К.А. Татосова, В.Е. Мосягин
Тюменский Государственный Университет, г.Тюмень

УДК 519.216

ПОСТРОЕНИЕ НА ОБЩЕМ ВЕРОЯТНОСТНОМ ПРОСТРАНСТВЕ ПУАССОНОВСКОЙ И БИНОМИАЛЬНОЙ ВЕЛИЧИН

Аннотация. С помощью метода одного вероятностного пространства установлено, что эмпирический процесс, построенный по выборке из равномерного на интервале [0, 1] распределения, можно так определить на общем вероятностном пространстве совместно со стандартным пуассоновским процессом, чтобы их траектории совпадали в логарифмической зоне $[0,\ \ln n]$.

Ключевые слова: Метод одного вероятностного пространства, теорема Пуассона, пуассоновский процесс, эмпирический процесс, локальная аппроксимация эмпирического процесса.

§1. Построение на одном вероятностном пространстве

Пусть B_n – случайная величина с биномиальным распределением:

$P(B_n = k) = C_n^k p^k (1-p)^{n-k}, \qquad k = 0,1,\ldots,n,$

а Π_n – случайная величина, имеющая пуассоновское распределение с параметром np:

$$P(\Pi_n = k) = \frac{(np)^k e^{-np}}{k!}.$$

Хорошо известно [1-2] следующее утверждение.

Теорема Пуассона. Если $p = p_n \to 0, np = O(1)$, то справедлива следующая оценка для расстояния по вариации:

$$Sup|P(B_n \in A) - P(\Pi_n \in A)| \le np^2, \tag{1}$$

где супремум берется по всем подмножествам числовой прямой.

Заметим, что из условия теоремы вытекает, что оценка скорости сходимости в (1) имеет порядок $O(1/n)$.

При доказательстве оценки (1) применяется конструкция, называемая «методом одного вероятностного пространства» [1, с. 124]. Суть этого метода заключается в том, что на общем вероятностном пространстве определяются биномиальная и пуассоновская величины так, чтобы

$$P(B_n \neq \Pi_n) \leq np^2 . \qquad (2)$$

В следующем параграфе величины B_n и Π_n будут так определены на общем вероятностном пространстве, чтобы выполнялось:

$$P(|\Pi_n - B_n| > 1) \leq np^3/6 + n^2p^4/4 . \qquad (3)$$

Другими словами, при выполнении условий теоремы Пуассона случайные величины B_n и Π_n могут отличаться друг от друга не более чем на 1 с вероятностью не меньшей, чем $1 - O(1/n^2)$.

§2. Точность пуассоновской аппроксимации

Идея изложенного ниже построения на общем вероятностном пространстве была подсказана нам А.И. Саханенко.

Пусть $\{\alpha_i\}, \{\beta_i\}, \{X_i\}$ – независимые между собой последовательности, состоящие из независимых случайных величин, определенные на общем вероятностном пространстве. Случайные величины α_i и X_i имеют равномерное распределение $\alpha_i, X_i \in \mathbb{U}[0,1]$, а величины β_i имеют «испорченное» пуассоновское распределение:

$$P(\beta_i = k) = \frac{p^{k-1}e^{-p}}{k!}, \qquad k = 1,2,\dots$$
$$P(\beta_i = 0) = 1 - \sum\nolimits_{k=1}^{\infty} \frac{p^{k-1}e^{-p}}{k!}. \qquad (4)$$

Определим последовательность величин $\gamma_i = I\{x_i \leq p\}$, $i = 1,2,\dots$, где $I(A)$ – индикаторная функция события A. Тогда каждая из величин γ_i последовательности распределена по закону Бернулли: $\gamma_i \in \mathbb{B}_p$,

$$p(\gamma_i = 1) = p, \;\; p(\gamma_i = 0) = 1 - p . \qquad (5)$$

Просуммировав γ_i по всем i от 1 до n, получим величину, имеющую биномиальное распределение с параметрами (p, n):

$$B_n = \sum\nolimits_{i=1}^{n} \gamma_i \in \mathbb{B}_p^n . \qquad (6)$$

Утверждение 1. Случайные величины

$$\Pi_n = \sum_{i=1}^{n} \beta_i \gamma_i \qquad (7)$$

имеют распределение Пуассона с параметром np .

