Сибирский математический журнал
Март—апрель, 2001. Том 42, № 2

УДК 519.21+519.85

О ПУАССОНОВСКОЙ АППРОКСИМАЦИИ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ

П. С. Рузанкин

Аннотация: Пусть Q — биномиальное распределение с параметрами n, p, а P — пуассоновское распределение с параметром $\lambda = np$. Оценивается расстояние

$$\rho(z, P, Q) = \inf_{\xi,\eta}\{\mathbf{P}(|\xi - \eta| > z)\},$$

где инфимум берется по всем случайным величинам ξ и η, заданным на одном вероятностном пространстве, с распределениями P и Q соответственно. Библиогр. 11.

Пусть P_1 и P_2 — произвольные вероятностные распределения на вещественной оси и z — произвольная неотрицательная постоянная. Обозначим через $\rho(z, P_1, P_2)$ так называемое *расстояние Дадли* между P_1 и P_2:

$$\rho(z, P_1, P_2) = \inf_{\xi,\eta}\{\mathbf{P}(|\xi - \eta| > z)\}, \tag{1}$$

где инфимум берется по всем случайным величинам ξ и η, заданным на одном вероятностном пространстве, с распределениями P_1 и P_2 соответственно. В 1968 г. Р. М. Дадли доказал [1], что

$$\rho(z, P_1, P_2) = \sup\{P_1(H) - P_2(H^z) : H \text{ замкнуто}\}, \tag{2}$$

где $H^z = \{y : \inf\{|w - y| : w \in H\} \le z\}$. Отсюда, в частности, следует, что $\rho(0, P_1, P_2)$ — это расстояние полной вариации между P_1 и P_2, а $\inf\{z : \rho(z, P_1, P_2) \le z\}$ — расстояние Прохорова, метризующее слабую сходимость мер (см. [2]).

Обозначим через F и G функции распределений P_1 и P_2 соответственно. В дальнейшем нам потребуется также следующая теорема (см. [3]).

Теорема A. *Пусть* $z > 0$. *Тогда*

$$\rho(z, P_1, P_2) = \lim_{y\to\infty} S(y) - 1 = \lim_{y\to\infty} T(y) - 1,$$

где функции S *и* T *задаются соотношениями*

$$\lim_{y\to-\infty} S(y) = \lim_{y\to-\infty} T(y) = 0, \tag{3}$$

$$dS(y) = \max\{dF(y), T(y + dy - z) - S(y)\}, \tag{4}$$

$$dT(y) = \max\{dG(y), S(y + dy - z) - T(y)\} \tag{5}$$

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (коды проектов 99–01–00504, 00–01–00802) и INTAS (код проекта 99–01317).

© 2001 Рузанкин П. С.

при всех y и требованием непрерывности функций S и T слева. Функции S и T, удовлетворяющие перечисленным условиям, существуют и задаются этими условиями единственным образом.

Положим $P_1 = P$, $P_2 = Q$, где Q — биномиальное распределение с параметрами n, p, а P — пуассоновское распределение с параметром $\lambda = np$. Тогда F и G будут функциями распределений P и Q соответственно. В настоящей работе оценивается величина $\rho(z, P, Q)$.

Один из первых результатов, связанных с данной проблемой, получен Ю. В. Прохоровым в 1953 г. (см. [4]). Им доказано, что $\rho(0, P, Q) \le cp$, где c — абсолютная постоянная. Позже Л. Ле Кам установил (см. [5]), что также $\rho(0, P, Q) \le np^2$. Комбинируя эти два неравенства, получим

$$\rho(0, P, Q) \le \min\{cp, np^2\}.$$

А. Барбур и П. Холл в [6] показали, что имеет место аналогичная оценка снизу.

В работах [7–11] при исследовании потраекторной близости процесса частных сумм и соответствующего пуассоновского процесса были получены как частные случаи оценки расстояния $\rho(z, P, Q)$ для произвольного z. Наиболее сильные оценки потраекторной близости содержатся в работе И. С. Борисова [11]. В частности,

$$\rho(z, P, Q) \le \begin{cases} (np^2)^{\lfloor z \rfloor + 1} \exp\{-C_1 z \log\log(z+2) + C_2\}, & \text{если } np \ge 1, \\ np^{\lfloor z \rfloor + 2} \exp\{-C_3 z + C_4\}, & \text{если } np \le 1, \end{cases}$$

где C_1, C_2, C_3, C_4 — абсолютные положительные постоянные, а символом $\lfloor z \rfloor$ обозначено наибольшее целое число, не превосходящее z.

