

A HIGHER ORDER LARGE-DEVIATION APPROXIMATION FOR THE DISCRETE DISTRIBUTIONS

Masafumi Akahira* and Kunihiro Takahashi**

For a sum of independent discrete random variables, its higher order large-deviation approximation is discussed. An approximation to the tail probability of the distribution of the sum is provided, and its numerical comparison with other approximations is done in the binomial case. Consequently, the approximation formula is seen to be more accurate.

Key words and phrases: Large-deviation approximation; Edgeworth expansion; Saddlepoint approximation; Binomial distribution.

1. Introduction

In the higher order asymptotics, the Edgeworth expansion for the distribution of statistics is very useful to compare statistical procedures like estimators and tests (see, e.g. Akahira and Takeuchi (1981), Pfanzagl and Wefelmeyer (1985), Ghosh (1994)). On the other hand, the large-deviation approximation plays an important role in the asymptotic efficiency of tests (see, e.g. Bahadur (1971) and Bucklew (1990)).

Recently, the higher order large-deviation approximation for the distribution of the sum of independent discrete random variables was discussed by Akahira *et al.* (1999). The approximation is closely connected with the saddlepoint approximation (see, e.g. Daniels (1954, 1987), Lugannani and Rice (1980), Barndorff-Nielsen and Cox (1989), Jensen (1995)).

In this paper, a higher order large-deviation approximation to the tail probability for the distribution of the sum of independent discrete random variables is given. A numerical comparison with others, including the saddlepoint approximation, is done in the binomial case. Consequently, the higher order approximation gives sufficiently accurate results.

2. The large-deviation approximation

In this section, the higher order large-deviation approximation in Akahira *et al.* (1999) is summarized as follows. Assume that $X_1, X_2, \dots, X_n, \dots$ is a sequence of independent integer-valued random variables and, for each $j =$

Received December 22, 2000. Revised March 26, 2001. Accepted April 9, 2001.

*Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan.

E-mail: akahira@math.tsukuba.ac.jp

**Institute of Mathematics, University of Tsukuba, Ibaraki 305-8571, Japan.

E-mail: kunihiro@math.tsukuba.ac.jp

$1, 2, \dots, n, \dots, X_j$ is distributed according to a probability function

$$p_j(x) := P\{X_j = x\} \quad \text{for } x = 0, \pm 1, \pm 2, \dots$$

Letting $S_n := \sum_{j=1}^n X_j$, we denote a probability function of S_n by

$$(2.1) \quad p_n^*(y) := P\{S_n = y\}$$

for $y = 0, \pm 1, \pm 2, \dots$. We also denote the moment generating function (m.g.f.) of X_j by

$$(2.2) \quad M_j(\theta) := E[e^{\theta X_j}]$$

for $j = 1, 2, \dots, n, \dots$, assuming that $M_j(\theta)$'s exist for values of θ in an open interval Θ which includes 0. Now, for each j , we consider a discrete exponential family $\mathcal{P}_j := \{p_{j,\theta}(x) : \theta \in \Theta\}$ of probability functions

$$(2.3) \quad p_{j,\theta}(x) := P_\theta\{X_j = x\} = p_j(x)e^{\theta x} M_j(\theta)^{-1}$$

for $x = 0, \pm 1, \pm 2, \dots$, where $p_{j,0}(x) = p_j(x)$. Denote a probability function of S_n by

$$(2.4) \quad p_{n,\theta}^*(y) := P_\theta\{S_n = y\} \quad (y = 0, \pm 1, \pm 2, \dots)$$

for $\theta \in \Theta$, where $p_{n,0}^*(y) = p_n^*(y)$.