Доказательство. Из (4) и (5) получаем

$$P(\beta_i\gamma_i = k) = P(\beta_i = k;\ \gamma_i = k) = P(\beta_i = k)P(\gamma_i = k) =$$

$$\frac{p^{k-1}e^{-p}}{k!}p = \frac{p^k e^{-p}}{k!} \in \text{Ш}_p \ .$$

Ввиду независимости β_i и γ_i случайная величина Π_n будет иметь распределение Пуассона с параметром np.

Утверждение 2. При $p \to 0$ справедлива оценка

$$P(|\Pi_n - B_n| > 1) \le np^3/6 + n^2p^4/4 \ . \qquad (8)$$

Доказательство. Рассмотрим разность определенных в (6) и (7) пуассоновской и биномиальной величин

$$\Pi_n - B_n = \sum_{i=1}^{n} \beta_i\gamma_i - \sum_{i=1}^{n} \gamma_i =$$

$$\gamma_1(\beta_1 - 1) + \gamma_2(\beta_2 - 1) + \cdots + \gamma_n(\beta_n - 1) = \delta_1 + \delta_2 + \cdots \delta_n \ ,$$

где $\delta_i = \gamma_i(\beta_i - 1)$. Величины δ_i могут принимать значения -1, 0, 1, либо значения, большие единицы. Оценим вероятности наступления каждого из четырех событий, учитывая, что γ_i и $(\beta_i - 1)$ являются независимыми.

$$P(\delta_i = 1) = P(\gamma_i = 1)P[(\beta_i - 1) = 1] =$$

$$P(\gamma_i = 1)P(\beta_i = 2) = \frac{p^2 e^{-p}}{2}, \qquad (9)$$

$$P(\delta_i = -1) = P(\gamma_i = 1)P[(\beta_i - 1) = -1] =$$

$$P(\gamma_i = 1)P(\beta_i = 0) = pP(\beta_i = 0) = p\left(\sum_{k=1}^{\infty} \frac{p^{k-1}e^{-p}}{k!}\right) =$$

$$p\left[1 - \frac{e^{-p}}{p}\left(\sum_{k=0}^{\infty} \frac{p^k}{k!} - \frac{p^0}{k}\right)\right] = p\left[1 - \frac{e^{-p}}{p}(e^p - 1)\right] \qquad (10)$$

$$= p - 1 + e^{-p} \ ,$$

$$P(\delta_i = 0) = P(\gamma_i = 0) + P(\gamma_i = 1)P[(\beta_i - 1) = 0] =$$
$$= (1 - p) + pP(\beta_i = 1) = 1 - p + pe^{-p}\,, \tag{11}$$

$$P(\delta_i \geq 2) = 1 - P(\delta_i = 1) - P(\delta_i = -1) - P(\delta_i = 0) =$$
$$= 1 - \left(\frac{p^2 e^{-p}}{2} + e^{-p} + pe^{-p}\right). \tag{12}$$

Так как $p \to 0$, воспользуемся разложением

$$e^{-p} = 1 - p + \frac{p^2}{2} - \frac{p^3}{6} + O(p^4)\,,$$

значит,

$$P(\delta_i \geq 2) \leq 1 - \left[\left(\frac{p^2}{2} - \frac{p^3}{2}\right) + \left(1 - p + \frac{p^2}{2} - \frac{p^3}{6}\right) + \left(p - p^2 + \frac{p^3}{2}\right)\right];$$
$$P(\delta_i \geq 2) \leq \frac{p^3}{6}\,. \tag{13}$$