Однако для аппроксимации биномиального распределения в некотором диапазоне значений n, p, z эти оценки уже не являются оптимальными. Так, П. Майор в [7] доказал, что для любого $p \le n^{-2/3}$ найдутся абсолютные положительные постоянные C и K, для которых имеет место неравенство

$$\rho(C, P, Q) < K \exp\left(-\frac{1}{8}\sqrt{n}\log n\right). \tag{6}$$

При доказательстве (6) оценивается вероятность $\mathbf{P}(|\xi - \eta| > C)$, где в качестве случайных величин ξ и η берутся квантильные преобразования $F^{-1}(\omega)$ и $G^{-1}(\omega)$ случайной величины ω, равномерно распределенной на отрезке $(0, 1)$. Такое задание случайных величин может и не быть оптимальным в смысле (1) при различных значениях параметров n, p, z. Доказательство следующего утверждения опирается на теорему А.

Теорема 1. *Пусть $p \le 1/2$ и $z \ge 1$ целое. Тогда*

$$\rho(z, P, Q) < \frac{7}{3}\exp\left\{-\frac{1}{27}\frac{z^2}{np^3}\right\} \quad \text{при } z \le np^2, \tag{7}$$

$$\rho(z, P, Q) < \exp\left\{-\frac{1}{15}np\right\} \quad \text{при } z \ge np^2, \tag{8}$$

$$\rho(z, P, Q) < \exp\left\{-\frac{1}{4}\sqrt{nz\left(\log\frac{z}{np^2} - 2\right)^3}\right\} \quad \text{при } e^4 np^2 \le z \le \frac{n(1-p)}{\log\frac{1}{p}}, \tag{9}$$

$$\rho(z, P, Q) = P([n+z+1, \infty)) < \exp\left\{-(n+z)\log\frac{n+z}{enp}\right\} \quad \text{при } z \ge \frac{n(1-p)}{\log\frac{1}{p}}. \tag{10}$$

Кроме того,

$$\rho(z,P,Q) > \exp\left\{-4\sqrt{nz\left(\log\frac{z}{np^2}+4\right)^3}\right\},\ \textit{если } p \le \frac{1}{10} \textit{ и } np^2 \le z \le \frac{n}{10\log\frac{1}{p}}. \tag{11}$$

Замечание. Неравенство (11) показывает, что в некотором диапазоне значений n, p, z оценки (8), (9) имеют неулучшаемый порядок логарифмической асимптотики. Оценка (10) также точна в этом смысле, так как

$$P([n+z+1,\infty)) = \frac{\theta}{\sqrt{2\pi(n+z+1)}}\exp\left\{-(n+z+1)\log\frac{n+z+1}{enp}-np\right\}, \tag{12}$$

где $e^{-1/12(n+z+2)} < \theta < \frac{1}{1-p}e^{-1/12(n+z+1)}$. Последнее равенство установлено в доказательстве теоремы 1.

Проведем сравнение утверждения теоремы 1 с оценкой (6).

1. В отличие от результата П. Майора теорема 1 содержит оценки расстояния $\rho(z,P,Q)$ для всех $z \ge 1$.

2. Для того чтобы применение теоремы 1 имело смысл, достаточно, чтобы $np^3 \to 0$, в то время как для выполнения неравенства (6) необходимо, чтобы $np^{3/2} \le 1$.

3. Для любого фиксированного z в силу неравенств (9), (10) равномерно для всех $p \le n^{-2/3}$ имеем

$$\rho(z,P,Q) < K_1\exp\{-K_2\sqrt{n(\log n)^3}\},$$

где K_1 и K_2 — легко вычисляемые абсолютные положительные постоянные. Следовательно, при фиксированном z и $p \le n^{-2/3}$ оценка теоремы 1 имеет большую, чем в неравенстве (6), скорость убывания.

Доказательство теоремы 1. Обозначим $f(k) = P(\{k\})$, $g(k) = Q(\{k\})$.

Лемма 1. *Имеют место следующие оценки:*

$$\frac{g(k)}{f(k)} \ge \exp\left\{-\frac{(k-np)^2}{2n(1-p)}\right\} \quad \textit{при } k \le np, \tag{13}$$

$$\frac{g(k)}{f(k)} \ge \exp\left\{-\frac{(k-np)^2}{n(1-p)}\right\} \quad \textit{при } np \le k \le n, \tag{14}$$

$$\frac{g(k)}{f(k)} \le \frac{1}{1-p}\exp\left\{-\frac{(k-1-np)^2}{2n(1-p)}\right\} \quad \textit{при } np \le k \le n. \tag{15}$$