From (2.1) to (2.4) we have

$$(2.5) \quad p_{n,\theta}^*(y) = p_n^*(y) e^{\theta y} \prod_{j=1}^n M_j(\theta)^{-1}.$$

On the other hand, it follows from (2.2) that, for each j , the characteristic function of X_j , under the family \mathcal{P}_j , is given by

$$(2.6) \quad E_\theta[e^{itX_j}] = \sum_x e^{itx} p_{j,\theta}(x) = M_j(\theta)^{-1} M_j(\theta + it).$$

Since, by (2.6), the characteristic function of S_n , under the families of $\{\mathcal{P}_j\}$, is

$$E_\theta[e^{itS_n}] = \prod_{j=1}^n M_j(\theta + it) \prod_{j=1}^n M_j(\theta)^{-1},$$

we have by the Fourier inverse transform

$$(2.7) \quad p_{n,\theta}^*(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \prod_{j=1}^n M_j(\theta + it) \prod_{j=1}^n M_j(\theta)^{-1} e^{-ity} dt.$$

From (2.5) and (2.7) we obtain

$$(2.8) \quad p_n^*(y) = e^{-\theta y} \prod_{j=1}^n M_j(\theta) p_{n,\theta}^*(y) \\ = e^{-\theta y} \prod_{j=1}^n M_j(\theta) \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \prod_{j=1}^n M_j(\theta + it) \prod_{j=1}^n M_j(\theta)^{-1} e^{-ity} dt.$$

Letting $K_n(\theta) := \sum_{j=1}^n \log M_j(\theta)$, we have

$$(2.9) \quad \prod_{j=1}^n M_j(\theta) = e^{K_n(\theta)}.$$

From (2.8) and (2.9) we obtain

$$(2.10) \quad p_n^*(y) = e^{K_n(\theta) - \theta y} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{K_n(\theta + it) - K_n(\theta) - ity} dt.$$

For small $|t|$, we have the Taylor expansion

$$K_n(\theta + it) - K_n(\theta) = K_n^{(1)}(\theta)it + \frac{1}{2}K_n^{(2)}(\theta)(it)^2 + \frac{1}{6}K_n^{(3)}(\theta)(it)^3 + \cdots,$$

where $K_n^{(\alpha)}(\theta) := (d^\alpha/d\theta^\alpha)K_n(\theta)$ for $\alpha = 1, 2, \dots$. Then we consider an estimator $\hat{\theta} := \hat{\theta}(S_n)$ for θ such that

$$(2.11) \quad K_n^{(1)}(\hat{\theta}) = y,$$

where $S_n = y$ for $y = 0, \pm 1, \pm 2, \dots$. The following theorems are given in Akahira *et al.* (1999).

THEOREM 1. *If $K_n^{(j)}(\hat{\theta}) = O(n)$ ($j = 2, 3, \dots$), then the probability function $p_n^*(y)$ of the sum S_n is asymptotically given by*

$$(2.12) \quad p_n^*(y) = \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta})}} e^{K_n(\hat{\theta}) - \hat{\theta}y} \left[1 + \frac{K_n^{(4)}(\hat{\theta})}{8\{K_n^{(2)}(\hat{\theta})\}^2} \right. \\ \left. - \frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24\{K_n^{(2)}(\hat{\theta})\}^3} + O\left(\frac{1}{n^2}\right) \right].$$

THEOREM 2. *If $K_n^{(j)}(\hat{\theta}) = O(n)$ ($j = 2, 3, \dots$), then*

$$(2.13) \quad P\{S_n \geq y\} = \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta})}} e^{K_n(\hat{\theta}) - \hat{\theta}y} \sum_{z=0}^{\infty} e^{-\hat{\theta}z} \\ \cdot \left[1 - \frac{z^2}{2K_n^{(2)}(\hat{\theta})} - \frac{K_n^{(3)}(\hat{\theta})z}{2\{K_n^{(2)}(\hat{\theta})\}^2} \right. \\ \left. + \frac{K_n^{(4)}(\hat{\theta})}{8\{K_n^{(2)}(\hat{\theta})\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24\{K_n^{(2)}(\hat{\theta})\}^3} + O\left(\frac{1}{n^2}\right) \right]$$

for all $y > E(S_n)$.

The formula (2.12) is numerically shown to be sufficiently accurate in the binomial and negative-binomial cases, and (2.13) seems to have an advantage in comparison with others, including the saddlepoint approximation. In the next section we try to improve (2.13) to get a more numerically stable and accurate formula for the tail probability than (2.13).

