APPROXIMATION FOR EXPECTATIONS OF UNBOUNDED FUNCTIONS OF DEPENDENT INTEGER-VALUED RANDOM VARIABLES

PAVEL S. RUZANKIN,* *Sobolev Institute of Mathematics*

Abstract

Expectations of unbounded functions of dependent nonnegative integer-valued random variables are approximated by the expectations of the functions of independent copies of these random variables. The Lindeberg method is used.

Keywords: Unbounded function; reliability system; Poisson approximation; Lindeberg's method

2010 Mathematics Subject Classification: Primary 60F05
Secondary 60E07

1. The main results

Let $\xi_1, \xi_2, \ldots, \xi_n$ be nonnegative integer-valued random variables, and let $\eta_1, \ldots, \eta_n$ be independent copies of $\xi_1, \ldots, \xi_n$, respectively (i.e. η_j coincides in distribution with ξ_j for each j). Our main goal is to approximate $\mathrm{E}\,F(\xi_1, \ldots, \xi_n)$ by $\mathrm{E}\,F(\eta_1, \ldots, \eta_n)$, where the function F can be unbounded.

An example where this problem occurs is as follows. Consider a point process Ξ on an interval $[A, B]$. If we want to construct a compound Poisson approximation for $\Xi([A, B])$, we can use the Bernstein block technique. In order to do this, we present $\Xi([A, B])$ as the following sum:

$$\Xi([A, B]) = \sum_{j=1}^n \Xi(U_j),$$

where the U_j are intervals such that $[A, B] = \bigcup_{j=1}^n U_j$ and $U_j \cap U_i = \varnothing$ for $i \neq j$. Our next step is to approximate each $\xi_j = \Xi(U_j)$ with $\gamma_j = \Xi(V_j)$, where $V_j \subset U_j$, $\gamma_1, \ldots, \gamma_n$ are weakly dependent, and $\mathrm{P}(\xi_j \neq \gamma_j)$ is small. Then $\sum_{j=1}^n \gamma_j$ can be approximated by the corresponding compound Poisson distribution. Barbour *et al.* (2002) employed this technique for approximation in terms of the total variation distance and some Kantorovich (Wasserstein) distances (in fact, not only was $\Xi([A, B])$ approximated, but a more complex approximation was built). However, this approach becomes unsuitable if we have to approximate $\mathrm{E}\,F(\xi_1, \ldots, \xi_n)$ for unbounded functions F, because even if $\sum_{j=1}^n \mathrm{P}(\xi_j \neq \gamma_j)$ is small, the contribution of $\Xi(U_j \setminus V_j)$ to the expectation may be large. In the present paper we show that, for unbounded functions, the Lindeberg method can be applied. Example 1, below, illustrates which sort of conditions can be used in this case.

To state the main result, we need the following notation. As above, let $\xi_1, \ldots, \xi_n$ be (dependent) nonnegative integer-valued random variables, and let the random variables $\eta_1, \ldots, \eta_n$ be

independent, independent of $\xi_1, \ldots, \xi_n$, and such that, for each j, η_j coincides in distribution with ξ_j. We will use the difference operator $\Delta f(j) \equiv \Delta_j f(j) := f(j+1) - f(j)$ and the so-called 'factorial power' $a_{[r]} := a(a-1) \cdots (a-r+1)$, where we assume that $a_{[0]} = 1$. We denote by $I(A)$ the indicator of an event A. The main result of the present paper is the following theorem.

Theorem 1. *If, for some $k \geq 1$, the expectations below exist, then the following equality holds:*

$$\mathrm{E}\, F(\xi_1, \ldots, \xi_n) - \mathrm{E}\, F(\eta_1, \ldots, \eta_n) = \sum_{j=2}^{n} \left(\sum_{r=1}^{k} \frac{1}{r!} \mathrm{cov}(\Delta^r F_j(0), (\xi_j)_{[r]} I(\xi_j \leq k)) \right.$$
$$\left. + \mathrm{E}\, F_j(\xi_j) I(\xi_j > k) - \mathrm{E}\, F_j(\eta_j) I(\eta_j > k) \right),$$

where

$$F_j(i) = F(\xi_1, \ldots, \xi_{j-1}, i, \eta_{j+1}, \ldots, \eta_n).$$

To prove this theorem, we need the following lemma.

