon}(B).
\]
Let
\[
Q_{\varepsilon} = \frac{1}{\lambda} \sum_{i=1}^{n} p_i Q_{i,\varepsilon}, \quad f_{\varepsilon} = \frac{1}{\lambda} \sum_{i=1}^{n} p_i f_{i,\varepsilon}.
\]
Then \(Q_{\varepsilon}\) has the \(Q\)-density \(f_{\varepsilon}\) and, for all \(r \in \mathbb{N}\), we have
\[
Q_{\varepsilon}(B_{r,\varepsilon}) = \frac{1}{\lambda} \sum_{i=1}^{n} p_i \frac{a_{i,r,\varepsilon}}{m_{i,\varepsilon}} Q(B_{r,\varepsilon}) = \frac{1}{\lambda} \sum_{i=1}^{n} p_i q_{i,r,\varepsilon} =: q_{r,\varepsilon}.
\]
Let
\[
R_{\varepsilon}^{(1)} = \sum_{i=1}^{n} ((1 - p_i) I_0 + p_i Q_{i,\varepsilon}) \quad \text{and} \quad R_{\varepsilon}^{(2)} = \text{CPo}(\lambda, Q_{\varepsilon}).
\]
Then
\[
d_{r} \leq d_{TV}(\mathcal{L}(S_n), \ R_{\varepsilon}^{(1)}) + d_{TV}(\ R_{\varepsilon}^{(1)}, \ R_{\varepsilon}^{(2)}) + d_{TV}(\ R_{\varepsilon}^{(2)}, \ \text{CPo}(\lambda, Q))
\]
\[
=:\ T_{\varepsilon}^{(1)} + T_{\varepsilon}^{(2)} + T_{\varepsilon}^{(3)}, \quad \text{say.}
\]
In what follows, we use some basic properties of the total variation distance: Firstly, it is subadditive, that is, if $W_i, \tilde{W}_i, (i \in \{1, \ldots, n\})$ are probability measures on $(\mathbb{R}^\ell, \mathcal{B}^\ell)$, then
\[
\text{d}_{TV}(\star_{i=1}^n W_i, \star_{i=1}^n \tilde{W}_i) \leq \sum_{i=1}^n \text{d}_{TV}(W_i, \tilde{W}_i).
\]
Secondly, if $W_1$ and $W_2$ have densities $w_1$ and $w_2$ with respect to a measure $\mu$ on $(\mathbb{R}^\ell, \mathcal{B}^\ell)$, then
\[
\text{d}_{TV}(W_1, W_2) = \frac{1}{2} \int |w_1 - w_2| d\mu.
\]
Now, we obtain
\[
T^{(1)}_\varepsilon \leq \sum_{i=1}^n p_i \text{d}_{TV}(Q_i, Q_i, \varepsilon),
\]
where
\[
\text{d}_{TV}(Q_i, Q_i, \varepsilon) = \frac{1}{2} \int |f_i - f_i, \varepsilon| dQ \leq \frac{1}{2} \int \left(|f_i - h_i, \varepsilon| + h_i, \varepsilon \right) dQ \leq \frac{1}{2}(\varepsilon + 1 - m_i, \varepsilon) \leq \varepsilon.
\]
This gives $T^{(1)}_\varepsilon \leq \lambda \varepsilon$. On the other hand, we have
\[
T^{(3)}_\varepsilon \leq \sum_{m=1}^\infty e^{-\lambda} \frac{\lambda^m}{m!} \text{d}_{TV}(Q^{*m}_\varepsilon, Q^{*m}),
\]
where, in view of the above, we see that
\[
\text{d}_{TV}(Q^{*m}_\varepsilon, Q^{*m}) \leq m \text{d}_{TV}(Q_\varepsilon, Q) \leq \frac{m}{\lambda} \sum_{i=1}^n p_i \text{d}_{TV}(Q_i, Q_i, Q_i) \leq m \varepsilon.
\]
Hence
\[
T^{(3)}_\varepsilon \leq \varepsilon \sum_{m=1}^\infty e^{-\lambda} \frac{\lambda^m}{(m-1)!} = \lambda \varepsilon.
\]
Estimating $T^{(2)}_\varepsilon$ with the help of (4) and using the inequalities already proved, we get
\[
d_r \leq 2\lambda \varepsilon + 8.8 \sum_{i=1}^n p_i^2 \min \left\{ \frac{1}{\lambda} \sum_{r=1}^\infty q_{i,r,\varepsilon}^2 q_{r,\varepsilon}, 1 \right\}.
\]
A similar inequality can be written down by using (7). By letting $\varepsilon \to 0$, we see that, in order to prove the assertion, we have to verify that, for all $i \in \{1, \ldots, n\}$,
\[ \sum_{r=1}^{\infty} \frac{q_{i,r,\varepsilon}^2}{q_{r,\varepsilon}} \rightarrow \int f_i^2 \, dQ, \] as \( \varepsilon \to 0. \) Note that, for \( i, j \in \{1, \ldots, n\}, \) \( \int f_i f_j \, dQ \leq \lambda/p_j < \infty, \) since

\[
1 = \int f_i \, dQ = \frac{1}{\lambda} \sum_{j=1}^{\infty} p_j \int f_i \, dQ_j = \frac{1}{\lambda} \sum_{j=1}^{\infty} p_j \int f_i f_j \, dQ.
\]