3. A higher order approximation to the tail probability

First, we consider two estimators $\hat{\theta}_0 = \hat{\theta}_0(S_n)$ and $\hat{\theta}_k = \hat{\theta}_k(S_n + k)$ ($k = 1, 2, \dots$) for θ such that

$$(3.1) \quad K_n^{(1)}(\hat{\theta}_0) = y, \quad K_n^{(1)}(\hat{\theta}_k) = y + k,$$

when $S_n = y$ for $y = 0, \pm 1, \pm 2, \dots$. For each $k = 1, 2, \dots$, we put

$$(3.2) \quad \Delta_k := \hat{\theta}_k - \hat{\theta}_0.$$

Then we have the following.

THEOREM 3. *If $K_n^{(j)}(\hat{\theta}_0) = O(n)$ ($j = 2, 3, \dots$), then the upper tail probability of the distribution of S_n is given by*

$$(3.3) \quad P\{S_n \geq y\} = p_n^*(y) \sum_{k=0}^{\infty} \exp \left\{ -k\hat{\theta}_0 - \frac{k^2}{2K_n^{(2)}(\hat{\theta}_0)} - \frac{K_n^{(3)}(\hat{\theta}_0)k}{2\{K_n^{(2)}(\hat{\theta}_0)\}^2} + O\left(\frac{1}{n^2}\right) \right\}$$

for all $y > E(S_n)$, where $p_n^*(y)$ is given by (2.12).

PROOF. From (2.12) we have

$$\begin{aligned} p_n^*(y) &= \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta}_0)}} e^{K_n(\hat{\theta}_0) - y\hat{\theta}_0} \\ &\quad \cdot \left[1 + \frac{K_n^{(4)}(\hat{\theta}_0)}{8\{K_n^{(2)}(\hat{\theta}_0)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_0)\}^2}{24\{K_n^{(2)}(\hat{\theta}_0)\}^3} + O\left(\frac{1}{n^2}\right) \right], \\ p_n^*(y+k) &= \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta}_k)}} e^{K_n(\hat{\theta}_k) - (y+k)\hat{\theta}_k} \\ &\quad \cdot \left[1 + \frac{K_n^{(4)}(\hat{\theta}_k)}{8\{K_n^{(2)}(\hat{\theta}_k)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_k)\}^2}{24\{K_n^{(2)}(\hat{\theta}_k)\}^3} + O\left(\frac{1}{n^2}\right) \right] \end{aligned}$$

for $y = 0, \pm 1, \pm 2, \dots$. Then we obtain

$$\begin{aligned}
 (3.4) \quad & \log(p_n^*(y+k)/p_n^*(y)) \\
 &= K_n(\hat{\theta}_k) - K_n(\hat{\theta}_0) - \{(y+k)\hat{\theta}_k - y\hat{\theta}_0\} \\
 &\quad - \frac{1}{2}\{\log K_n^{(2)}(\hat{\theta}_k) - \log K_n^{(2)}(\hat{\theta}_0)\} \\
 &\quad + \log \left[1 + \frac{K_n^{(4)}(\hat{\theta}_k)}{8\{K_n^{(2)}(\hat{\theta}_k)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_k)\}^2}{24\{K_n^{(2)}(\hat{\theta}_k)\}^3} + O\left(\frac{1}{n^2}\right) \right] \\
 &\quad - \log \left[1 + \frac{K_n^{(4)}(\hat{\theta}_0)}{8\{K_n^{(2)}(\hat{\theta}_0)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_0)\}^2}{24\{K_n^{(2)}(\hat{\theta}_0)\}^3} + O\left(\frac{1}{n^2}\right) \right].
 \end{aligned}$$