Доказательство. Для $k \le n$ имеем

$$\frac{g(k)}{f(k)} = \frac{n(n-1)\dots(n-k+1)}{n^k(1-p)^k}(1-p)^n e^{np}$$

$$= \exp\left\{n(p+\log(1-p)) - k\log(1-p) + \sum_{j=0}^{k-1}\log\left(1-\frac{j}{n}\right)\right\}$$

$$\ge \exp\left\{n(p+\log(1-p)) - k\log(1-p) + n\int_0^{k/n}\log(1-x)\,dx\right\} = \exp\{-nH_p(k/n)\},$$

где
$$H_p(x) = -p + x + (1-x)\log\frac{1-x}{1-p}.$$
Отметим очевидные свойства функции H_p:
$$H_p(x) \geq 0 \text{ при } 0 \leq x \leq 1, \quad H_p(1) = 1-p, \quad H_p(p) = 0,$$
$$\frac{d}{dx}H_p(p) = 0, \quad \frac{d^2}{dx^2}H_p(x) = \frac{1}{1-x}.$$
Отсюда легко построить мажоранты для H_p:
$$H_p(x) \leq \frac{(p-x)^2}{2(1-p)} \text{ при } 0 \leq x \leq p, \quad H_p(x) \leq \frac{(x-p)^2}{(1-p)} \text{ при } p \leq x \leq 1.$$
Это и влечет неравенства (13) и (14). Для доказательства неравенства (15) можно сделать аналогичную оценку сверху:
$$\frac{g(k)}{f(k)} \leq \exp\Biggl\{-\log(1-p) + n(p+\log(1-p)) - (k-1)\log(1-p)$$
$$+ n\int\limits_0^{(k-1)/n}\log(1-x)\,dx\Biggr\} \leq \exp\biggl\{-\log(1-p) - nH_p\biggl(\frac{k-1}{n}\biggr)\biggr\}.$$
Отсюда следует неравенство (15), так как $H_p(x) \geq (x-p)^2/2(1-p)$ при $p \leq x \leq 1$.

Лемма 2. *Пусть* $z \leq np^2$, $p \leq 1/2$. *Тогда*
$$g(k) \geq f(k-z) \quad \textit{при } np - z/p \leq k \leq np, \tag{16}$$
$$g(k) \geq f(k-z) \quad \textit{при всех } k \leq np, \quad \textit{если } z \geq \frac{1}{2}np^2, \tag{17}$$
$$g(k) > f(k+z) \quad \textit{при } np \leq k \leq np + \frac{1}{3}\frac{z}{p}. \tag{18}$$

Доказательство. В силу неравенства (13) при $z \leq k \leq np$ имеем
$$\frac{g(k)}{f(k-z)} = \frac{g(k)}{f(k)}\frac{f(k)}{f(k-z)} \geq \exp\biggl\{-\frac{(k-np)^2}{n}\biggr\}\biggl(\frac{np}{k}\biggr)^z.$$
Последнее выражение не меньше 1, если
$$-\frac{(k-np)^2}{n} + z\log\frac{np}{k} \geq 0.$$
Обозначив $\delta = 1 - \frac{k}{np}$, перепишем это неравенство:
$$\frac{\delta^2}{-\log(1-\delta)} \leq \frac{z}{np^2}, \tag{19}$$
где значение выражения $\frac{\delta^2}{-\log(1-\delta)}$ в точке 0 полагается равным 0. Для выполнения последнего неравенства достаточно, чтобы $\delta \leq \frac{z}{np^2}$, т. е. чтобы k было не меньше, чем $np - \frac{z}{p}$. Таким образом, доказана оценка (16) при $k \geq z$. Правильность этой оценки при $0 \leq k < z$ очевидна.

Заметим также, что $\frac{\delta^2}{-\log(1-\delta)} \le \frac{1}{2}$ при всех δ, $0 \le \delta < 1$, а значит, верна и оценка (17).

Аналогично в силу неравенства (13) при $k \ge np$ имеем

$$\frac{g(k)}{f(k+z)} = \frac{g(k)}{f(k)}\frac{f(k)}{f(k+z)} > \exp\left\{-\frac{(k-np)^2}{(1-p)n}\right\}\left(\frac{k}{np}\right)^z.$$

Последнее выражение не меньше 1, если

$$-\frac{(k-np)^2}{n(1-p)} + z\log\frac{k}{np} \ge 0.$$

Обозначим $\zeta = \frac{k}{np} - 1$ и перепишем это неравенство в виде

$$\frac{\zeta^2}{\log(1+\zeta)} \le (1-p)\frac{z}{np^2}, \tag{20}$$

где при $\zeta = 0$ полагаем $\frac{\zeta^2}{\log(1+\zeta)} = 0$. Так как $\frac{z}{np^2} \le 1$, то $\zeta < \frac{4}{5}$. При $0 \le \zeta \le \frac{4}{5}$ выполнено $\frac{\zeta^2}{\log(1+\zeta)} \le \frac{3}{2}\zeta$. Значит, неравенство (20) верно, если

$$\zeta \le \frac{1}{3}\frac{z}{np^2},$$

т. е. $k \le np + \frac{1}{3}\frac{z}{np^2}$.