Lemma 1. *For any function f, and all integers $k \geq 0$ and $j \geq 0$,*

$$\sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) j_{[r]} = f(j) \quad \text{if } k \geq j, \tag{1}$$

$$\sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) j_{[r]} = \frac{j_{[k+1]}}{k!} \Delta_s^k \left(\frac{f(s)}{j-s} \right) \bigg|_{s=0} \quad \text{if } k < j,$$

where, for a function g, $\Delta_s g(s) := g(s+1) - g(s)$.

Proof. We can assume that $k \leq j$, because $j_{[r]} = 0$ for $r > j$. We have

$$\sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) j_{[r]} = \sum_{r=0}^{k} \sum_{s=0}^{r} (-1)^{r-s} C_r^s f(s) C_j^r$$
$$= \sum_{s=0}^{k} f(s) \sum_{r=s}^{k} \frac{(-1)^{r-s} j!}{s!\,(r-s)!\,(j-r)!}$$
$$= \frac{1}{k!} f(k) j_{[k]} + \sum_{s=0}^{k-1} f(s) C_j^s \sum_{t=0}^{k-s} (-1)^t C_{j-s}^t,$$

where in the last equality we have used the change of variable $t = r - s$. To prove (1), it remains to note that $\sum_{t=0}^{k-s} (-1)^t C_{j-s}^t = 0$ if $j = k$.

If $j > k$ then $\sum_{t=0}^{k-s} (-1)^t C_{j-s}^t = (-1)^{k-s} C_{j-s-1}^{k-s}$. Thus, for $j > k$,

$$\sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) j_{[r]} = \sum_{s=0}^{k} f(s)(-1)^{k-s} C_j^s C_{j-s-1}^{k-s} = \frac{j_{[k+1]}}{k!} \sum_{s=0}^{k} (-1)^{k-s} C_k^s \frac{f(s)}{j-s}.$$

This completes the proof.

Corollary 1. *Let ψ be an arbitrary nonnegative integer-valued random variable, and let f be an arbitrary function. If the expectation $\mathrm{E} f(\psi)$ exists then, for any $k \geq 0$,*

$$\mathrm{E} f(\psi) = \sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) \, \mathrm{E} \, \psi_{[r]} I(\psi \leq k) + \mathrm{E} f(\psi) I(\psi > k). \tag{2}$$

Moreover,

$$\mathrm{E} f(\psi) = \sum_{r=0}^{\infty} \frac{1}{r!} \Delta^r f(0) \, \mathrm{E} \, \psi_{[r]}, \tag{3}$$

if all these expectations exist and the series converges absolutely.

Proof. By Lemma 1 we have

$$\mathrm{E} f(\psi) = \sum_{j=0}^{k} \mathrm{E} f(\psi) I(\psi = j) + \mathrm{E} f(\psi) I(\psi > k)$$

$$= \sum_{j=0}^{k} \sum_{r=0}^{k} \frac{1}{r!} \Delta^r f(0) \, \mathrm{E} \, \psi_{[r]} I(\psi = j) + \mathrm{E} f(\psi) I(\psi > k).$$

This implies (2). Relation (3) follows from (2) because, as $k \to \infty$,

$$\mathrm{E} \, \psi_{[r]} I(\psi \leq k) \to \mathrm{E} \, \psi_{[r]}, \qquad \mathrm{E} f(\psi) I(\psi > k) \to 0.$$

This completes the proof.

The following corollary is an immediate consequence of Corollary 1.

