数理解析研究所講究録
1055 巻 1998 年 45-67

Large-deviation approximations for the discrete distribution

筑波大・数学　高橋 邦彦 (Kunihiko Takahashi)
筑波大・数学　赤平 昌文 (Masafumi Akahira)

1. はじめに

統計的推測の高次漸近理論では Edgeworth(展開による) 近似がよく用いられ, 推定量, 検定などの高次の漸近的性質が論じられてきた. 実際, 推定量の真の母数の周りでの集中確率を漸近的に求めるときには漸近分布の中心部分を必要とし, そのとき Edgeworth 近似が重要な役割を果たした ([BC89]). 一方, Bahadur 効率は漸近分布の裾の部分に基づいていて, 大偏差確率を必要とする. また, 裾確率の近似についてもいろいろ論じられている ([D87], [J95]).

本論では, 独立な離散型確率変数の和の分布に対する大偏差 (原理による) 近似を高次の次数まで求めて, 通常用いられている正規近似, Edgeworth 近似と比較検討する ([ATT98]).

2. 大偏差近似

$X_1, \cdots, X_n, \cdots$ を互いに独立に, 各 $j = 1, 2, \cdots$ について X_j が確率関数

$$p_j(x) = P\{X_j = x\} \qquad (x = 0, \pm 1, \pm 2, \cdots)$$

に従う離散型確率変数列とする. それらの和 $S := \sum_{j=1}^{n} X_j$ の確率関数を

$$p_n^*(y) := P\{S_n = y\} \qquad (y = 0, \pm 1, \pm 2, \cdots)$$

とする. また各 j について, X_j の積率母関数を $M_j(\theta) := E\left[e^{\theta X_j}\right]$ $(\theta \in \Theta)$ とする. ただし Θ は原点を含むある開区間とする. さて各 j について, 確率関数

$$p_{j,\theta}(x) := P_\theta\left\{X_j = x\right\} = p_j(x)e^{\theta x}M_j(\theta)^{-1} \qquad (x = 0, \pm 1, \pm 2, \cdots)$$

の離散指数型分布族 $\mathcal{P}_j := \{p_{j,\theta}(x) : \theta \in \Theta\}$ を考える. ここで $p_{j,0}(x) = p_j(x)$ になる. また S_n の確率関数を

$$p_{n,\theta}^*(y) := P_\theta\left\{S_n = y\right\} \qquad (y = 0, \pm 1, \pm 2, \cdots)$$

で表わす．ただし $p^*_{n,0}(y)=p^*_n(y)$ とする．このとき

$$p^*_{n,\theta}(y) = p^*_n(y)e^{\theta y}\prod_{j=1}^{n} M_j(\theta)^{-1} \tag{2.1}$$

となる．また各 j について，$\mathcal{P}_j$の下で，X_jの特性関数は

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

になる．従って，$\{\mathcal{P}_j\}$ の下で，S_nの特性関数は

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

となるから，フーリエ逆変換を用いて

$$p^*_{n,\theta}(y) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{n} M_j(\theta+it)\prod_{j=1}^{n} M_j(\theta)^{-1}e^{-ity}dt \tag{2.2}$$

になる．そこで (2.1), (2.2) から

$$p^*_n(y) = e^{-\theta y}\prod_{j=1}^{n} M_j(\theta)\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi}\prod_{j=1}^{n} M_j(\theta+it)\prod_{j=1}^{n} M_j(\theta)^{-1}e^{-ity}dt$$

を得る．ただし i は虚数単位とする．また $K_n(\theta) := \sum_{j=1}^{n}\log M_j(\theta)$ とおくと

$$p^*_n(y) = e^{K_n(\theta)-\theta y}\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{K_n(\theta+it)-K_n(\theta)-ity}dt \tag{2.3}$$

になる．さらに Taylor 展開によって，十分に小さい $|t|$ に対して

$$K_n(\theta+it) - K_n(\theta) = K_n^{(1)}(\theta)it + \frac{1}{2}K_n^{(2)}(it)^2 + \frac{1}{6}K_n^{(3)}(\theta)(it)^3 + \frac{1}{24}K_n^{(4)}(\theta)(it)^4 + \cdots \tag{2.4}$$

になる．ただし各 $\alpha = 1, 2, \cdots$ について

$$K_n^{(\alpha)}(\theta) = \left(d^\alpha/d\theta^\alpha\right)K_n(\theta)$$

とする．このとき，$S_n = y$ $(y = 0, \pm 1, \pm 2, \cdots)$ に対して $K_n^{(1)}(\hat\theta) = y$ となる θ の推定量 $\hat\theta := \hat\theta(S_n)$ を考える．ここで $K_n^{(\alpha)}(\hat\theta) = O(n)$ $(\alpha = 2, 3, \cdots)$ を仮定する．

定理 1. S_n の確率関数 $p^*_n(y)$ は漸近的に次のように与えられる．

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

証明については,

$$\int_{-\pi}^{\pi} e^{K_n(\hat{\theta}+it)-K_n(\hat{\theta})-ity}dt$$
$$= \frac{1}{\sqrt{K_n^{(2)}(\hat{\theta})}} \int_{-\pi\sqrt{K_n^{(2)}(\hat{\theta})}}^{\pi\sqrt{K_n^{(2)}(\hat{\theta})}} \exp\left\{K_n\left(\hat{\theta}+\frac{iz}{\sqrt{K_n^{(2)}(\hat{\theta})}}\right) - K_n(\hat{\theta}) - \frac{iyz}{\sqrt{K_n^{(2)}(\hat{\theta})}}\right\}dz$$

となるから, (2.4) と$\hat{\theta}$の定義を用いて計算すれば結論を得る.

各 $j=1,2,\cdots$ について, 確率関数 $p_{j,\theta}(\cdot)$ と $p_j(\cdot)$ の間の Kullback-Leibler 情報量を

$$I_j(\theta,0) := \sum_x p_{j,\theta}(x)\log\frac{p_{j,\theta}(x)}{p_j(x)}$$

で定義する. このとき確率関数 $p^*_{j,\theta}(\cdot)$ と $p^*_j(\cdot)$ の間の Kullback-Leibler 情報量は

$$I^*_n(\theta,0) = \sum_{j=1}^{n} I_j(\theta,0) = \theta K_n^{(1)}(\theta) - K_n(\theta)$$

となる. 従って定理 1 から次の系を得る.

系. S_n の確率関数 $p^*_n(y)$ は漸近的に次のように与えられる.

$$p^*_n(y) = \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta})}}e^{-I^*_n(\hat{\theta},0)}\left[1+\frac{K_n^{(4)}(\hat{\theta})}{8\{K_n^{(2)}(\hat{\theta})\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24\{K_n^{(2)}(\hat{\theta})\}^3} + O\left(\frac{1}{n^2}\right)\right].$$

定理 2. S_n の裾確率は漸近的に次のように与えられる.

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

証明については, (2.3) から任意の $z \geq 0$ に対して

$$p^*_n(y+z) = e^{K_n(\hat{\theta})-\hat{\theta}(y+z)} \cdot \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{K_n(\hat{\theta}+it)-K_n(\hat{\theta})-it(y+z)}dt$$

となるから, 定理 1 の証明と同様に変形して計算すれば結論を得る.