Лемма 3. *Пусть $z \ge np^2$, $p \le \frac{1}{2}$. Тогда*

$$g(k) \ge f(k-z) \quad \textit{при всех } k \le np, \tag{21}$$

$$g(k) > f(k+z) \quad \textit{при } np \le k \le \frac{7}{5}np, \tag{22}$$

$$g(k) > f(k+z) \textit{ при } e^4np^2 \le z \le \frac{n(1-p)}{\log\frac{1}{p}}, \quad np \le k \le np + \frac{1}{2}\sqrt{nz\log\frac{z}{np^2}}. \tag{23}$$

$$g(k) > f(k+z) \textit{ при } z \ge \frac{n(1-p)}{\log\frac{1}{p}}, \quad np \le k \le n. \tag{24}$$

Доказательство. Вывод оценки (21) аналогичен выводу (17) в лемме 2.

Чтобы проверить неравенства (22) и (23) обозначим, как и в доказательстве леммы 2, $\zeta = \frac{k}{np} - 1$. Для того чтобы при $k \ge np$ выполнялось $g(k) > f(k+z)$ достаточно, чтобы выполнялось (20). При $\frac{z}{np^2} \ge \frac{1}{2}$ и $p \le \frac{1}{2}$ последнее неравенство верно при всех ζ, $0 \le \zeta \le \frac{2}{5}$, т. е. при всех k, $np \le k \le \frac{7}{5}np$.

Если

$$\frac{z}{np^2} \ge e^4, \quad p \le \frac{1}{2}, \quad \zeta \le \frac{1}{2}\sqrt{\frac{z}{np^2}\log\frac{z}{np^2}},$$

то

$$\frac{\zeta^2}{\log(1+\zeta)} \le \frac{\frac{1}{2}\frac{z}{np^2}\log\frac{z}{np^2}}{\log\frac{z}{np^2} + \log\log\frac{z}{np^2}} \le (1-p)\frac{z}{np^2},$$

т. е. выполнено (20).

Заметим, что так как $z \le \frac{n(1-p)}{\log\frac{1}{p}}$, из неравенства (20) следует $\zeta \le \frac{1-p}{p}$, а значит, $np + \frac{1}{2}\sqrt{nz\log\frac{z}{np^2}} \le n$.

Для доказательства (24) достаточно заметить, что при $z \ge \frac{n(1-p)}{\log\frac{1}{p}}$ неравенство (20) выполняется при всех ζ, $0 \le \zeta \le \frac{1-p}{p}$, т. е. при всех k, $np \le k \le n$.

Лемма 4. *Пусть*
$$p \le \frac{1}{10}, \quad np^2 \le z \le \frac{n}{10\log\frac{1}{p}}.$$

Тогда
$$\frac{5}{4}(1-G(k)) \le 1-F(k+z) \quad \textit{при}$$
$$np + 2\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)} - 1 \le k \le \min\Big\{n,\ np + \frac{5}{2}\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)}\Big\}. \quad (25)$$

Доказательство леммы 4. При $k \ge np + \sqrt{nz\big(\log\frac{z}{np^2}+4\big)}$ и $z \ge np^2$ имеем
$$1-G(k) \le \frac{3}{2}g(k), \quad 1-F(k) \ge f(k).$$
Для того чтобы выполнялось $\frac{5}{4}(1-G(k)) \le 1-F(k+z)$ при таких k, что
$$np + \sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)} \le k \le np + \frac{5}{2}\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)},$$
достаточно показать, что
$$2\frac{g(k)}{f(k)}\,\frac{f(k)}{f(k+z)} \le 1,$$
а ввиду неравенства (15) — что
$$\log\frac{2}{1-p} - \frac{(k-1-np)^2}{2n} + z\log\frac{k+z}{np} \le 0.$$
Так как при $z \le n$ имеем $k+z \le \frac{7}{2}\sqrt{nz\big(\log\frac{z}{np^2}+4\big)}$, достаточно, чтобы
$$k-1-np \ge \sqrt{2n\Big(\frac{1}{2}z\log\frac{z}{np^2} + \frac{1}{2}\log\Big(\log\frac{z}{np^2}+4\Big) + z\log\frac{7}{2} + \log\frac{2}{1-p}\Big)},$$
для выполнения последнего неравенства — чтобы
$$k-1-np \ge \sqrt{2nz\Big(\log\frac{z}{np^2}+4\Big)}.$$
Последнее соотношение верно, если
$$k \ge np + 2\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)} - 1.$$