Corollary 2. *Let ϕ be an arbitrary random variable, let ψ be an arbitrary nonnegative integer-valued random variable, and let $f_y(j) \equiv f(y, j)$ be an arbitrary function. If, for some $k \geq 1$, all the expectations in (4) below exist then*

$$\mathrm{E} f_\phi(\psi) = \sum_{r=0}^{k} \frac{1}{r!} \mathrm{E} \, \Delta^r f_\phi(0) \psi_{[r]} I(\psi \leq k) + \mathrm{E} f_\phi(\psi) I(\psi > k). \tag{4}$$

Moreover,

$$\mathrm{E} f_\phi(\psi) = \sum_{r=0}^{\infty} \frac{1}{r!} \mathrm{E} \, \Delta^r f_\phi(0) \psi_{[r]},$$

if all these expectations exist and the series converges absolutely.

Theorem 1 follows immediately from relation (4) and the following identity, which is an application of the Lindeberg method,

$$\mathrm{E} F(\xi_1, \ldots, \xi_n) - \mathrm{E} F(\eta_1, \ldots, \eta_n) = \sum_{j=2}^{n} (\mathrm{E} F_j(\xi_j) - \mathrm{E} F_j(\eta_j)).$$

The following corollary, which is presumably well known, will be used in Example 2, below.

Corollary 3. *If $m \geq 0$ is an integer, and a and b are arbitrary real numbers, then*

$$(a + b)_{[m]} = \sum_{r=0}^{m} C_m^r a_{[r]} b_{[m-r]}. \tag{5}$$

Proof. Relation (5) for $b \in \{0, 1, \ldots, m\}$ follows from Lemma 1 if we set

$$f(j) = (a + j)_{[m]}, \qquad k = m.$$

Both sides of (5) are polynomials of b of order m; hence, (5) is valid for any real b. This completes the proof.

Example 1. Set $k = 1$ in Theorem 1. We have

$$\mathrm{E}\, F(\xi_1, \ldots, \xi_n) - \mathrm{E}\, F(\eta_1, \ldots, \eta_n)$$

$$= \sum_{j=2}^{n} (\mathrm{cov}(\Delta F_j(0), I(\xi_j = 1)) + \mathrm{E}\, F_j(\xi_j) I(\xi_j \geq 2) - \mathrm{E}\, F_j(\eta_j) I(\eta_j \geq 2)).$$

In the last expression

$$\mathrm{cov}(\Delta F_j(0), I(\xi_j = 1))$$
$$= \mathrm{E}(\Delta F_j(0) \mid \xi_j = 1)\, \mathrm{P}(\xi_j = 1) - \mathrm{E}(\Delta F_j(0) \mid \xi_j = 1)(\mathrm{P}(\xi_j = 1))^2$$
$$\quad - \mathrm{E}(\Delta F_j(0) \mid \xi_j \neq 1)\, \mathrm{P}(\xi_j \neq 1)\, \mathrm{P}(\xi_j = 1)$$
$$= p_j(1 - p_j)(\mathrm{E}(\Delta F_j(0) \mid \xi_j = 1) - \mathrm{E}(\Delta F_j(0) \mid \xi_j \neq 1)),$$

where $p_j = \mathrm{P}(\xi_j = 1)$.

Now let

$$F(j_1, \ldots, j_n) = (j_1 + \cdots + j_n)_{[m]}.$$

Then $\Delta F_j(0) = m(\xi_1 + \cdots + \xi_{j-1} + \eta_{j+1} + \cdots + \eta_n)_{[m-1]}$. We assume that

$$(\xi_1 + \cdots + \xi_{j-1} \mid \xi_j = 1) \leq_{\mathrm{st}} (\xi_1 + \cdots + \xi_{j-1} \mid \xi_j \neq 1) + c,$$
$$(\xi_1 + \cdots + \xi_{j-1} \mid \xi_j \neq 1) \leq_{\mathrm{st}} (\xi_1 + \cdots + \xi_{j-1} \mid \xi_j = 1) + c,$$

where $c \geq 1$ is a constant, $\alpha \mid A$ denotes a random variable distributed by the conditional distribution of the random variable α under condition A, and $\alpha \leq_{\mathrm{st}} \beta$ means that there can be constructed on a common probability space random variables $\tilde{\alpha}$ and $\tilde{\beta}$ such that $\tilde{\alpha} \leq \tilde{\beta}$, $\tilde{\alpha}$ coincides in distribution with α and $\tilde{\beta}$ coincides in distribution with β. Then

$$|\mathrm{E}(\Delta F_j(0) \mid \xi_j = 1) - \mathrm{E}(\Delta F_j(0) \mid \xi_j \neq 1)|$$
$$\leq cm(m - 1)\, \mathrm{E}(\xi_1 + \cdots + \xi_{j-1} + \eta_{j+1} + \cdots + \eta_n + c - 1)_{[m-2]}.$$