In particular \( \int f_i^2 \, dQ < \lambda/p_i < \infty. \) Now, we derive

\[
\left| \int f_i^2 \, dQ - \sum_{r=1}^{\infty} \frac{q_{i,r,\varepsilon}^2}{q_{r,\varepsilon}} \right| \leq \left| \int (f_i^2 - f_{i,\varepsilon}^2) \, dQ \right| + \left| \int f_{i,\varepsilon}^2 \, dQ - \sum_{r=1}^{\infty} \frac{q_{i,r,\varepsilon}^2}{q_{r,\varepsilon}} \right|
=: J_{i,\varepsilon}^{(1)} + J_{i,\varepsilon}^{(2)}, \text{ say.}
\]

On the one hand,

\[
J_{i,\varepsilon}^{(1)} \leq \int \left| f_i - f_{i,\varepsilon}\right| (f_i + f_{i,\varepsilon}) \, dQ
\leq \int \left( \varepsilon + h_{i,\varepsilon}\left( \frac{1}{m_{i,\varepsilon}} - 1 \right) \right) (f_i + f_{i,\varepsilon}) \, dQ
\leq 2\varepsilon \left( 1 + \int h_{i,\varepsilon}(f_i + f_{i,\varepsilon}) \, dQ \right).
\]

Using the inequalities \( h_{i,\varepsilon} \leq f_i \) and \( f_{i,\varepsilon} \leq 2h_{i,\varepsilon} \leq 2f_i, \) we get

\[
J_{i,\varepsilon}^{(1)} \leq 2\varepsilon \left( 1 + 3 \int f_i^2 \, dQ \right) \leq 2\varepsilon \left( 1 + 3\lambda/p_i \right) (\varepsilon \to 0) \to 0.
\]

On the other hand,

\[
J_{i,\varepsilon}^{(2)} = \left| \int \sum_{r=1}^{\infty} \frac{a_{i,r,\varepsilon}^2}{m_{i,\varepsilon}^2} \mathbf{1}(B_{r,\varepsilon}) \, dQ - \sum_{r=1}^{\infty} \frac{q_{i,r,\varepsilon}^2}{q_{r,\varepsilon}} \right|
\leq \sum_{r \in \mathbb{N} : q_{r,\varepsilon} > 0} \frac{a_{i,r,\varepsilon}^2}{m_{i,\varepsilon}^2} |Q_{r,\varepsilon} - Q(B_{r,\varepsilon})|
\leq \frac{\lambda}{p_i} \sum_{r=1}^{\infty} \frac{a_{i,r,\varepsilon}}{m_{i,\varepsilon}} |Q(\varepsilon B_{r,\varepsilon}) - Q(B_{r,\varepsilon})|
\leq \frac{1}{p_i} \sum_{j=1}^{\infty} p_j \sum_{r=1}^{\infty} \frac{a_{i,r,\varepsilon}}{m_{i,\varepsilon}} \int_{B_{r,\varepsilon}} \left| f_{j,\varepsilon} - f_j \right| \, dQ
= \frac{1}{p_i} \sum_{j=1}^{\infty} p_j \int \sum_{r=1}^{\infty} \frac{a_{i,r,\varepsilon}}{m_{i,\varepsilon}} \mathbf{1}(B_{r,\varepsilon}) \left| f_{j,\varepsilon} - f_j \right| \, dQ
\leq \frac{\varepsilon}{p_i m_{i,\varepsilon}} \sum_{j=1}^{\infty} p_j \int f_i (1 + 2f_j) \, dQ
\leq \frac{2\varepsilon}{p_i} \sum_{j=1}^{\infty} p_j \left( 1 + \frac{2\lambda}{p_i} \right) (\varepsilon \to 0) \to 0.
\]
This completes the proof of Theorem 1. □

**Proof of Theorem 2.** We use the assumptions and definitions from the proof of Theorem 1. Letting

\[ R_\varepsilon^{(3)} = \left( I_0 - \frac{1}{2} \sum_{i=1}^{n} p_i^2 (Q_{i, \varepsilon} - I_0) \ast \ast^2 \right) \ast \text{CPo}(\lambda, Q_\varepsilon), \]

we obtain

\[
\begin{align*}
\text{d}_{\text{TV}}(\mathcal{L}(S_n), \text{CPo}_2(\lambda, Q)) & \leq \text{d}_{\text{TV}}(\mathcal{L}(S_n), R_\varepsilon^{(1)}) \\
& + \text{d}_{\text{TV}}(R_\varepsilon^{(1)}, R_\varepsilon^{(3)}) + \text{d}_{\text{TV}}(R_\varepsilon^{(3)}, \text{CPo}_2(\lambda, Q)) \\
& =: T_{\varepsilon}^{(1)} + T_{\varepsilon}^{(4)} + T_{\varepsilon}^{(5)}, \text{ say.}
\end{align*}
\]