Для завершения доказательства леммы осталось заметить, что при $p \le \frac{1}{10}$ и $z \le \frac{n}{10\log\frac{1}{p}}$ имеют место неравенства
$$np + 2\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)} \le \frac{5}{2}\sqrt{nz\Big(\log\frac{z}{np^2}+4\Big)} \le n.$$

Доказательство теоремы. Символом $\lceil y\rceil$ будем обозначать наименьшее целое число, не меньшее чем y: $\lceil y\rceil = -\lfloor -y\rfloor$.

Обозначим

$$R(k) = \frac{g(k+1)}{f(k+1)} : \frac{g(k)}{f(k)}.$$

Имеем

$$R(k) = \frac{1-k/n}{1-p}.$$

Значит, $f(k)/g(k)$ достигает максимума при $k = \lceil np \rceil$, и $g(\lceil np \rceil) > f(\lceil np \rceil)$. Так как $f(k)$ убывает при $k \geq np$, то

$$g(\lceil np \rceil) > f(\lceil np \rceil + z). \tag{26}$$

Обозначим также

$$R_1(k) = \frac{g(k+1)}{f(k-z+1)} : \frac{g(k)}{f(k-z)}.$$

Тогда

$$R_1(k) = \frac{1-k/n}{1-p}\left(1 - \frac{z}{k+1}\right), \quad \frac{d}{dk}R_1(k) = \frac{1}{1-p}\left(\frac{z}{(k+1)^2}\left(1+\frac{1}{n}\right) - \frac{1}{n}\right).$$

Значит, существует такое $k_1 \geq z$, что $R_1(k)$ возрастает при $z \leq k < k_1$ и убывает при $k_1 \leq k \leq n$, причем $R_1(n) = 0$. Теперь нетрудно убедиться в том, что всевозможные соотношения между $f(k-z)$ и $g(k)$ могут быть заданы следующими неравенствами:

$$f(k-z) \leq g(k) \quad \text{при } z \leq k < k_2, \tag{27}$$

$$f(k-z) > g(k) \quad \text{при } k_2 \leq k < k_3, \tag{28}$$

$$f(k-z) \leq g(k) \quad \text{при } k_3 \leq k < k_4, \tag{29}$$

$$f(k-z) > g(k) \quad \text{при } k \geq k_4, \tag{30}$$

где $z \leq k_2 \leq k_3 \leq k_4$; при этом некоторые из интервалов изменения аргумента в (27)–(29) могут быть пустыми (например, если $z = k_2$). Эти соотношения можно доказать, рассмотрев всевозможные случаи поведения $R_1(k)$ при $z \leq k \leq n$. Заметим, что $R_1(k)$ может пересечь единичный уровень не более двух раз. Отсюда можно заключить, что отношение $g(k)/f(k-z)$ пересечет единичный уровень не более трех раз, что и записано в (27)–(30).

Заметим также, что $R_1(k) < 1$ при $k \geq np$, а значит, можно считать, что

$$k_3 \leq np. \tag{31}$$

Аналогично для

$$R_2(k) = \frac{g(k+1)}{f(k+z+1)} : \frac{g(k)}{f(k+z)}$$

имеем

$$R_2(k) = \frac{1-k/n}{1-p}\left(1 + \frac{z}{k+1}\right), \quad \frac{d}{dk}R_2(k) = \frac{1}{1-p}\left(-\frac{z}{(k+1)^2}\left(1+\frac{1}{n}\right) - \frac{1}{n}\right) < 0.$$

Следовательно, существуют такие k_5 и k_6, что $z \leq k_5 \leq k_6$ и

$$g(k-z) \leq f(k) \quad \text{при } k < k_5, \tag{32}$$

$$g(k-z) > f(k) \quad \text{при } k_5 \leq k < k_6, \tag{33}$$

$$g(k-z) \leq f(k) \quad \text{при } k \geq k_6. \tag{34}$$

В силу неравенства (26) имеем

$$k_6 > np + z. \tag{35}$$

Перейдем к рассмотрению функций S и T, определенных в теореме А. Нам будет удобно использовать следующее эквивалентное определение этих функций:

$$S(k) = T(k) = 0 \quad \text{при } k \le 0,$$
$$S(k+1) = \max\{S(k) + f(k), T(k+1-z)\} \quad \text{при } k \ge 0,$$
$$T(k+1) = \max\{T(k) + g(k), S(k+1-z)\} \quad \text{при } k \ge 0.$$

Здесь функции S и T определяются только в целых k, но из теоремы А следует, что в случае целочисленных распределений для всех y, $0 < y < 1$, и всех целых k выполнено

$$S(k-y) = S(k), \quad T(k-y) = T(k).$$

Поэтому далее значения функций S и T в нецелых точках рассматриваться не будут.