Finally, we obtain

$$|\mathrm{E}F(\xi_1, \ldots, \xi_n) - \mathrm{E}F(\eta_1, \ldots, \eta_n)|$$

$$\leq \sum_{j=2}^{n} cm^2 p_j(1 - p_j)\, \mathrm{E}(\xi_1 + \cdots + \xi_{j-1} + \eta_{j+1} + \cdots + \eta_n + c)_{[m-2]}$$

$$+ \left| \sum_{j=2}^{n} (\mathrm{E}\, F_j(\xi_j) I(\xi_j \geq 2) - \mathrm{E}\, F_j(\eta_j) I(\eta_j \geq 2)) \right|.$$

2. Dependent Bernoulli random variables

In this section $\xi_1, \ldots, \xi_n$ will be Bernoulli random variables with success probabilities $P(\xi_j = 1) = 1 - P(\xi_j = 0) = p_j$. Repeating the reasoning of Example 1 for this case we obtain the following corollary.

Corollary 4. *We have*

$$\text{E}\, F(\xi_1, \ldots, \xi_n) - \text{E}\, F(\eta_1, \ldots, \eta_n)$$

$$= \sum_{k=2}^{n} p_k \{\text{E}(\Delta F_j(0) \mid \xi_k = 1) - \text{E}(\Delta F_j(0))\}$$

$$= \sum_{k=2}^{n} p_k (1 - p_k) \{\text{E}(\Delta F_j(0) \mid \xi_k = 1) - \text{E}(\Delta F_j(0) \mid \xi_k = 0)\},$$

where the random functions F_j are defined in Theorem 1.

Example 2. (*Reliability systems.*)   Many reliability systems can be described as follows. Consider m independent Bernoulli random variables $\zeta_1, \ldots, \zeta_m$ with $P(\zeta_j = 1) = q_j$. Let ξ_k, $k = 1, \ldots, n$, be the products of the corresponding families of ζ_js:

$$\xi_k = \prod_{i=1}^{d(k)} \zeta_{l(i,k)}.$$

Each of the random variables $\xi_1, \ldots, \xi_n$ is responsible for the failure of the corresponding element of the system: $\xi_k = 1$ if the kth element fails. Define

$$\Gamma(k) = \{j : \xi_k \text{ depends on } \xi_j\}, \qquad p_k = \text{E}\, \xi_k = \prod_{i=1}^{d(k)} q_{l(i,k)}.$$

Note that $\eta_1, \ldots, \eta_n$ are the independent Bernoulli random variables with $P(\eta_j = 1) = p_k$.

First let us consider the total variation distance between the distributions of $(\xi_1, \ldots, \xi_n)$ and $(\eta_1, \ldots, \eta_n)$:

$$d_{\text{TV}}(\mathcal{L}(\xi_1, \ldots, \xi_n), \mathcal{L}(\eta_1, \ldots, \eta_n)) := \sup_{\{F : |F| \le 1/2\}} |\text{E}F(\xi_1, \ldots, \xi_n) - \text{E}F(\eta_1, \ldots, \eta_n)|.$$

Applying Corollary 4 we obtain

$$d_{\text{TV}}(\mathcal{L}(\xi_1, \ldots, \xi_n), \mathcal{L}(\eta_1, \ldots, \eta_n)) \le \sum_{k=2}^{n} \sum_{j \in \Gamma(k) \cap \{1, \ldots, k-1\}} \text{E}\, \xi_k \xi_j. \qquad (6)$$