Since, from (3.1),

$$\begin{aligned}
 k &= K_n^{(1)}(\hat{\theta}_k) - K_n^{(1)}(\hat{\theta}_0) \\
 &= K_n^{(2)}(\hat{\theta}_0)\Delta_k + \frac{1}{2}K_n^{(3)}(\hat{\theta}_0)\Delta_k^2 + \frac{1}{6}K_n^{(4)}(\hat{\theta}_0)\Delta_k^3 + \dots,
 \end{aligned}$$

it follows that

$$\begin{aligned}
 (3.5) \quad \Delta_k &= \frac{k}{K_n^{(2)}(\hat{\theta}_0)} - \frac{K_n^{(3)}(\hat{\theta}_0)k^2}{2\{K_n^{(2)}(\hat{\theta}_0)\}^3} \\
 &\quad + \left[-\frac{K_n^{(4)}(\hat{\theta}_0)}{6\{K_n^{(2)}(\hat{\theta}_0)\}^4} + \frac{\{K_n^{(3)}(\hat{\theta}_0)\}^2}{2\{K_n^{(2)}(\hat{\theta}_0)\}^5} \right] k^3 + \dots,
 \end{aligned}$$

which implies

$$\begin{aligned}
 (3.6) \quad K_n(\hat{\theta}_k) - K_n(\hat{\theta}_0) &= K_n^{(1)}(\hat{\theta}_0)\Delta_k + \frac{1}{2}K_n^{(2)}(\hat{\theta}_0)\Delta_k^2 + \frac{1}{6}K_n^{(3)}(\hat{\theta}_0)\Delta_k^3 + \dots \\
 &= y\Delta_k + \frac{1}{2}K_n^{(2)}(\hat{\theta}_0)\Delta_k^2 + O\left(\frac{1}{n^2}\right).
 \end{aligned}$$

From (3.2) we have

$$(3.7) \quad k\hat{\theta}_k = k\hat{\theta}_0 + k\Delta_k,$$

and

$$\begin{aligned}
 (3.8) \quad & \frac{1}{2}\{\log K_n^{(2)}(\hat{\theta}_k) - \log K_n^{(2)}(\hat{\theta}_0)\} \\
 &= \frac{1}{2} \left[\frac{K_n^{(3)}(\hat{\theta}_0)}{K_n^{(2)}(\hat{\theta}_0)}\Delta_k + \frac{1}{2} \left\{ \frac{K_n^{(4)}(\hat{\theta}_0)}{K_n^{(2)}(\hat{\theta}_0)} - \left(\frac{K_n^{(3)}(\hat{\theta}_0)}{K_n^{(2)}(\hat{\theta}_0)} \right)^2 \right\} \Delta_k^2 + \dots \right], \\
 (3.9) \quad & \log \left[1 + \frac{K_n^{(4)}(\hat{\theta}_k)}{8\{K_n^{(2)}(\hat{\theta}_k)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_k)\}^2}{24\{K_n^{(2)}(\hat{\theta}_k)\}^3} + O\left(\frac{1}{n^2}\right) \right] \\
 &= \log \left[1 + \frac{K_n^{(4)}(\hat{\theta}_0)}{8\{K_n^{(2)}(\hat{\theta}_0)\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta}_0)\}^2}{24\{K_n^{(2)}(\hat{\theta}_0)\}^3} \right] + \frac{g^{(1)}(\hat{\theta}_0)}{1+g(\hat{\theta}_0)}\Delta_k + O\left(\frac{1}{n^2}\right),
 \end{aligned}$$

where

$$\begin{aligned} g(\theta) &:= \frac{K_n^{(4)}(\theta)}{8\{K_n^{(2)}(\theta)\}^2} - \frac{5\{K_n^{(3)}(\theta)\}^2}{24\{K_n^{(2)}(\theta)\}^3} = O\left(\frac{1}{n}\right), \\ g^{(1)}(\theta) &:= \frac{dg(\theta)}{d\theta} \\ &= \frac{K_n^{(5)}(\theta)}{8\{K_n^{(2)}(\theta)\}^2} - \frac{K_n^{(3)}(\theta)K_n^{(4)}(\theta)}{4\{K_n^{(2)}(\theta)\}^3} - \frac{5K_n^{(3)}(\theta)K_n^{(4)}(\theta)}{12\{K_n^{(2)}(\theta)\}^3} + \frac{5\{K_n^{(3)}(\theta)\}^3}{8\{K_n^{(2)}(\theta)\}^4} \\ &= O\left(\frac{1}{n}\right). \end{aligned}$$