Поскольку $\lim\limits_{k\to\infty} S(k) = \lim\limits_{k\to\infty} T(k) = \rho(z, P, Q) + 1 > 1$, конечны следующие значения:

$$l_1 = \min\{k \ : \ T(k-z) > F(k)\}, \quad m_1 = \min\{k \ : \ S(k-z) > G(k)\}.$$

В силу определения l_1 имеем $T(l_1-z) > F(l_1) \ge F(l_1-2z) = S(l_1-2z)$. Поэтому $T(l_1-z) = T(l_1-z-1)+g(l_1-z-1)$. Следовательно, так как $T(l_1-z-1) \le F(l_1-1)$ и $T(l_1-z-1)+g(l_1-z-1) > F(l_1)$, то $f(l_1-1) < g(l_1-1-z)$, а значит, в силу неравенств (32)–(34) выполнено $k_5 < l_1 \le k_6$ и, в частности, $k_5 < k_6$. Из неравенств (32)–(34) также следует, что

$$S(k) = F(k) \quad \text{при } k < l_1,$$

$$S(k) = T(k-z) \quad \text{при } l_1 \le k \le k_6,$$
$$S(k) = F(k) + T(k_6-z) - F(k_6) \quad \text{при } k > k_6.$$

Отметим, что $l_1 \ne m_1$, ибо при всяком k выполняется не более одного из неравенств $F(k-z) > G(k)$, $G(k-z) > F(k)$.

В случае $l_1 < m_1$ имеем

$$m_1 > k_6 + z. \tag{36}$$

Из (36) и (35) вытекает, что $m_1 > np$. Так как $k_3 \le np$ (31), то $m_1 > k_4$. Следовательно,

$$T(k) = G(k) \quad \text{при } k < m_1,$$

$$T(k) = S(k-z) \quad \text{при } k \ge m_1.$$

Если же $l_1 > m_1$, то в силу (27)–(30) либо $k_2 < m_1 \le k_3$, либо $m_1 > k_4$. Последнее неравенство невозможно ввиду того, что если бы $m_1 > k_4$, то выполнялись бы соотношения

$$T(k) = G(k) \quad \text{при } k < m_1,$$

$$T(k) = S(k-z) \quad \text{при } k \ge m_1,$$

$$S(k) = F(k) \quad \text{при всех} \quad k,$$

последнее из которых означало бы, что $l_1 = \infty$. Следовательно, $k_2 < m_1 \le k_3$ и

$$T(k) = G(k) \quad \text{при } k < m_1,$$
$$T(k) = S(k-z) = F(k-z) \quad \text{при } m_1 \le k \le k_3,$$
$$T(k) = G(k) + F(k_3 - z) - G(k_3) \quad \text{при } k_3 \le k < m_2,$$
$$T(k) = S(k-z) \quad \text{при } k \ge m_2,$$

где $m_2 = \min\{k > k_3 \ : \ T(k) > G(k) + F(k_3 - z) - G(k_3)\}$. При этом

$$m_2 > k_6 + z. \tag{37}$$

Оценим $\rho(z,P,Q)$ в терминах $f(k)$. Если $l_1 < m_1$, то

$$S(k) = F(k) + G(k_6 - z) - F(k_6) \quad \text{при } k \ge m_1 - z,$$
$$T(k) = G(k) \quad \text{при } k < m_1,$$
$$T(k) = S(k-z) \quad \text{при } k \ge m_1.$$

Тогда ввиду (36) имеет место оценка

$$\rho(z,P,Q) = \lim_{k\to\infty} T(k) - 1 \le \sum_{k=m_1-z-1}^{\infty} f(k) \le \sum_{k=k_6}^{\infty} f(k).$$

Если же $l_1 > m_1$, то

$$T(k) = G(k) \quad \text{при } k < m_1,$$
$$T(k) = F(k-z) \quad \text{при } m_1 \le k \le k_3,$$
$$T(k) = G(k) + F(k_3 - z) - G(k_3) \quad \text{при } k_3 < k < m_2,$$
$$T(k) = F(k-z) + S(k_6 - z) - F(k_6 - z) \quad \text{при } k \ge m_2.$$