Note that

$$d_{\text{TV}}(\mathcal{L}(\xi_1 + \cdots + \xi_n), \mathcal{L}(\eta_1 + \cdots + \eta_n)) \le d_{\text{TV}}(\mathcal{L}(\xi_1, \ldots, \xi_n), \mathcal{L}(\eta_1, \ldots, \eta_n)),$$

and, hence, bound (6) can be applied to approximate the distribution of $\xi_1 + \cdots + \xi_n$. In particular, applying the well-known estimate $d_{\text{TV}}(\mathcal{L}(\eta_1 + \cdots + \eta_n), \mathcal{P}) \le \min\{1, 1/\lambda\} \sum_k p_k^2$ (see Barbour and Hall (1984)), we can obtain bounds on $d_{\text{TV}}(\mathcal{L}(\xi_1 + \cdots + \xi_n), \mathcal{P})$, where $\lambda = \sum_{k=1}^{n} p_k$ and $\mathcal{P}$ is the Poisson distribution with parameter λ.

It is interesting to compare the estimate in (6) with the corresponding results that are derived with the Stein–Chen method. Let us consider the so-called connected-s systems, i.e. when $d(k) = s$ for all k (and no other restrictions are imposed). In a number of works (see Barbour and Chryssaphinou (2001) and the references therein), compound Poisson approximation for the sum $\xi_1 + \cdots + \xi_n$ was studied using the Stein–Chen method. When λ is upper bounded, the estimates are of the following form:

$$d_{\mathrm{TV}}(\mathcal{L}(\xi_1 + \cdots + \xi_n), \mathcal{Q}) \leq c q_{\max}^R \lambda, \tag{7}$$

where $\mathcal{Q}$ is the corresponding compound Poisson distribution, c is some constant, $q_{\max} := \max_j q_j$, and the parameter $R \geq 1$ corresponds to the complexity of the approximation (the bigger R is, the more complex the distribution $\mathcal{Q}$). If, roughly speaking, $q_{\max}^R \max_j p_j >> \max_{j \neq k} \mathrm{E}\, \xi_j \xi_k$ then the bound in (6) turns out to be better than the bound in (7), despite the more complex nature of the compound Poisson approximation.

Now let us, as in Example 1, consider the approximation for

$$F(j_1, \ldots, j_n) = (j_1 + \cdots + j_n)_{[m]}.$$

Note that $\mathrm{E}\, F(\xi_1, \ldots, \xi_n) \geq \mathrm{E}\, F(\eta_1, \ldots, \eta_n)$ by Corollary 3. We have

$$\mathrm{E}\, F(\xi_1, \ldots, \xi_n) - \mathrm{E}\, F(\eta_1, \ldots, \eta_n) = \sum_{k=2}^{n} p_j (\mathrm{E}(\Delta F_k(0) \mid \xi_k = 1) - \mathrm{E}(\Delta F_k(0)))$$

$$\leq \sum_{k=2}^{n} \left(\sum_{j \in \Gamma(k) \cap \{1, \ldots, k-1\}} \mathrm{E}\, \xi_k \xi_j \right) \mathrm{E}\, \Delta^2 F_k(\nu_k - 1),$$

where ν_j is the number of elements in the set $\Gamma(j) \cap \{1, \ldots, j-1\}$. Furthermore,

$$\mathrm{E}\, \Delta^2 F_k(\nu_k - 1) = m(m-1) \mathrm{E}(\xi_1 + \cdots + \xi_{k-1} + \eta_{k+1} + \cdots + \eta_n + \nu_k - 1)_{[m-2]}$$

$$\leq m(m-1) \mathrm{E}(\xi_1 + \cdots + \xi_n + \nu_k - 1)_{[m-2]},$$

where the last inequality follows from Corollary 3. Hence, finally,

$$\mathrm{E}\, F(\xi_1, \ldots, \xi_n) - \mathrm{E}\, F(\eta_1, \ldots, \eta_n)$$

$$\leq m^2 \sum_{k=2}^{n} \left(\sum_{j \in \Gamma(k) \cap \{1, \ldots, k-1\}} \mathrm{E}\, \xi_k \xi_j \right) \mathrm{E}(\xi_1 + \cdots + \xi_n + \nu_k - 1)_{[m-2]}.$$

3. Application to Poisson approximation

The results of the present paper can be applied to Poisson approximations. Firstly, we approximate the random variables $\xi_1, \ldots, \xi_n$ by their independent copies $\eta_1, \ldots, \eta_n$. Secondly, we approximate $\eta_1, \ldots, \eta_n$ by the accompanying compound Poisson random variables.