From (3.4) to (3.9) we obtain

$$\log \frac{p_n^*(y+k)}{p_n^*(y)} = -k\hat{\theta}_0 - k\Delta_k + \frac{1}{2}K_n^{(2)}(\hat{\theta}_0)\Delta_k^2 - \frac{K_n^{(3)}(\hat{\theta}_0)}{2K_n^{(2)}(\hat{\theta}_0)}\Delta_k + O\left(\frac{1}{n^2}\right),$$

hence

$$(3.10) \quad p_n^*(y+k) = p_n^*(y) \exp \left\{ -k\hat{\theta}_0 - k\Delta_k + \frac{1}{2}K_n^{(2)}(\hat{\theta}_0)\Delta_k^2 - \frac{K_n^{(3)}(\hat{\theta}_0)}{2\{K_n^{(2)}(\hat{\theta}_0)\}}\Delta_k + O\left(\frac{1}{n^2}\right) \right\}.$$

Since, by (3.5),

$$\Delta_k = \frac{k}{K_n^{(2)}(\hat{\theta}_0)} + O\left(\frac{1}{n^2}\right),$$

it follows from (3.10) that

$$p_n^*(y+k) = p_n^*(y) \exp \left\{ -k\hat{\theta}_0 - \frac{k^2}{2K_n^{(2)}(\hat{\theta}_0)} - \frac{K_n^{(3)}(\hat{\theta}_0)k}{2\{K_n^{(2)}(\hat{\theta}_0)\}^2} + O\left(\frac{1}{n^2}\right) \right\},$$

which implies (3.3) from $P\{S_n \geq y\} = \sum_{k=0}^{\infty} p_n^*(y+k)$. \square

In a similar way to the above we have a higher order approximation to the lower tail probability of the distribution of S_n . Indeed, we consider two estimators $\hat{\theta}_0 = \hat{\theta}_0(S_n)$ and $\tilde{\theta}_k = \tilde{\theta}_k(S_n - k)$ ($k = 1, 2, \dots$) for θ such that

$$K_n^{(1)}(\hat{\theta}_0) = y, \quad K_n^{(1)}(\tilde{\theta}_k) = y - k,$$

when $S_n = y$ for $y = 0, \pm 1, \pm 2, \dots$. Then we have

$$p_n^*(y-k) = p_n^*(y) \exp \left\{ k\hat{\theta}_0 - \frac{k^2}{2K_n^{(2)}(\hat{\theta}_0)} + \frac{K_n^{(3)}(\hat{\theta}_0)k}{2\{K_n^{(2)}(\hat{\theta}_0)\}^2} + O\left(\frac{1}{n^2}\right) \right\}.$$

COROLLARY. If $K_n^{(j)}(\hat{\theta}_0) = O(n)$ ($j = 2, 3, \dots$), then the lower tail probability of the distribution of S_n is given by

$$(3.11) \quad P\{S_n \leq y\} = p_n^*(y) \sum_{k=0}^{\infty} \exp \left\{ k\hat{\theta}_0 - \frac{k^2}{2K_n^{(2)}(\hat{\theta}_0)} + \frac{K_n^{(3)}(\hat{\theta}_0)k}{2\{K_n^{(2)}(\hat{\theta}_0)\}^2} + O\left(\frac{1}{n^2}\right) \right\}$$

for all $y < E(S_n)$, where $p_n^*(y)$ is given by (2.12).

4. Binomial cases

Suppose that $X_1, X_2, \dots, X_n, \dots$ is a sequence of independent random variables and, for each $j = 1, 2, \dots, n, \dots$, X_j is distributed as the binomial distribution $B(1, p_j)$ with a probability function

$$f_{X_j}(x) = P\{X_j = x\} = p_j^x q_j^{1-x}$$

for $x = 0, 1$, where $0 < p_j < 1$ and $q_j = 1 - p_j$. Let $S_n := \sum_{j=1}^n X_j$. Then we consider the large-deviation approximations (2.13), (3.3), the Edgeworth expansion and the Lugannani and Rice approximation to the upper tail probability of the distribution of S_n . First, in a similar way to the binomial case in Akahira *et al.* (1999), from (2.11) we take $\hat{\theta} = \hat{\theta}(S_n)$ such that