Тогда с учетом (37)

$$\rho(z,P,Q) = \lim_{k\to\infty} T(k) - 1 \le \sum_{k=m_1-1}^{k_3-1} f(k) + \sum_{k=m_2-z-1}^{\infty} f(k) \le \sum_{k=0}^{k_3-1} f(k) + \sum_{k=k_6}^{\infty} f(k).$$

Таким образом,

$$\rho(z,P,Q) \le \sum_{k=0}^{k_3-1} f(k) + \sum_{k=k_6}^{\infty} f(k). \tag{38}$$

В условиях леммы 2 имеем $k_3 \le np - z/p$, $k_6 \ge np + z/3p$ в силу неравенств (31), (16)–(18). Если $\frac{z}{np^2} \ge \frac{1}{2}$, то ввиду (17) $k_3 = 0$. Если же $\frac{z}{np^2} \le \frac{1}{2}$, то $np - z/p \ge \frac{1}{2}np \ge 1$, а значит,

$$\sum_{k=0}^{k_3-1} f(k) \le \frac{2np^2}{z} f(\lfloor np - z/p \rfloor) < \frac{2}{\sqrt{\pi}} \sqrt{\frac{np^3}{z^2}} \exp\left\{-\frac{z}{p} + \left(np - \frac{z}{p}\right) \log\left(1 - \frac{z}{np^2}\right)\right\}$$
$$< \frac{2}{\sqrt{\pi}} \sqrt{\frac{np^3}{z^2}} \exp\left\{-\frac{z}{p} + \left(np - \frac{z}{p}\right)\left(\frac{z}{np^2} + \frac{7}{9}\left(\frac{z}{np^2}\right)^2\right)\right\} \le \frac{2}{\sqrt{\pi}} \sqrt{\frac{np^3}{z^2}} \exp\left\{-\frac{2}{9}\, \frac{z^2}{np^3}\right\}.$$

Аналогично

$$\sum_{k=k_6}^{\infty} f(k) \le (3 + z/np^2) \frac{np^2}{z} f\left(\left\lceil np + \frac{1}{3}\, \frac{z}{np^2} \right\rceil\right)$$

$$< \frac{3}{\sqrt{\pi}}\sqrt{\frac{np^3}{z^2}}\exp\left\{\frac{1}{3}\,\frac{z}{p} - \left(np + \frac{1}{3}\,\frac{z}{p}\right)\left(\frac{1}{3}\,\frac{z}{np^2} - \frac{1}{18}\left(\frac{z}{np^2}\right)^2\right)\right\}$$
$$< \frac{3}{\sqrt{\pi}}\sqrt{\frac{np^3}{z^2}}\exp\left\{-\frac{1}{27}\,\frac{z^2}{np^3}\right\}.$$

Таким образом, в силу (38)

$$\rho(z, P, Q) < \frac{7}{3}\exp\left\{-\frac{1}{27}\,\frac{z^2}{np^3}\right\},$$

т. е. доказано неравенство (7) теоремы.

Далее, в условиях леммы 3 имеем $k_3 = 0$, $k_6 \geq \frac{7}{5}np$. Значит,

$$\sum_{k=k_6}^{\infty} f(k) \leq \frac{7}{2} f\left(\left\lceil \frac{7}{5}\,np \right\rceil\right) < \exp\left\{-\frac{1}{15}\,np\right\}.$$

Последняя оценка и доказывает неравенство (8) теоремы.

Если кроме условий леммы 3 выполнено неравенство $z \geq e^4 np^2$, то ввиду (23) имеем

$$\sum_{k=k_6}^{\infty} f(k) < \frac{7}{6} f\left(\left\lceil np + \frac{1}{2}\sqrt{nz\log\frac{z}{np^2}}\,\right\rceil\right)$$
$$< \exp\left\{\frac{1}{2}\sqrt{nz\log\frac{z}{np^2}} - \left(np + \frac{1}{2}\sqrt{nz\log\frac{z}{np^2}}\right)\log\left(1 + \frac{1}{2}\sqrt{\frac{z}{np^2}\log\frac{z}{np^2}}\right)\right\}$$
$$< \exp\left\{-\frac{1}{2}\sqrt{nz\log\frac{z}{np^2}}\left(\frac{1}{2}\log\frac{z}{np^2} - 1\right)\right\} < \exp\left\{-\frac{1}{4}\sqrt{nz\left(\log\frac{z}{np^2} - 2\right)^3}\right\},$$

что и доказывает неравенство (9) теоремы.