定理2の結果は $y-E(S_n)=O(n)$ のとき意味を持つことに注意したい．しかしこのことは，通常，鞍部点近似 (saddlepoint approximation) として $y-E(S_n)=o(n)$ のときに正規分布に基づいて得た結果 (例えば Lugannani and Rice [LR80], Robinson [R82]) とは異なっている．

3. 二項分布の場合

$X_1,\cdots,X_n,\cdots$ を互いに独立に，各 $j=1,2,\cdots$ について X_j が二項分布 $B(1,p_j)$ に従う確率変数とする．ただし $0<p_j<1$ とし，また $q_j=1-p_j$ $(j=1,2,\cdots)$ とする．このとき 各 $y=0,1,\cdots,n$ について

$$K_n^{(1)}(\theta)=\sum_{j=1}^{n}\frac{p_je^{\theta}}{p_je^{\theta}+q_j}=y$$

となる θ を $\hat{\theta}$ とする．このとき $\hat{\tau}:=e^{\hat{\theta}}$ とおいて，各 j について $\hat{p}_j:=p_j\hat{\tau}/(p_j\hat{\tau}+q_j)$, $\hat{q}_j:=1-\hat{p}_j$ とすれば，$K_n^{(2)}(\hat{\theta})=\sum_{j=1}^n\hat{p}_j\hat{q}_j$, $K_n^{(3)}(\hat{\theta})=\sum_{j=1}^n\hat{p}_j\hat{q}_j(\hat{q}_j-\hat{p}_j)$, $K_n^{(4)}(\hat{\theta})=\sum_{j=1}^n\hat{p}_j\hat{q}_j(1-6\hat{p}_j\hat{q}_j)$ になる．従って定理1から大偏差近似

$$p_n^*(y)=\frac{1}{\sqrt{2\pi\sum_{j=1}^n\hat{p}_j\hat{q}_j}}\left(\prod_{j=1}^n\frac{q_j}{\hat{q}_j}\right)\hat{\tau}^{-y}$$
$$\cdot\left[1+\frac{K_n^{(4)}(\hat{\theta})}{8\left(\sum_{j=1}^n\hat{p}_j\hat{q}_j\right)^2}-\frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24\left(\sum_{j=1}^n\hat{p}_j\hat{q}_j\right)^3}+O\left(\frac{1}{n^2}\right)\right]$$

を得る．ここで上式の右辺の第1項目までの近似を1次の大偏差近似として LD_1 と表わし，右辺すべての項による近似を2次の大偏差近似として LD_2 で表わす．

一方，S_n のキュムラントは

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

となるから, S_nの分布の Edgeworth 展開は

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

になる. ただし $y := (t-\mu_n)\big/\sqrt{v_n}$, $\phi(y) = \left(1\big/\sqrt{2\pi}\right)e^{-y^2/2}$ とする. 上式の右辺の第1項目までの近似を正規近似といい, 右辺のすべての項による近似を Edgeworth 近似という. このとき大偏差近似 LD_1, LD_2はそれぞれ正規近似, Edgeworth 近似より端において精確であることが分かる. 実際, 次の場合について数値計算を行って比較した. (i) $\{p_j\}_{j=1}^{9} = 0.1(0.1)0.9$, (ii) $\{p_j\}_{j=1}^{19} = 0.05(0.05)0.95$, (iii) $\{p_j\}_{j=1}^{10} = 0.05(0.05)0.50$, (iv) $\{p_j\}_{j=1}^{20} = 0.03(0.03)0.60$, (v) $\{p_j\}_{j=1}^{n}$ が区間 $(0,1)$ 上の一様乱数 $(n=10,20)$, (vi) $\{p_j\}_{j=1}^{n}$ が区間 $(0.05, 0.95)$ 上の一様乱数 $(n=10,20)$ (表 3.1 〜 3.8 参照).

また定理 2 より S_nの裾確率は $y > E(S_n)$ に対して

$$
\begin{aligned}
&P\{S_n \geq y\} \\
&= \frac{1}{\sqrt{2\pi\sum_{j=1}^{n}\hat{p}_j\hat{q}_j}}\left(\prod_{j=1}^{n}\frac{q_j}{\hat{q}_j}\right)\hat{\tau}^{-y} \\
&\quad \cdot\left[\frac{1-\hat{\tau}^{-(m+1)}}{1-\hat{\tau}^{-1}}\left\{1 + \frac{K_n^{(4)}(\hat{\theta})}{8(\sum_{j=1}^{n}\hat{p}_j\hat{q}_j)^2} - \frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24(\sum_{j=1}^{n}\hat{p}_j\hat{q}_j)^3}\right\}\right. \\
&\quad - \frac{1}{2\sum_{j=1}^{n}\hat{p}_j\hat{q}_j}\left\{-\frac{(m+1)^2\hat{\tau}^{-(m+1)}}{1-\hat{\tau}^{-1}}\right. \\
&\quad \left. + \frac{1}{(1-\hat{\tau}^{-1})^3}\left(-(2m+3)\hat{\tau}^{-(m+2)} + (2m+1)\hat{\tau}^{-(m+3)} + \hat{\tau}^{-1}(1+\hat{\tau}^{-1})\right)\right\} \\
&\quad \left. - \frac{K_n^{(3)}(\hat{\theta})}{2(\sum_{j=1}^{n}\hat{p}_j\hat{q}_j)^2}\left\{-\frac{(m+1)\hat{\tau}^{-(m+1)}}{1-\hat{\tau}^{-1}} + \frac{(1-\hat{\tau}^{-(m+1)})\hat{\tau}^{-1}}{(1-\hat{\tau}^{-1})^2}\right\} + O\left(\frac{1}{n^2}\right)\right]
\end{aligned}
$$

となる. ただし $m = n - y$ とし,

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

とする. またこの裾確率の 2 次の大偏差近似を LD_2で表わす.

一方, Edgeworth 展開を用いると, その裾確率の連続補正した Edgeworth 近似

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

を得る. ただし $z = (y-0.5-\mu_n)\big/\sqrt{v_n}$ とし $\Phi(y) = \int_{-\infty}^{y}\phi(x)dx$ とする. この場合にも 2 次の大偏差近似 LD_2は Edgeworth 近似より精確であることが分かる (表 3.9 ～ 3.16 参照).