From Theorem 3 in Roos (2003), it follows that

\[
T_{\varepsilon}^{(4)} \leq \frac{4}{3} \gamma_\varepsilon^2 + \alpha_\varepsilon^2 \left( 1 + \frac{0.82 \alpha_\varepsilon}{1 - 2^{-1/2} e^{\alpha_\varepsilon}} \right),
\]

where

\[
\begin{align*}
\gamma_\varepsilon &= \sum_{i=1}^{n} g_2(2p_i) p_i^3 \min \left\{ 0.46 \left( \frac{1}{\lambda} \sum_{r=1}^{\infty} q_{i,r, \varepsilon} q_{r, \varepsilon} \right)^{3/2}, 1 \right\}, \\
\alpha_\varepsilon &= \sum_{i=1}^{n} g_1(2p_i) p_i^2 \min \left\{ \frac{1}{\lambda} \sum_{r=1}^{\infty} q_{i,r, \varepsilon} q_{r, \varepsilon}, 1 \right\},
\end{align*}
\]

and we assume that \( \alpha_\varepsilon < 2^{1/2} e^{-1} \). In the proof of Theorem 1, we have shown that

\[
\lim_{\varepsilon \to 0} T_{\varepsilon}^{(1)} = 0 \quad \text{and} \quad \lim_{\varepsilon \to 0} \sum_{r=1}^{\infty} \frac{q_{r, \varepsilon}^2}{q_{r, \varepsilon}} = \int f_i^2 \, dQ.
\]

Further, it is easily proved that \( \lim_{\varepsilon \to 0} T_{\varepsilon}^{(5)} = 0 \). This completes the proof. □

The proof of Theorem 3 is based on Theorem 1 in Roos (2002). It is similar to the above and is therefore omitted. For the proof of Proposition 1, we need the following lemma.

**Lemma 1** For \( i \in \{1, \ldots, n\} \) and \( r \in \mathbb{N} \), let \( b_{i,r} \in [0, \infty) \) with \( \tilde{b}_i = \sum_{r=1}^{\infty} b_{i,r} < \infty \), \( \sum_{i=1}^{n} \tilde{b}_i > 0 \), and \( b'_r = \sum_{i=1}^{n} b_{i,r} \). Then

\[
\frac{\sum_{i=1}^{n} \tilde{b}_i^2}{\sum_{i=1}^{n} \tilde{b}_i} \leq \sum_{r \in \mathbb{N} : b'_r > 0} \sum_{i=1}^{n} \frac{b_{i,r}^2}{b'_r}.
\]
Proof. It is not difficult to show that, if $b_1', b_2' > 0$ and $i \in \{1, \ldots, n\}$,

$$\frac{(b_{i,1} + b_{i,2})^2}{b_1' + b_2'} \leq \frac{b_{i,1}^2}{b_1'} + \frac{b_{i,2}^2}{b_2'}.$$

The proof is easily completed. \(\square\)

Proof of Proposition 1. The Radon–Nikodym theorem says that, for all $r \in \mathbb{N}$, $U_r$ has a density $u_r : \mathbb{R}^\ell \to [0, \infty)$ with respect to the probability measure $\mu = e^{-1} \sum_{r=1}^\infty U_r/(r - 1)!$. Then $Q_i$ and $Q$ have the $\mu$-densities $h_i := \sum_{r=1}^\infty q_{i,r} u_r$ and $h := \sum_{r=1}^\infty q_r u_r$, respectively. Using Lemma 1, we obtain

$$\sum_{i=1}^n p_i^2 \int f_i^2 dQ = \sum_{i=1}^n p_i^2 \int_{\{h > 0\}} \frac{h_i^2}{h} d\mu$$

$$= \lambda \int_{\{h > 0\}} \frac{\sum_{i=1}^n (\sum_{r=1}^\infty p_i q_{i,r} u_r)^2}{\sum_{i=1}^n \sum_{r=1}^\infty p_i q_{i,r} u_r} d\mu$$

$$\leq \lambda \sum_{r=1}^\infty \int 1(\{\sum_{i=1}^n p_i q_{i,r} u_r > 0\}) \frac{\sum_{i=1}^n (p_i q_{i,r} u_r)^2}{\sum_{i=1}^n p_i q_{i,r} u_r} d\mu$$

$$= \sum_{i=1}^n p_i^2 \nu_i.$$  

The proposition is shown. \(\square\)

Acknowledgement

The author thanks the referee for his remarks, which led to an improvement of the presentation of the paper.

References