Предположим теперь, что выполнены условия леммы 3 и $z \geq \frac{n(1-p)}{\log\frac{1}{p}}$. Тогда в силу соотношения (24) имеем $k_6 = n + z + 1$,

$$S(n + z + 1) = G(n + 1) = 1,$$
$$S(k) = F(k) - F(n + z + 1) + 1 \quad \text{при } k > n + z + 1.$$

Значит,

$$\lim_{k\to\infty} S(k) - 1 = 1 - F(n + z + 1) = \sum_{k=n+z+1}^{\infty} f(k)$$
$$< \exp\left\{-np + n + z - (n + z)\log\frac{n+z}{np}\right\} < \exp\left\{-(n + z)\log\frac{n+z}{enp}\right\},$$

т. е. верна оценка (10) теоремы. Оценка (11) следует из неравенств

$$f(n + z + 1) < 1 - F(n + z + 1) < \frac{1}{1-p} f(n + z + 1).$$

Установим теперь оценку снизу для $\rho(z, P, Q)$ (11). В условиях леммы 4, используя неравенства (25) и (14), получим

$$\rho(z, P, Q) \geq 1 - F\left(\left\lfloor np + 2\sqrt{nz\left(\log\frac{z}{np^2} + 4\right)}\right\rfloor + z\right)$$

$$-\left(1-G\left(\left\lfloor np+2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}\right\rfloor\right)\right)\geq\frac{1}{4}g\left(\left\lfloor np+2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}\right\rfloor\right)$$
$$\geq\frac{1}{4}\exp\left\{-\frac{4z\left(\log\frac{z}{np^2}+4\right)}{1-p}\right\}f\left(\left\lfloor np+2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}\right\rfloor\right)$$
$$>\frac{1}{11}\exp\left\{\left(-\frac{4z\left(\log\frac{z}{np^2}+4\right)}{1-p}-\left(np+2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}\right)\right.\right.$$
$$\times\log\left(2\sqrt{\frac{z}{np^2}\left(\log\frac{z}{np^2}+4\right)}\right)\Bigg)$$
$$\left.+\left(2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}-2\log\left(np+2\sqrt{nz\left(\log\frac{z}{np^2}+4\right)}\right)\right)\right\}$$
$$>\exp\left\{-\frac{40}{9}z\left(\log\frac{z}{np^2}+4\right)-3\sqrt{nz\left(\log\frac{z}{np^2}+4\right)^3}\right\}$$
$$>\exp\left\{-4\sqrt{nz\left(\log\frac{z}{np^2}+4\right)^3}\right\}.$$

Теорема доказана.

Автор выражает искреннюю благодарность И. С. Борисову за постановку задачи и полезные замечания.

ЛИТЕРАТУРА

1. *Dudley R. M.* Distances of probability measures and random variables // Ann. Math. Stat. 1968. V. 39, N 5. P. 1563–1572.
2. *Прохоров Ю. В.* Сходимость случайных процессов и предельные теоремы теории вероятностей // Теория вероятностей и ее применения. 1956. Т. 1, № 2. С. 177–237.
3. *Рузанкин П. С.* Решение задачи Монжа — Канторовича для одного класса функционалов // Докл. РАН. 2000. Т. 370, № 4. С. 443–444.
4. *Прохоров Ю. В.* Асимптотическое поведение биномиального распределения // Успехи мат. наук. 1953. Т. 8, № 3. С. 135–142.
5. *Le Cam L.* Remargues sur le théorème limit central dans les espaces localement convexes // Les Probabilitiés sur les Structures Algébriques. Paris: C.N.R.S., 1970. P. 233–249.
6. *Barbour A. D., Hall P.* On the rate of Poisson convergence // Math. Proc. Cambridge Philos. Soc. 1984. V. 95. P. 473–480.
7. *Major P.* A note on the approximation of the uniform empirical process // Ann. Probab. 1990. V. 18, N 1. P. 129–139.
8. *Horváth L.* A note on the rate of Poisson approximation of empirical processes // Ann. Probab. 1990. V. 2. P. 724–726.
9. *Borisov I. S.* Strong Poisson and mixed approximations of sums of independent random variables in Banach spaces // Siberian Adv. Math. 1993. V. 3, N 2. P. 1–13.
10. *Adell J. A., Le Cal J. de* Coupling methods in Poisson approximation of binomial processes. Zaragoza, 1994 (Preprint / Univ. de Zaragoza).
11. *Борисов И. С.* Пуассоновская аппроксимация процесса частных сумм в банаховых пространствах // Сиб. мат. журн. 1996. Т. 37, № 4. С. 723–731.

Статья поступила 29 сентября 2000 г.,

Рузанкин Павел Сергеевич
Институт математики им. С. Л. Соболева СО РАН, Новосибирск 630090
ruzankin@math.nsc.ru