表 3.1. $\{p_j\}_{j=1}^{9} = 0.1(0.1)0.9$ のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth(%)	LD_1(%)	LD_2(%)
0	0.0363	0.0309	0.0021	0.0000	0.0000
1	0.7000	0.0587	0.0035	0.0484	−0.0067
2	4.8208	−0.1473	−0.0127	0.1420	−0.0123
3	15.9749	−0.2692	0.0102	0.3096	−0.0144
4	28.4580	0.3237	−0.0021	0.4615	−0.0138
5	28.4680	0.3237	−0.0021	0.4615	−0.0138
6	15.9749	−0.2692	0.0102	0.3096	−0.0144
7	4.8208	−0.1473	−0.0127	0.1420	−0.0123
8	0.7000	0.0587	0.0035	0.0484	−0.0067
9	0.0363	0.0309	0.0021	0.0000	0.0000

表 3.2. $\{p_j\}_{j=1}^{19} = 0.05(0.05)0.95$ のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0000	11.0363	—	0.0000	0.0000
1	0.0001	2.4706	−0.5734	0.0742	−0.0080
2	0.0025	0.8305	−0.0435	0.0321	−0.0024
3	0.0293	0.2991	0.0076	0.0196	−0.0010
4	0.2122	0.0905	0.0040	0.0141	−0.0005
5	1.0319	0.0091	−0.0000	0.0112	−0.0002
6	3.5182	−0.0144	−0.0006	0.0095	−0.0002
7	8.6510	−0.0120	−0.0000	0.0085	−0.0001
8	15.6199	−0.0140	0.0001	0.0079	−0.0001
9	20.9349	0.0065	−0.0000	0.0076	−0.0001
10	20.9349	0.0065	−0.0000	0.0076	−0.0001
11	15.6199	−0.0140	0.0001	0.0079	−0.0001
12	8.6510	−0.0120	−0.0000	0.0085	−0.0001
13	3.5182	−0.0144	−0.0006	0.0095	−0.0002
14	1.0319	0.0091	−0.0000	0.0112	−0.0002
15	0.2122	0.0905	0.0040	0.0141	−0.0005
16	0.0293	0.2991	0.0076	0.0196	−0.0010
17	0.0025	0.8305	−0.0435	0.0321	−0.0024
18	0.0001	2.4706	−0.5734	0.0742	−0.0080
19	0.0000	11.0363	—	0.0000	0.0000

表 3.3. $\{p_j\}_{j=1}^{10} = 0.05(0.05)0.50$ のときの S_nの分布 $P\{S_n = y\}$ の値と，その値の正規近似，Edgeworth 近似，1 次の大偏差近似 LD_1，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	3.2737	0.0991	−0.0130	0.0000	0.0000
1	14.3236	−0.1155	−0.0122	0.0835	−0.0061
2	26.7408	−0.0466	0.0073	0.0410	−0.0014
3	27.9306	0.0498	−0.0037	0.0272	−0.0008
4	17.9848	0.0717	0.0016	0.0210	−0.0006
5	7.4151	−0.0235	0.0085	0.0185	−0.0005
6	1.9680	−0.2100	−0.0028	0.0186	−0.0006
7	0.3289	−0.4200	−0.0022	0.0220	−0.0011
8	0.0327	−0.5912	0.2494	0.0330	−0.0025
9	0.0017	−0.6885	1.1587	0.0739	−0.0084
10	0.0000	−0.6533	4.6822	0.0000	0.0000

表 3.4. $\{p_j\}_{j=1}^{20} = 0.03(0.03)0.60$ のときの S_nの分布 $P\{S_n = y\}$ の値と，その値の正規近似，Edgeworth 近似，1 次の大偏差近似 LD_1，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0263	2.7822	0.0933	0.0000	0.0000
1	0.2970	0.5926	0.0512	0.0840	−0.0059
2	1.5463	0.1126	0.0032	0.0416	−0.0013
3	4.9248	−0.0290	−0.0058	0.0273	−0.0006
4	10.7461	−0.0548	−0.0014	0.0201	−0.0003
5	17.0540	−0.0334	0.0018	0.0159	−0.0002
6	20.3938	0.0024	0.0005	0.0132	−0.0002
7	18.7873	0.0312	−0.0014	0.0114	−0.0002
8	13.5171	0.0378	−0.0002	0.0102	−0.0001
9	7.6552	0.0139	0.0025	0.0094	−0.0001
10	3.4236	−0.0416	0.0011	0.0089	−0.0001
11	1.2081	−0.1226	−0.0071	0.0087	−0.0001
12	0.3348	−0.2185	−0.0131	0.0089	−0.0001
13	0.0722	−0.3169	0.0087	0.0094	−0.0001
14	0.0120	−0.4057	0.1017	0.0105	−0.0002
15	0.0015	−0.4748	0.3331	0.0122	−0.0003
16	0.0001	−0.5146	0.8263	0.0152	−0.0005
17	0.0000	−0.5092	1.8805	0.0210	−0.0010
18	0.0000	−0.4175	4.4459	0.0339	−0.0022
19	0.0000	−0.0774	12.7190	0.0764	−0.0075
20	0.0000	1.6449	60.6571	0.0000	0.0000

表 3.5. $\{p_j\}_{j=1}^{10}$が区間$(0,1)$上の一様乱数$(p_1,\cdots,p_{10}$ =0.525118, 0.351263,0.000803728, 0.0528506, 0.972966, 0.682429,0.245028,0.219989, 0.812036, 0.812036, 0.569428) のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0119	3.3322	0.1245	0.0000	0.0000
1	0.5499	0.2372	0.0287	0.0055	−0.0228
2	4.7866	−0.0287	−0.0063	0.0295	0.0059
3	16.9828	−0.0287	0.0005	0.0244	0.0008
4	30.0907	0.0098	0.0007	0.0196	−0.0002
5	28.5760	0.0169	−0.0010	0.0178	−0.0010
6	14.6380	−0.0144	0.0016	0.0217	0.0002
7	3.8805	−0.0416	−0.0033	0.0284	0.0018
8	0.4671	0.0657	0.0011	0.0365	−0.0115
9	0.0165	1.0968	0.1420	0.0450	0.2980
10	0.0000	95.2134	——	0.0000	0.0000

表 3.6. $\{p_j\}_{j=1}^{20}$が区間$(0,1)$上の一様乱数$(p_1,\cdots,p_{10} =$ 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) のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0000	36.8118	——	0.0000	0.0000
1	0.0001	5.5998	——	0.0564	−0.0122
2	0.0015	1.5567	−0.0944	0.0186	−0.0036
3	0.0210	0.5223	0.0311	0.0123	−0.0012
4	0.1670	0.1642	0.0138	0.0119	−0.0000
5	0.8584	0.0291	0.0001	0.0120	0.0003
6	3.0390	−0.0142	−0.0025	0.0117	0.0003
7	7.7151	−0.0172	−0.0007	0.0110	0.0003
8	14.4053	−0.0057	0.0007	0.0102	0.0002
9	20.0831	0.0055	0.0006	0.0093	0.0001
10	21.0635	0.0087	−0.0002	0.0085	0.0000
11	16.6429	0.0025	−0.0006	0.0077	−0.0000
12	9.8669	−0.0091	−0.0003	0.0072	−0.0002
13	4.3488	−0.0167	0.0006	0.0070	−0.0003
14	1.4041	−0.0059	0.0019	0.0074	−0.0004
15	0.3251	0.0457	0.0024	0.0091	−0.0005
16	0.0524	0.1788	−0.0065	0.0125	−0.0007
17	0.0056	0.4856	−0.0668	0.0192	−0.0011
18	0.0004	1.2369	−0.3813	0.0338	−0.0022
19	0.0000	3.5355	——	0.0789	−0.0064
20	0.0000	15.4703	——	0.0000	0.0000