$$K_n^{(1)}(\hat{\theta}) = \sum_{j=1}^n \frac{p_j e^{\hat{\theta}}}{p_j e^{\hat{\theta}} + q_j} = y,$$

where $S_n = y$ for $y = 0, \pm 1, \pm 2, \dots$, hence

$$(4.1) \quad \begin{aligned} K_n^{(2)}(\hat{\theta}) &= \sum_{j=1}^n \hat{p}_j \hat{q}_j, & K_n^{(3)}(\hat{\theta}) &= \sum_{j=1}^n \hat{p}_j \hat{q}_j (\hat{q}_j - \hat{p}_j), \\ K_n^{(4)}(\hat{\theta}) &= \sum_{j=1}^n \hat{p}_j \hat{q}_j (1 - 6\hat{p}_j \hat{q}_j), \end{aligned}$$

where $\hat{p}_j = p_j e^{\hat{\theta}} / (p_j e^{\hat{\theta}} + q_j)$ and $\hat{q}_j = 1 - \hat{p}_j$ ($j = 1, \dots, n$). From (4.1) we can calculate the large-deviation approximations (2.13) and (3.3) to the upper tail probability of the distribution of S_n . Second, the cumulants of S_n are given by

$$\begin{aligned} \mu_n &:= E(S_n) = \sum_{j=1}^n p_j, & v_n &:= V(S_n) = \sum_{j=1}^n p_j q_j, \\ \kappa_{3,n} &:= \kappa_3(S_n) = \sum_{j=1}^n p_j q_j (q_j - p_j), \\ \kappa_{4,n} &:= \kappa_4(S_n) = \sum_{j=1}^n p_j q_j (1 - 6p_j q_j), \end{aligned}$$

Table 1. The exact values of the tail probability $P\{S_n \geq y\}$ of S_n , and the relative errors of its E-approx. (4.3) and its second order LD ones (2.13) and (3.3) when (i) $\{p_j\}_{j=1}^{19} = 0.05(0.05)0.95$, (ii) $\{p_j\}_{j=1}^{20} = 0.03(0.03)0.60$, and (iii) $\{p_j\}_{j=1}^{20}$ is uniformly distributed on the interval $(0, 1)$, that is p_1, \dots, p_{20} are given by 0.305146, 0.715095, 0.612101, 0.672283, 0.447648, 0.268358, 0.434328, 0.552620, 0.608603, 0.130255, 0.941095, 0.141198, 0.164085, 0.693920, 0.565611, 0.977985, 0.0513902, 0.877854, 0.451323, 0.0628465, respectively (see Tables 3.4, 3.5 and 3.6 in Akahira *et al.* (1999)).

(i) $\{p_j\}_{j=1}^{19} = 0.05(0.05)0.95$				
y	Exact(%)	E-approx.	LD (2.13)	LD (3.3)
10	50.0000	0.0000	—	-0.0029
11	29.0651	-0.0000	-0.8356	0.0001
12	13.4452	0.0000	-0.2010	0.0009
13	4.7942	0.0004	-0.0627	0.0009
14	1.2761	0.0014	-0.0232	0.0007
15	0.2442	-0.0006	-0.0097	0.0003
16	0.0320	-0.0293	-0.0048	-0.0003
17	0.0027	0.2154	-0.0038	-0.0018
18	0.0001	—	-0.0083	-0.0077
19	0.0000	—	—	—

(ii) $\{p_j\}_{j=1}^{20} = 0.03(0.03)0.60$				
y	Exact(%)	E-approx.	LD (2.13)	LD (3.3)
7	45.0120	0.0002	—	-0.0213
8	26.2247	0.0003	-0.9933	-0.0095
9	12.7076	-0.0003	-0.2467	-0.0043
10	5.0524	-0.0032	-0.0840	-0.0020
11	1.6288	-0.0067	-0.0346	-0.0009
12	0.4207	0.0005	-0.0161	-0.0003
13	0.0859	0.0451	-0.0082	-0.0001
14	0.0136	0.1758	-0.0044	-0.0001
15	0.0016	0.4772	-0.0026	-0.0001
16	0.0001	1.1191	-0.0018	-0.0003
17	0.0000	2.5290	-0.0017	-0.0008
18	0.0000	6.0939	-0.0025	-0.0020
19	0.0000	18.1006	-0.0076	-0.0073
20	0.0000	91.0675	—	—