Используя (9)-(12), получим оценку вероятности в (8):

$$P(|\Pi_n - B_n| > 1) = P(\text{хотя бы две } \delta_i = 1) +$$
$$P(\text{хотя бы две } \delta_i = -1) + P(\text{хотя бы одна } \delta_i \geq 2)\,. \tag{14}$$

Из (13) следует, что

$$P(\text{хотя бы одна } \delta_i \geq 2) = nP(\delta_i \geq 2) \leq \frac{np^3}{6}\,. \tag{15}$$

Всего пар элементов C_n^2, тогда, с учетом (9) и (10):

$$P(\text{хотя бы две } \delta_i = 1) = C_n^2 P(\delta_i = 1) P\left(\delta_j = 1\right) =$$
$$\frac{n!}{(n-2)!\,2!}\left(\frac{p^2 e^{-p}}{2}\right) \leq \frac{n^2 p^4}{8} - \frac{np^4}{8}\,, \tag{16}$$

$$P(\text{хотя бы две } \delta_i = -1) = C_n^2 P(\delta_i = -1) P\left(\delta_j = -1\right) =$$
$$\frac{n!}{(n-2)!\,2!}(p - 1 + e^{-p}) \leq \frac{n^2 p^4}{8} - \frac{np^4}{8}\,. \tag{17}$$

Из (14)-(17) устанавливаем справедливость (8):

$$P(|\Pi_n - B_n| < 1) = C_n^2 P(\delta_i = 1) P\left(\delta_j = 1\right) + C_n^2 P(\delta_i = -1) P\left(\delta_j = -1\right) +$$
$$+nP(\delta_i \geq 2) \leq np^3/6 + n^2 p^4/4\,.$$

§3. Локальная аппроксимация эмпирического процесса

Пусть $\{T_n\}_{n=1,2,\dots}$ - неограниченно возрастающая числовая последовательность с условием: $T_n/n \to 0$ при $n \to \infty$.

Определим случайные процессы:

$$E_n^*(t) = \sum_{i=1}^{B_n} I\{\alpha_i \le t/T_n\},$$

$$\Pi_n^*(t) = \sum_{i=1}^{\Pi_n} I\{\alpha_i \le t/T_n\}, \qquad 0 \le t \le T_n, \tag{18}$$

считая их равными нулю при $t > T_n$. Здесь B_n и Π_n - биномиальная и пуассоновская случайные величины, построенные на общем вероятностном пространстве в §2. Из курса теории случайных процессов хорошо известно, что $E_n^*(t)$ является сужением на интервал $[0, T_n]$ эмпирического процесса, построенного по выборке из равномерного на интервале $[0,1]$ распределения, а процесс $\Pi_n^*(t)$ является стандартным пуассоновским на интервале $[0, T_n]$.

Из утверждения 2 вытекает следующая

Теорема 1. Если $T_n = \sqrt[4]{n}$, то

$$E_n^*(t) = \Pi_n^*(t) + \Delta_n(t), \qquad 0 \le t \le T_n \tag{19}$$

с вероятностью $1 - O(1/n)$, где

$$\Delta_n(t) = I\{\alpha_{\Pi_n+1} \le t/T_n\}, \qquad \text{если } B_n = \Pi_n + 1,$$

$$\Delta_n(t) = -I\{\alpha_{\Pi_n} \le t/T_n\}, \qquad \text{если } B_n = \Pi_n - 1,$$

и $\Delta_n(t) = 0$ в остальных случаях.

Применяя теорему 1, можно доказать основное утверждение работы.

Теорема 2. Эмпирический процесс, построенный по выборке из равномерного $\mathbb{U}[0,1]$ распределения и стандартный пуассоновский процесс можно так задать на общем вероятностном пространстве, чтобы они потраекторно совпадали на интервале $[0, \ln n]$ с вероятностью $1 - O(1/n)$.

СПИСОК ЛИТЕРАТУРЫ

1. Боровков А.А. Теория вероятностей. М.: Наука, 1976. 352 с.

2. Ширяев А.Н. Вероятность. М.: Наука 1980. 480 с.