表 3.7. $\{p_j\}_{j=1}^{10}$が区間 $(0.05, 0.95)$ 上の一様乱数 $(p_1, \cdots, p_{10}$ = 0.120824, 0.58353, 0.441495, 0.17778, 0.447765, 0.185215, 0.197315, 0.484553, 0.206519, 0.747038) のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.6283	0.5471	0.0644	0.0000	0.0000
1	5.0155	−0.0144	−0.0116	0.0764	−0.0079
2	15.9047	−0.0659	−0.0009	0.0351	−0.0020
3	26.7978	−0.0154	0.0028	0.0225	−0.0008
4	26.7458	0.0355	−0.0026	0.0169	−0.0009
5	16.5743	0.0364	0.0012	0.0154	−0.0012
6	6.4835	−0.0288	0.0047	0.0184	−0.0009
7	1.5928	−0.1436	−0.0143	0.0258	−0.0007
8	0.2373	−0.2640	−0.0182	0.0405	−0.0013
9	0.0195	−0.3229	0.1729	0.0833	−0.0061
10	0.0007	−0.1274	1.4002	0.0000	0.0000

表 3.8. $\{p_j\}_{j=1}^{20}$が区間 $(0.05, 0.95)$ 上の一様乱数 $(p_1, \cdots, p_{10}$ = 0.865169, 0.760194, 0.163742, 0.862702, 0.903511, 0.504386, 0.706864, 0.921407, 0.150022, 0.139438, 0.587975, 0.469021, 0.941708, 0.191839, 0.168768, 0.500518, 0.138714, 0.217245, 0.060993, 0.462495) のときの S_nの分布 $P\{S_n = y\}$ の値と, その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0000	12.0023	—	0.0000	0.0000
1	0.0000	3.0662	−0.9444	0.0815	−0.0061
2	0.0007	1.1739	−0.1700	0.0373	−0.0014
3	0.0105	0.5048	0.0255	0.0215	−0.0007
4	0.0970	0.2109	0.0005	0.0130	−0.0007
5	0.5876	0.0718	0.0020	0.0080	−0.0007
6	2.4259	0.0094	0.0004	0.0053	−0.0006
7	6.9806	−0.0117	−0.0002	0.0043	−0.0004
8	14.2468	−0.0113	−0.0001	0.0040	−0.0001
9	20.8930	−0.0025	−0.0000	0.0040	0.0000
10	22.2171	0.0059	−0.0001	0.0039	−0.0001
11	17.2238	0.0082	0.0001	0.0041	−0.0003
12	9.7565	0.0023	0.0003	0.0050	−0.0005
13	4.0361	−0.0112	−0.0000	0.0068	−0.0004
14	1.2151	−0.0286	−0.0017	0.0097	−0.0002
15	0.2640	−0.0417	−0.0023	0.0133	−0.0001
16	0.0407	−0.0345	0.0078	0.0181	−0.0001
17	0.0043	0.0248	0.0537	0.0251	−0.0005
18	0.0003	0.2134	0.1986	0.0389	−0.0016
19	0.0000	0.7953	0.6495	0.0811	−0.0066
20	0.0000	3.4273	2.6412	0.0000	0.0000

表 3.9. $\{p_j\}_{j=1}^{9} = 0.1(0.1)0.9$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と，その値の Edgeworth 近似，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
5	50.0000	0.0000	—
6	21.5320	−0.0002	−0.1551
7	5.5570	0.0012	−0.0270
8	0.7363	0.0077	−0.0121
9	0.0363	0.1107	—

表 3.10. $\{p_j\}_{j=1}^{19} = 0.05(0.05)0.95$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と，その値の Edgeworth 近似，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
10	50.0000	0.0000	—
11	29.0651	−0.0000	−0.8356
12	13.4452	0.0000	−0.2010
13	4.7942	0.0004	−0.0627
14	1.2761	0.0014	−0.0232
15	0.2442	−0.0006	−0.0097
16	0.0320	−0.0293	−0.0048
17	0.0027	0.2154	−0.0038
18	0.0001	—	−0.0083
19	0.0000	—	—

表 3.11. $\{p_j\}_{j=1}^{10} = 0.05(0.05)0.95$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と，その値の Edgeworth 近似，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
3	55.6619	−0.0014	—
4	27.7313	0.0013	−0.6408
5	9.7465	0.0050	−0.1133
6	2.3314	0.0243	−0.0301
7	0.3634	0.0259	−0.0103
8	0.0345	0.3918	−0.0054
9	0.0018	1.6628	−0.0089
10	0.0000	7.0938	—

表 3.12. $\{p_j\}_{j=1}^{20} = 0.03(0.03)0.60$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と, その値の Edgeworth 近似, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
7	45.0120	0.0002	—
8	26.2247	0.0003	–0.9933
9	12.7076	–0.0003	–0.2467
10	5.0524	–0.0032	–0.0840
11	1.6288	–0.0067	–0.0346
12	0.4207	0.0004	–0.0161
13	0.0859	0.0451	–0.0082
14	0.0136	0.1758	–0.0044
15	0.0016	0.4772	–0.0026
16	0.0001	1.1191	–0.0018
17	0.0000	2.5290	–0.0017
18	0.0000	6.0939	–0.0025
19	0.0000	18.1006	–0.0076
20	0.0000	91.0675	—

表 3.13. $\{p_j\}_{j=1}^{10}$が区間 $(0,1)$ 上の一様乱数 ($p_1, \cdots, p_{10}$ =0.525118, 0.351263,0.000803728, 0.0528506, 0.972966, 0.682429,0.245028,0.219989, 0.812036, 0.812036, 0.569428) のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と, その値の Edgeworth 近似, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
5	47.5781	0.0000	—
6	19.0021	0.0002	–0.1892
7	4.3641	–0.0002	–0.0301
8	0.4836	0.0041	–0.0162
9	0.0165	–0.1031	0.2942
10	0.0000	—	—

表 3.14. $\{p_j\}_{j=1}^{20}$が区間 $(0,1)$ 上の一様乱数 ($p_1, \cdots, p_{10} = 0.305146$, 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) のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と, その値の Edgeworth 近似, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
10	53.7097	–0.0002	—
11	32.6462	–0.0003	—
12	16.0033	0.0001	–0.2793
13	6.1365	0.0012	–0.0836
14	1.7877	0.0023	–0.0306
15	0.3836	–0.0014	–0.0129
16	0.0585	–0.0307	–0.0061
17	0.0060	0.1676	–0.0036
18	0.0004	–0.7757	–0.0032
19	0.0000	—	–0.0067
20	0.0000	—	—

表 3.15. $\{p_j\}_{j=1}^{10}$が区間 $(0.05, 0.95)$ 上の一様乱数 $(p_1, \cdots, p_{10} = 0.120824,\ 0.58353,\ 0.441495,\ 0.17778, 0.447765, 0.185215, 0.197315, 0.484553, 0.206519, 0.747038)$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と, その値の Edgeworth 近似, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
4	51.6537	0.0006	—
5	24.9080	0.0008	-0.4092
6	8.3337	-0.0026	-0.0815
7	1.8502	-0.0129	-0.0219
8	0.2575	0.0121	-0.0073
9	0.0202	0.2682	-0.0073
10	0.0007	1.8232	—