(iii) $\{p_j\}_{j=1}^{20}$ is uniformly distributed on the interval $(0, 1)$				
y	Exact(%)	E-approx.	LD (2.13)	LD (3.3)
10	53.7097	-0.0002	—	-0.0059
11	32.6462	-0.0003	—	-0.0010
12	16.0033	0.0001	-0.2793	0.0005
13	6.1365	0.0012	-0.0836	0.0008
14	1.7877	0.0023	-0.0306	0.0006
15	0.3836	-0.0014	-0.0129	0.0003
16	0.0585	-0.0307	-0.0061	-0.0000
17	0.0060	0.1676	-0.0036	-0.0006
18	0.0004	-0.7757	-0.0032	-0.0018
19	0.0000	—	-0.0067	-0.0062
20	0.0000	—	—	—

Table 2. The exact values of the tail probabilities $P\{S_n \geq y\}$ for $y > E(S_n)$ and $P\{S_n \leq y\}$ for $y < E(S_n)$, and the relative errors of its normal approx. the saddlepoint approx. with the first term only, the saddlepoint (sp.) expansion, the L-R approx., and the LD approx. (3.3) and (3.11) when $p_j = p = 0.15$ ($j = 1, \dots, n$) for $n = 10, 20$ and $p_j = p = 0.5$ ($j = 1, \dots, n$) for $n = 10$. The values except the LD approx., are referred from Table 2.4.4 of Jensen (1995, page 44).

(i) $n = 10, p = 0.15$						
y	1	2	4	5	8	9
Exact(%)	54.43	45.57	5.00	0.987	8.67×10^{-4}	3.33×10^{-5}
Normal	-0.081	0.097	-0.234	-0.601	-0.994	-1.000
Saddlepoint	-0.251	-0.137	-0.050	-0.035	-0.007	0.021
Sp.-expansion	0.050	0.004	0.004	0.005	0.010	0.024
Lugannani-Rice	0.051	0.004	0.008	0.013	0.038	0.081
LD (3.3), (3.11)	-0.008	-0.069	-0.008	-0.003	-0.002	-0.006

(ii) $n = 20, p = 0.15$						
y	1	2	6	8	18	19
Exact(%)	17.56	40.49	6.73	0.592	2.07×10^{-11}	3.80×10^{-13}
Normal	-0.010	-0.069	-0.128	-0.591	-1.000	-1.000
Saddlepoint	-0.150	-0.196	-0.043	-0.027	0.010	0.039
Sp.-expansion	0.055	0.021	0.001	0.002	0.005	0.021
Lugannani-Rice	0.065	0.023	0.003	0.005	0.039	0.084
LD (3.3), (3.11)	-0.006	0.018	-0.010	-0.003	-0.004	-0.006

(iii) $n = 10, p = 0.5$				
y	1	2	3	4
Exact(%)	1.07	5.47	17.19	37.70
Normal	0.252	0.040	-0.003	-0.003
Saddlepoint	-0.047	-0.090	-0.116	-0.149
Sp.-expansion	0.047	0.020	0.013	0.008
Lugannani-Rice	0.075	0.032	0.017	0.009
LD (3.11)	-0.007	0.000	0.001	-0.003