If $\xi_1, \ldots, \xi_n$ are Bernoulli random variables then the following theorem (see Borisov and Ruzankin (2002)) can be used.

Theorem 2. *Let $\zeta_1, \ldots, \zeta_n$ be independent Poisson random variables with parameters $\mathrm{E}\, \zeta_j = p_j := \mathrm{P}(\xi_j = 1)$. If $\mathrm{E}\,|F(\zeta_1, \ldots, \zeta_n)| < \infty$ then*

$$\mathrm{E}\, F(\zeta_1, \ldots, \zeta_n) - \mathrm{E}\, F(\eta_1, \ldots, \eta_n) = \sum_{j=1}^{n} \sum_{r=2}^{\infty} \frac{p_j^r}{r!} \mathrm{E}\, \Delta_{(j)}^r F(\eta_1, \ldots, \eta_{j-1}, 0, \zeta_{j+1}, \ldots, \zeta_n),$$

$$\tag{8}$$

where, for each j, the corresponding series in (8) converges absolutely. The notation $\Delta_{(j)}$ means that the corresponding difference is taken with respect to the jth argument.

Moreover, for $k \geq 2$,

$$\left| \sum_{r=k}^{\infty} \frac{p_j^r}{r!} \, \mathrm{E} \, \Delta_{(j)}^r F(\eta_1, \ldots, \eta_{j-1}, 0, \zeta_{j+1}, \ldots, \zeta_n) \right|$$

$$\leq \mathrm{e}^{p_j} \frac{p_j^k}{k!} \, \mathrm{E} \, |\Delta_{(j)}^k F(\eta_1, \ldots, \eta_{j-1}, \zeta_j, \ldots, \zeta_n)|,$$

where the right-hand side is finite if and only if $\mathrm{E} \, \zeta_j^k |F(\eta_1, \ldots, \eta_{j-1}, \zeta_j, \ldots, \zeta_n)| < \infty$.

If $\xi_1, \ldots, \xi_n$ are not Bernoulli distributed, but have large atoms at 0, a compound Poisson approximation can be used. First we apply Theorem 1. Then we approximate each η_j by the compound Poisson distribution $\mathrm{e}^{p_j(\mathcal{L}_j - 1)}$, where $p_j = \mathrm{P}(\xi_j \neq 0)$ and $\mathcal{L}_j$ is the conditional distribution of ξ_j under the condition $\xi_j \neq 0$. The error estimates (complete asymptotic expansions) for this approximation can be found in Borisov and Ruzankin (2002) (see also the references therein) and Barbour (1987).

References

BARBOUR, A. D. (1987). Asymptotic expansions in the Poisson limit theorem. *Ann. Prob.* **15,** 748–766.

BARBOUR, A. D. AND CHRYSSAPHINOU, O. (2001). Compound Poisson approximation: a user's guide. *Ann. Appl. Prob.* **11,** 964–1002.

BARBOUR, A. D. AND HALL, P. (1984). On the rate of Poisson convergence. *Math. Proc. Camb. Phil. Soc.* **95,** 473–480.

BARBOUR, A. D., NOVAK, S. Y. AND XIA, A. (2002). Compound Poisson approximation for the distribution of extremes. *Adv. Appl. Prob.* **34,** 223–240.

BORISOV, I. S. AND RUZANKIN, P. S. (2002). Poisson approximation for expectations of unbounded functions of independent random variables. *Ann. Prob.* **30,** 1657–1680.