表 3.16. $\{p_j\}_{j=1}^{20}$が区間 $(0.05, 0.95)$ 上の一様乱数 $(p_1, \cdots, p_{10} = 0.865169,\ 0.760194,\ 0.163742,\ 0.862702, 0.903511, 0.504386, 0.706864, 0.921407, 0.150022, 0.139438, 0.587975, 0.469021, 0.941708,\ 0.191839,\ 0.168768,\ 0.500518,\ 0.138714,\ 0.217245,\ 0.060993,\ 0.462495)$ のときの S_nの裾確率 $P\{S_n \geq y\}$ の値と, その値の Edgeworth 近似, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	Edgeworth	LD_2
10	54.7579	0.0001	—
11	32.5408	0.0003	—
12	15.3170	0.0001	-0.2653
13	5.5606	-0.0007	-0.0778
14	1.5244	-0.0020	-0.0284
15	0.3093	-0.0017	-0.0119
16	0.0453	0.0083	-0.0055
17	0.0046	0.0476	-0.0030
18	0.0003	0.1645	-0.0027
19	0.0000	0.5073	-0.0069
20	0.0000	1.8578	—

4. 負の二項分布の場合

確率変数 $X_1, \cdots, X_n, \cdots$ を互いに独立に, 各 $j = 1, 2, \cdots$ について, X_jが

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

を確率関数とする負の二項分布 $NB(r_j, p_j)$ に従うとする. ただし $0 < p_j < 1$ とし, また $q_j = 1 - p_j (j = 1, 2, \cdots)$ とする. このとき $S_n = \sum_{j=1}^n X_j$の分布とその大偏差近似, 正規近似, Edgeworth 近似をそれぞれ求めて数値的に比較する.

(i)S_nの分布 : S_nの分布は次のようにして再帰的に得られる. 例えば $n = 20,\ r_j = r$ $(j = 1, \cdots, 20)$ で $q_1 = \cdots = q_5 = a,\ q_6 = \cdots = q_{10} = b,\ q_{11} = \cdots = q_{15} = c,$

$q_{16} = \cdots = q_{20} = d$ の場合,

$$P\{X_j = x\} = \frac{r(r+1)\cdots(r+x-1)}{x!}a^r(1-a)^x \quad (x = 0,1,2,\ldots;j=1,\ldots,5),$$
$$P\{X_j = x\} = \frac{r(r+1)\cdots(r+x-1)}{x!}b^r(1-b)^x \quad (x = 0,1,2,\ldots;j=6,\ldots,10),$$
$$P\{X_j = x\} = \frac{r(r+1)\cdots(r+x-1)}{x!}c^r(1-c)^x \quad (x = 0,1,2,\ldots;j=11,\ldots,15),$$
$$P\{X_j = x\} = \frac{r(r+1)\cdots(r+x-1)}{x!}d^r(1-d)^x \quad (x = 0,1,2,\ldots;j=16,\ldots,20)$$

となる. ここで $S'_1 := S_5 = \sum_{j=1}^{5} X_j$, $S'_2 := \sum_{j=6}^{10} X_j$, $S'_3 := \sum_{j=11}^{15} X_j$, $S'_4 := \sum_{j=16}^{20} X_j$ とおくと S'_1, S'_2, S'_3, S'_4 はそれぞれ負の二項分布 $NB(5r,a)$, $NB(5r,b)$, $NB(5r,c)$, $NB(5r,d)$ に従う. S'_1 と S'_2 は互いに独立なので $S'_1 + S'_2$ の確率関数は次のようになる.

$$
\begin{aligned}
&P\{S'_1 + S'_2 = 0\} = P\{S'_1 = 0\}P\{S'_2 = 0\} = a^{5r}b^{5r},\\
&P\{S'_1 + S'_2 = 1\}\\
&\quad = P\{S'_1 = 1\}P\{S'_2 = 0\} + P\{S'_1 = 0\}P\{S'_2 = 1\}\\
&\quad = 5ra^{5r}b^{5r}(2-a-b),\\
&P\{S'_1 + S'_2 = 2\}\\
&\quad = P\{S'_1 = 2\}P\{S'_2 = 0\} + P\{S'_1 = 1\}P\{S'_2 = 1\} + P\{S'_1 = 0\}P\{S'_2 = 2\}\\
&\quad = 5ra^{5r}b^{5r}\left\{\frac{5r+1}{2}(1-a)^2 + 5r(1-a)(1-b) + \frac{5r+1}{2}(1-b)^2\right\}\\
&P\{S'_1 + S'_2 = 3\}\\
&\quad = P\{S'_1 = 3\}P\{S'_2 = 0\} + P\{S'_1 = 2\}P\{S'_2 = 1\} + P\{S'_1 = 1\}P\{S'_2 = 2\}\\
&\qquad + P\{S'_1 = 0\}P\{S'_2 = 3\}\\
&\quad = 5r(5r+1)a^{5r}b^{5r}\left\{\frac{1}{6}(5r+2)(1-a)^3 + \frac{5}{2}r(1-a)^2(1-b)\right.\\
&\qquad \left. + \frac{5}{2}r(1-a)(1-b)^2 + \frac{1}{6}(5r+2)(1-b)^3\right\}\\
&\cdots.
\end{aligned}
$$

ここで $S'_{12} := S'_1 + S'_2$ とおくと

$$S'_1 + S'_2 + S'_3 = S'_{12} + S'_3.$$

となり．S'_{12} と S'_3 も互いに独立なので上と同様にして $S'_{12} + S'_3 = S'_1 + S'_2 + S'_3$ の確率関数が得られる．さらに同様の手順を繰り返すことで $S'_1 + S'_2 + S'_3 + S'_4$ の確率関数を得る．

(ii) 大偏差近似：各 $j = 1, 2, \cdots$ について, $|\theta|$ が小さいとき X_jの積率母関数は

$$M_j(\theta) = E(e^{\theta X_j}) = p_j^{r_j}(1 - q_j e^{\theta})^{-r_j}$$

で与えられる．このとき

$$K_n(\theta) = \sum_{j=1}^{n} \log M_j(\theta) = \sum_{j=1}^{n} \{r_j \log p_j - r_j \log(1 - q_j e^{\theta})\}$$

であるから, 各 $y = 0, 1, \cdots$ について

$$K_n^{(1)}(\theta) = \sum_{j=1}^{n} \frac{r_j q_j e^{\theta}}{1 - q_j e^{\theta}} = y$$

となるθを$\hat{\theta}$とする．ここで, $\hat{\tau} = e^{\hat{\theta}}$ とおくと

$$y = \sum_{j=1}^{n} \frac{r_j q_j \hat{\tau}}{1 - q_j \hat{\tau}}$$

になる．また

$$K_n^{(2)}(\hat{\theta}) = \sum_{j=1}^{n} \frac{r_j q_j \hat{\tau}}{(1 - q_j \hat{\tau})^2},$$

$$K^{(3)}(\hat{\theta}) = \sum_{j=1}^{n} \frac{r_j q_j \hat{\tau}(1 + q_j \hat{\tau})}{(1 - q_j \hat{\tau})^3},$$

$$K^{(4)}(\hat{\theta}) = \sum_{j=1}^{n} \frac{r_j q_j (1 + 4 q_j \hat{\tau} + q_j^2 \hat{\tau}^2)}{(1 - q_j \hat{\tau})^4}$$