it follows that the Edgeworth expansion of the distribution of S_n is given by

$$\begin{aligned}
 (4.2) \quad P\{S_n = t\} &= P\left\{\frac{S_n - \mu_n}{\sqrt{v_n}} = \frac{t - \mu_n}{\sqrt{v_n}}\right\} \\
 &= \frac{1}{\sqrt{v_n}} \phi(y) \left\{ 1 + \frac{\kappa_{3,n}}{6v_n^{3/2}}(y^3 - 3y) + \frac{\kappa_{4,n}}{24v_n^2}(y^4 - 6y^2 + 3) \right. \\
 &\quad \left. + \frac{\kappa_{3,n}^2}{72v_n^3}(y^6 - 15y^4 + 45y^2 - 15) \right\} + o\left(\frac{1}{n\sqrt{n}}\right),
 \end{aligned}$$

where $y := (t - \mu_n)/\sqrt{v_n}$. From (4.2) we have

$$(4.3) \quad P\{S_n \geq y\} = 1 - \Phi(z) + \phi(z) \cdot \left\{ \frac{\kappa_{3,n}}{6v_n^{3/2}}(z^2 - 1) + \frac{\kappa_{4,n}}{24v_n^2}(z^3 - 3z) + \frac{\kappa_{3,n}^2}{72v_n^2}(z^5 - 10z^3 + 15z) - \frac{1}{24v_n}z + o\left(\frac{1}{n}\right) \right\},$$

where $z = (y - 0.5 - \mu_n)/\sqrt{v_n}$.

Third, the Lugannani and Rice approximation to the upper tail probability of the distribution of S_n based on the i.i.d. sample (X_1, \dots, X_n) is given by

$$(4.4) \quad P\{S_n \geq y\} = \{1 - \Phi(r)\} \{1 + O(n^{-2})\} + \phi(r) \left\{ \frac{1}{\tilde{\lambda}} - \frac{1}{r} + O(n^{-3/2}) \right\}$$

where

$$r^2 = 2n \left\{ \frac{y}{n} \log \left(\frac{(y/n)(1-p)}{(1-(y/n))p} \right) - \log \frac{1-p}{1-(y/n)} \right\},$$

$$\tilde{\lambda} = \sqrt{n} \left\{ 1 - \frac{(1-(y/n))p}{(y/n)(1-p)} \right\} \sqrt{\frac{y}{n} \left(1 - \frac{y}{n} \right)}.$$

Now, we compare the Edgeworth approximation (E-approx.) (4.3), the second order large-deviation (LD) ones (2.13) and (3.3) to the upper tail probability of the distribution of S_n . As the numerical result, the LD approx. (3.3) is more accurate than the LD one (2.13), and (3.3) is not so worse than the E-approx. even in the central part of the distribution of S_n (see Table 1). Further, comparing the LD approx. (3.3) and (3.11) with some approximations in Jensen (1995), we see that (3.3) and (3.11) seems to be more accurate than the others in the numerical calculation (see Table 2).

REFERENCES

- Akahira, M. and Takeuchi, K. (1981). *Asymptotic Efficiency of Statistical Estimators: Concepts and Higher Order Asymptotic Efficiency*, Lecture Notes in Statistics, **7**, Springer, New York.
- Akahira, M., Takahashi, K. and Takeuchi, K. (1999). The higher order large-deviation approximation for the distribution of the sum of independent discrete random variables, *Commun. Statist.-Theory Meth.*, **28**(3&4), 706-726.
- Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*, Regional Conference Series in Applied Mathematics, SIAM, Philadelphia.
- Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*, Chapman and Hall, London.
- Bucklew, J. A. (1990). *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley, New York.
- Daniels, H. E. (1954). Saddlepoint approximations in statistics, *Ann. Math. Statist.*, **25**, 631-650.
- Daniels, H. E. (1987). Tail probability approximations, *Int. Statist. Review*, **55**, 37-48.
- Ghosh, J. K. (1994). *Higher Order Asymptotics*, NSF-CBMS Regional Conference Series, Probability and Statistics, **4**, IMS, Hayward, California.

- Jensen, J. L. (1995). *Saddlepoint Approximations*, Clarendon Press, Oxford.
- Lugannani, R. and Rice, S. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables, *Adv. Appl. Prob.*, **12**, 475–490.
- Pfanzagl, J. and Wefelmeyer, W. (1985). *Asymptotic Expansions for General Statistical Models*, Lecture Notes in Statistics, **31**, Springer, Berlin.