を得る．従って, 定理 1 から大偏差近似として

$$p_n^*(y) = \frac{1}{\sqrt{2\pi K_n^{(2)}(\hat{\theta})}} \left\{ \prod_{j=1}^{n} \left(\frac{p_j}{1 - q_j \hat{\tau}} \right)^{r_j} \right\} \hat{\tau}^{-y}$$

$$\cdot \left[1 + \frac{K_n^{(4)}(\hat{\theta})}{8\{K_n^{(2)}(\hat{\theta})\}^2} - \frac{5\{K_n^{(3)}(\hat{\theta})\}^2}{24\{K_n^{(2)}(\hat{\theta})\}^3} + O\left(\frac{1}{n^2}\right) \right]$$

を得る．ここで上式の右辺の第１項目までの近似を１次の大偏差近似として LD_1 と表わし，右辺すべての項による近似を２次の大偏差近似として LD_2 で表わす．

一方，S_n のキュムラントは

$$\mu_n := E(S_n) = \sum_{j=1}^{n} \frac{q_j r_j}{p_j}, \qquad v_n := V(S_n) = \sum_{j=1}^{n} \frac{q_j r_j}{p_j^2},$$
$$\kappa_{3,n} := \kappa_3(S_n) = \sum_{j=1}^{n} \frac{q_j(1+q_j) r_j}{p_j^3},$$
$$\kappa_{4,n} := \kappa_4(S_n) = \sum_{j=1}^{n} \frac{q_j r_j}{p_j^4}(1 + 4q_j + q_j^2)$$

となるから，S_n の分布の Edgeworth 展開は

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

になる．ただし $y := (t - \mu_n)\big/\sqrt{v_n}$ とする．上式の第１項目までの近似を正規近似といい，右辺のすべての項による近似を Edgeworth 近似という．

以上の結果を数値的に比較する．いま，(i) $n = 20$, $r_j = 1$ $(j = 1, \ldots, 20)$, $q_1 = \cdots = q_5 = 0.1$, $q_6 = \cdots = q_{10} = 0.2$, $q_{11} = \cdots = q_{15} = 0.3$, $q_{16} = \cdots = q_{20} = 0.4$, (ii) $n = 20$, $r_j = 2$ $(j = 1, \ldots, 20)$, $q_1 = \cdots = q_5 = 0.1$, $q_6 = \cdots = q_{10} = 0.2$, $q_{11} = \cdots = q_{15} = 0.3$, $q_{16} = \cdots = q_{20} = 0.4$, (iii) $n = 20$, $r_j = 1$ $(j = 1, \ldots, 20)$, $q_1 = \cdots = q_5 = 0.2$, $q_6 = \cdots = q_{10} = 0.4$, $q_{11} = \cdots = q_{15} = 0.6$, $q_{16} = \cdots = q_{20} = 0.8$, (iv) $n = 20$, $r_j = 2$ $(j = 1, \ldots, 20)$, $q_1 = \cdots = q_5 = 0.2$, $q_6 = \cdots = q_{10} = 0.4$, $q_{11} = \cdots = q_{15} = 0.6$, $q_{16} = \cdots = q_{20} = 0.8$ の各場合について，正規近似，Edgeworth 近似，j 次の大偏差近似 LD_j $(j = 1, 2)$ の数値計算を行った (表 4.1 〜 4.4 参照)．その結果，大偏差近似 LD_1, LD_2 が他の近似よりかなり精確であることが分かった．

表 4.1. $n=20$, $r_j=1$ $(j=1,\cdots,20)$, $q_1=\cdots=q_5=0.1$, $q_6=\cdots=q_{10}=0.2$, $q_{11}=\cdots=q_{15}=0.3$, $q_{16}=\cdots=q_{20}=0.4$ のときの S_nの分布 $P\{S_n=y\}$ の値と，その値の正規近似，Edgeworth 近似，1 次の大偏差近似 LD_1，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.2529	3.1201	0.3185	0.0000	0.0000
1	1.2644	0.5443	0.1379	0.0845	−0.0059
2	3.3506	−0.0044	0.0296	0.0422	−0.0012
3	6.2587	−0.1701	−0.0061	0.0281	−0.0005
4	9.2491	−0.2028	−0.0090	0.0210	−0.0002
5	11.5098	−0.1711	−0.0024	0.0168	−0.0001
6	12.5389	−0.1026	0.0017	0.0139	−0.0001
7	12.2775	−0.0147	0.0007	0.0119	−0.0000
8	11.0112	0.0766	−0.0015	0.0103	−0.0000
9	9.1744	0.1543	0.0002	0.0091	−0.0000
10	7.1794	0.2011	0.0059	0.0081	−0.0000
11	5.3226	0.2025	0.0080	0.0073	−0.0000
12	3.7646	0.1503	−0.0028	0.0066	−0.0000
13	2.5548	0.0453	−0.0267	0.0060	−0.0000
14	1.6716	−0.1018	−0.0475	0.0055	−0.0000
15	1.0586	−0.2733	−0.0387	0.0050	−0.0000
16	0.6512	−0.4481	0.0164	0.0046	−0.0000
17	0.3902	−0.6079	0.1057	0.0043	−0.0000
18	0.2283	−0.7399	0.1861	0.0040	−0.0001
19	0.1308	−0.8393	0.2047	0.0037	−0.0001
20	0.0735	−0.9078	0.1281	0.0035	−0.0001
21	0.0405	−0.9509	−0.0396	0.0033	−0.0001
22	0.0220	−0.9757	−0.2611	0.0032	−0.0002
23	0.0118	−0.9889	−0.4873	0.0030	−0.0002
24	0.0062	−0.9953	−0.6792	0.0029	−0.0002
25	0.0032	−0.9982	−0.8188	0.0028	−0.0002
26	0.0017	−0.9993	−0.9076	0.0027	−0.0002
27	0.0008	−0.9998	−0.9574	0.0027	−0.0003
28	0.0004	−0.9999	−0.9822	0.0027	−0.0003
29	0.0002	−0.9999	−0.9933	0.0027	−0.0003
30	0.0001	−1.0000	−0.9977	0.0027	−0.0004

表 4.2. $n = 20$, $r_j = 2$ $(j = 1, \cdots, 20)$, $q_1 = \cdots = q_5 = 0.1$, $q_6 = \cdots = q_{10} = 0.2$, $q_{11} = \cdots = q_{15} = 0.3$, $q_{16} = \cdots = q_{20} = 0.4$ のときの S_nの分布 $P\{S_n = y\}$ の値と，その値の正規近似，Edgeworth 近似，1 次の大偏差近似 LD_1，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0006	97.8639	—	0.0000	0.0000
1	0.0064	17.9615	—	0.0844	−0.0060
2	0.0329	5.7426	−0.6928	0.0422	−0.0012
3	0.1164	2.3357	−0.0674	0.0281	−0.0005
4	0.3173	1.0422	0.0456	0.0210	−0.0003
5	0.7115	0.4515	0.0460	0.0168	−0.0002
6	1.3660	0.1504	0.0266	0.0140	−0.0001
7	2.3082	−0.0110	0.0101	0.0120	−0.0001
8	3.5026	−0.0961	0.0006	0.0105	−0.0001
9	4.8463	−0.1349	−0.0031	0.0093	−0.0000
10	6.1873	−0.1433	−0.0030	0.0084	−0.0000
11	7.3591	−0.1306	−0.0015	0.0076	−0.0000
12	8.2185	−0.1028	0.0000	0.0069	−0.0000
13	8.6742	−0.0647	0.0006	0.0064	−0.0000
14	8.7003	−0.0205	0.0004	0.0059	−0.0000
15	8.3318	0.0259	−0.0003	0.0055	−0.0000
16	7.6491	0.0700	−0.0006	0.0051	−0.0000
17	6.7559	0.1076	0.0001	0.0048	−0.0000
18	5.7583	0.1342	0.0015	0.0045	−0.0000
19	4.7493	0.1460	0.0028	0.0043	−0.0000
20	3.7996	0.1397	0.0027	0.0040	−0.0000

表 4.3. $n = 20$, $r_j = 1$ $(j = 1, \cdots, 20)$, $q_1 = \cdots = q_5 = 0.2$, $q_6 = \cdots = q_{10} = 0.4$, $q_{11} = \cdots = q_{15} = 0.6$, $q_{16} = \cdots = q_{20} = 0.8$ のときの S_nの分布 $P\{S_n = y\}$ の値と，その値の正規近似, Edgeworth 近似, 1 次の大偏差近似 LD_1, 2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0000	7135.4238	—	0.0000	0.0000
1	0.0001	916.1823	—	0.0845	−0.0059
2	0.0004	219.6504	—	0.0422	−0.0012
3	0.0017	73.7118	—	0.0281	−0.0005
4	0.0049	30.7225	—	0.0210	−0.0002
5	0.0122	14.8687	—	0.0168	−0.0001
6	0.0265	7.9959	—	0.0139	−0.0001
7	0.0519	4.6291	−0.2211	0.0119	−0.0001
8	0.0931	2.8151	0.1071	0.0103	−0.0000
9	0.1551	1.7612	0.2037	0.0091	−0.0000
10	0.2427	1.1110	0.2100	0.0081	−0.0000
11	0.3599	0.6900	0.1831	0.0073	−0.0000
12	0.5091	0.4070	0.1469	0.0066	−0.0000
13	0.6910	0.2111	0.1113	0.0060	−0.0000
14	0.9042	0.0728	0.0801	0.0055	−0.0000
15	1.1454	−0.0261	0.0543	0.0050	−0.0000
16	1.4091	−0.0969	0.0341	0.0046	−0.0000
17	1.6887	−0.1471	0.0188	0.0043	−0.0000
18	1.9764	−0.1817	0.0078	0.0040	−0.0001
19	2.2640	−0.2043	0.0004	0.0037	−0.0001
20	2.5432	−0.2171	−0.0042	0.0035	−0.0001
21	2.8062	−0.2221	−0.0067	0.0033	−0.0001
22	3.0462	−0.2205	−0.0075	0.0032	−0.0001
23	3.2573	−0.2134	−0.0072	0.0030	−0.0002
24	3.4351	−0.2015	−0.0062	0.0029	−0.0002
25	3.5766	−0.1854	−0.0049	0.0028	−0.0002
26	3.6801	−0.1658	−0.0036	0.0028	−0.0002
27	3.7452	−0.1432	−0.0024	0.0027	−0.0003
28	3.7729	−0.1179	−0.0014	0.0027	−0.0003
29	3.7648	−0.0905	−0.0008	0.0027	−0.0003
30	3.7237	−0.0614	−0.0005	0.0027	−0.0004

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
31	3.6526	−0.0310	−0.0004	0.0027	−0.0004
32	3.5553	0.0001	−0.0004	0.0027	−0.0004
33	3.4357	0.0316	−0.0004	0.0028	−0.0004
34	3.2975	0.0628	−0.0003	0.0029	−0.0005
35	3.1447	0.0932	0.0000	0.0030	−0.0005
36	2.9810	0.1223	0.0007	0.0031	−0.0005
37	2.8098	0.1497	0.0018	0.0031	−0.0005
38	2.6344	0.1746	0.0032	0.0033	−0.0005
39	2.4575	0.1966	0.0048	0.0034	−0.0005
40	2.2815	0.2151	0.0064	0.0035	−0.0005
41	2.1086	0.2297	0.0078	0.0036	−0.0006
42	1.9404	0.2399	0.0087	0.0038	−0.0006
43	1.7785	0.2454	0.0087	0.0039	−0.0006
44	1.6237	0.2458	0.0075	0.0040	−0.0006
45	1.4770	0.2409	0.0048	0.0042	−0.0006
46	1.3289	0.2305	0.0004	0.0043	−0.0006
47	1.2097	0.2146	−0.0058	0.0045	−0.0006
48	1.0895	0.1931	−0.0136	0.0047	−0.0006
49	0.9782	0.1663	−0.0228	0.0048	−0.0006
50	0.8758	0.1343	−0.0330	0.0050	−0.0006
51	0.7820	0.0975	−0.0436	0.0051	−0.0007
52	0.6964	0.0563	−0.0536	0.0053	−0.0006
53	0.6186	0.0111	−0.0622	0.0055	−0.0006
54	0.5481	−0.0375	−0.0685	0.0056	−0.0007
55	0.4846	−0.0889	−0.0714	0.0058	−0.0006
56	0.4275	−0.1425	−0.0699	0.0060	−0.0006
57	0.3763	−0.1976	−0.0634	0.0062	−0.0006
58	0.3205	−0.2536	−0.0513	0.0063	−0.0006
59	0.2898	−0.3098	−0.0334	0.0065	−0.0006
60	0.2536	−0.3656	−0.0097	0.0067	−0.0006

表 4.4. $n = 20$, $r_j = 2$ $(j = 1, \cdots, 20)$, $q_1 = \cdots = q_5 = 0.2$, $q_6 = \cdots = q_{10} = 0.4$, $q_{11} = \cdots = q_{15} = 0.6$, $q_{16} = \cdots = q_{20} = 0.8$ のときの S_nの分布 $P\{S_n = y\}$ の値と，その値の正規近似，Edgeworth 近似，1 次の大偏差近似 LD_1，2 次の大偏差近似 LD_2の相対誤差

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
0	0.0000	—	—	0.0000	0.0000
1	0.0000	—	—	0.0844	−0.0059
2	0.0000	—	—	0.0422	−0.0012
3	0.0000	—	—	0.0281	−0.0005
4	0.0000	—	—	0.0210	−0.0003
5	0.0000	—	—	0.0168	−0.0002
6	0.0000	—	—	0.0140	−0.0001
7	0.0000	—	4261.7750	0.0120	−0.0001
8	0.0000	4883.9078	813.9298	0.0105	−0.0001
9	0.0000	2202.2702	21.5215	0.0093	−0.0000
10	0.0000	1071.1525	—	0.0083	−0.0000
11	0.0000	556.8065	—	0.0076	−0.0000
12	0.0000	306.8562	—	0.0069	−0.0000
13	0.0001	178.0666	—	0.0064	−0.0000
14	0.0002	108.1674	—	−0.0115	−0.0000
15	0.0003	68.4290	—	0.0055	−0.0000
16	0.0005	44.8772	—	0.0051	−0.0000
17	0.0010	30.3857	—	0.0048	−0.0000
18	0.0016	21.1616	—	0.0045	−0.0000
19	0.0027	15.1073	—	0.0043	0.0000
20	0.0043	11.0210	—	0.0040	−0.0000
21	0.0068	8.1921	—	0.0038	−0.0000
22	0.0102	6.1880	—	0.0036	0.0000
23	0.0151	4.7379	−0.8014	0.0034	0.0000
24	0.0219	3.6685	−0.4941	0.0033	0.0000
25	0.0309	2.8658	−0.2898	0.0031	0.0000
26	0.0428	2.2537	−0.1551	0.0030	−0.0000
27	0.0582	1.7802	−0.0675	0.0029	−0.0000
28	0.0777	1.4091	−0.0120	0.0027	−0.0000
29	0.1019	1.1149	0.0219	0.0026	−0.0000
30	0.1317	0.8791	0.0412	0.0025	−0.0000

y	真値 (%)	正規近似	Edgeworth	LD_1	LD_2
31	0.1675	0.6885	0.0506	0.0024	−0.0000
32	0.2101	0.5331	0.0535	0.0023	−0.0000
33	0.2598	0.4055	0.0524	0.0022	−0.0000
34	0.3173	0.3002	0.0487	0.0022	−0.0000
35	0.3826	0.2130	0.0436	0.0021	−0.0000
36	0.4561	0.1404	0.0378	0.0020	−0.0000
37	0.5376	0.0799	0.0318	0.0019	−0.0000
38	0.6269	0.0295	0.0261	0.0019	−0.0000
39	0.7237	−0.0124	0.0207	0.0018	−0.0000
40	0.8273	−0.0472	0.0159	0.0018	−0.0000
41	0.9369	−0.0758	0.0117	0.0017	−0.0000
42	1.0517	−0.0990	0.0081	0.0017	−0.0000
43	1.1705	−0.1177	0.0050	0.0016	−0.0000
44	1.2921	−0.1324	0.0026	0.0016	−0.0000
45	1.4152	−0.1435	0.0007	0.0016	−0.0000
46	1.5384	−0.1514	−0.0007	0.0015	−0.0000
47	1.6603	−0.1566	−0.0017	0.0015	−0.0000
48	1.7794	−0.1592	−0.0023	0.0015	−0.0001
49	1.8944	−0.1595	−0.0027	0.0015	−0.0001
50	2.0040	−0.1578	−0.0028	0.0015	−0.0001
51	2.1067	−0.1542	−0.0027	0.0014	−0.0001
52	2.2017	−0.1489	−0.0025	0.0014	−0.0001
53	2.2877	−0.1420	−0.0022	0.0014	−0.0001
54	2.3640	−0.1338	−0.0018	0.0014	−0.0001
55	2.4298	−0.1242	−0.0015	0.0014	−0.0001
56	2.4846	−0.1136	−0.0012	0.0014	−0.0001
57	2.5281	−0.1019	−0.0008	0.0014	−0.0001
58	2.5600	−0.0893	−0.0006	0.0014	−0.0001
59	2.5804	−0.0759	−0.0004	0.0014	−0.0001
60	2.5893	−0.0618	−0.0002	0.0014	−0.0001

y	Exact(%)	Normal	Edgeworth	LD_1	LD_2
61	2.5869	−0.0472	−0.0002	0.0014	−0.0001
62	2.5738	−0.0321	−0.0001	0.0014	−0.0001
63	2.5503	−0.0167	−0.0001	0.0015	−0.0001
64	2.5172	−0.0011	−0.0001	0.0015	−0.0001
65	2.4750	0.0145	−0.0001	0.0015	−0.0001
66	2.4246	0.0302	−0.0001	0.0015	−0.0001
67	2.3667	0.0456	−0.0001	0.0015	−0.0001
68	2.3022	0.0608	0.0000	0.0015	−0.0001
69	2.2320	0.0755	0.0002	0.0016	−0.0001
70	2.1568	0.0896	0.0004	0.0016	−0.0001
71	2.0776	0.1031	0.0006	0.0016	−0.0001
72	1.9952	0.1157	0.0010	0.0016	−0.0001
73	1.9103	0.1273	0.0013	0.0016	−0.0002
74	1.8237	0.1378	0.0017	0.0017	−0.0001
75	1.7361	0.1471	0.0021	0.0017	−0.0001
76	1.6483	0.1551	0.0025	0.0018	−0.0001
77	1.5607	0.1616	0.0028	0.0018	−0.0001
78	1.4739	0.1666	0.0030	0.0018	−0.0001
79	1.3884	0.1698	0.0031	0.0018	−0.0002
80	1.3047	0.1713	0.0030	0.0019	−0.0001
81	1.2231	0.1710	0.0026	0.0019	−0.0002
82	1.1440	0.1687	0.0020	0.0019	−0.0001
83	1.0675	0.1644	0.0010	0.0020	−0.0002
84	0.9939	0.1581	−0.0002	0.0020	−0.0001
85	0.9235	0.1498	−0.0018	0.0020	−0.0002
86	0.8562	0.1394	−0.0037	0.0021	−0.0001
87	0.7922	0.1016	−0.0059	0.0021	−0.0001
88	0.7315	0.1123	−0.0084	0.0022	−0.0002
89	0.6742	0.0957	−0.0110	0.0022	−0.0001
90	0.6201	0.0772	−0.0138	0.0023	−0.0001
91	0.5694	0.0567	−0.0166	0.0023	−0.0002
92	0.5218	0.0343	−0.0193	0.0023	−0.0002
93	0.4774	0.0102	−0.0219	0.0024	−0.0001
94	0.4360	−0.0155	−0.0240	0.0024	−0.0001
95	0.3975	−0.0428	−0.0257	0.0024	−0.0002
96	0.3619	−0.0715	−0.0267	0.0025	−0.0002
97	0.3288	−0.1015	−0.0270	0.0025	−0.0001
98	0.2984	−0.1325	−0.0263	0.0026	−0.0002
99	0.2703	−0.1646	−0.0246	0.0026	−0.0002
100	0.2445	−0.1974	−0.0217	0.0027	−0.0002

参考文献

[ATT98] Akahira, M. Takahashi, K. and Takeuchi, K. (1998). The higher order large-deviation approximation for the distribution of the sum of independent discrete random variables. To appear in the *Commun. Statist. - Theory Meth.*.

[BC89] Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics.* Chapman and Hall, London.

[D87] Daniels, H. E. (1987). Tail probability approximations. *Int. Statist. Review*, **55**, 37–48.

[J95] Jensen, J. L. (1995). *Saddlepoint Approximations.* Clarendon Press, Oxford.

[LR80] Lugannani, R. and Rice, S. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables. *Adv. Appl. Prob.*, **12**, 475–490.

[R82] Robinson, J. (1982). Saddlepoint approximations for permutation tests and confidence intervals. *J. R. Statist. Soc.*, **B44**, 91–101.