Journal of Siberian Federal University. Mathematics & Physics 1 (2008) 68-77

УДК 517.55

# Многомерные дискретные распределения

**Олег Ю.Воробьев***
**Лаврентий С.Головков** †
Институт вычислительного моделирования СО РАН,
Академгородок, 660036, Красноярск,
Россия

Получена 10.10.2007, принята 10.11.2007

*Вводятся два новых дискретных распределения: многомерное биномиальное распределение и многомерное распределение Пуассона. Эти распределения были созданы в рамках эвентологии как более корректные обобщения биномиального и пуассоновского распределений. Согласно эвентологии новые законы учитывают полные распределения вероятностей событий. Также в статье описываются их свойства и характеристики.*

Ключевые слова: *многомерное биномиальное распределение, многомерное распределение Пуассона, полное распределение вероятностей событий.*

## Введение

Распределение вероятностей — одно из центральных понятий теории вероятностей и математической статистики. Его определение равносильно заданию всех случайных событий, относящихся к выбранному явлению. Однако результаты испытаний крайне редко выражаются одним числом, и намного чаще — системой чисел, вектором, функцией. Если некая закономерность описывается несколькими случайными величинами, заданными на одном и том же вероятностном пространстве, то говорят о многомерном распределении. Таким образом, последние вводятся для отображения поведения случайного вектора, который служит для описания более или менее приближенных к реальности случайных явлений. В связи с возникшей научной необходимостью задания двух многомерных дискретных распределений и появилась эта работа.

Используемое в настоящее время полиномиальное распределение как обобщение биномиального не учитывает, фактически, основного специфического понятия теории вероятностей, а именно *независимости* событий, случайных величин, испытаний. Ведь без этого свойства теория вероятностей может рассматриваться лишь как часть общей теории меры множеств[1].

## 1. Биномиальное многомерное распределение

Пусть проводится конечная последовательность из $n$ независимых случайных экспериментов. В результате $i$-го эксперимента могут наступить или нет события из $N$-множества $\mathfrak{X}^{(\mathrm{i})}$ событий $x^{(i)} \in \mathfrak{X}^{(\mathrm{i})}$. Эвентологические распределения множеств событий $\mathfrak{X}^{(\mathrm{i})}, i = 1, ..., n$, совпадают с одним и тем же эвентологическим распределением $\{p(X), X \subseteq \mathfrak{X}\}$ некоторого $N$-множества $\mathfrak{X}$ событий $x \in \mathfrak{X}$, которое не меняется от эксперимента к эксперименту.

Такая схема проведения экспериментов называется *многомерной (эвентологической) схемой испытаний Бернулли с порождающим множеством событий* $\mathfrak{X}$, и каждая из случайных

*e-mail: vorob@akadem.ru
†e-mail: lavrentiy.golovkov@gmail.com


[1]Случайные величины предстают как измеримые функции, а их математические ожидания, дисперсии и другие моменты — как абстрактные интегралы Лебега и так далее.

величин

$$\xi_x(\omega) = \sum_{i=1}^{n} \mathbf{1}_{x^{(i)}}(\omega), x^{(i)} \in \mathfrak{X}^{(\mathrm{i})}, x \in \mathfrak{X}$$

подчиняется биномиальному распределению с параметрами $n,\ p_x = \mathbf{P}(x)$, а случайный вектор[2] $\hat{\xi} = (\xi_x, x \in \mathfrak{X})$ подчиняется *биномиальному многомерному ($N$-мерному) распределению* с параметрами $(n, \{p(X), \varnothing \neq X \subseteq \mathfrak{X}\})$.

Вероятности *биномиального многомерного распределения, порожденного $N$-множеством событий* $\mathfrak{X}$, определяются для любого целочисленного набора $\hat{n} = (n_x, x \in \mathfrak{X}) \in [0, n]^N$ формулой

$$b_{\hat{n}}\big(n; p(X), \varnothing \neq X \subseteq \mathfrak{X}\big) = \mathbf{P}(\hat{\xi} = \hat{n}) = \mathbf{P}(\xi_x = n_x, x \in \mathfrak{X}) = \sum_{\check{n}} m_{\check{n}}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big),$$

где

$$m_{\check{n}}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big) = \mathbf{P}(\check{\xi} = \check{n}) = \mathbf{P}\Big(\big(\xi(X), X \subseteq \mathfrak{X}\big) = \big(n(X), X \subseteq \mathfrak{X}\big)\Big) =$$

$$= \frac{n!}{\prod\limits_{X \subseteq \mathfrak{X}} n(X)!} \prod_{X \subseteq \mathfrak{X}} [p(X)]^{n(X)}$$

— вероятности $2^N$-мерного полиномиального распределения случайного вектора $\hat{\xi} = \big(\xi(X), X \subseteq \mathfrak{X}\big)$ с параметрами $(n; \{p(X), X \subseteq \mathfrak{X}\})$, порожденного $2^N$-множеством событий-террасок $\{ter(X), X \subseteq \mathfrak{X}\}$, которое взаимно-однозначно соответствует данному биномиальному многомерному распределению; а суммирование производится по всем $2^N$-мерным наборам $\hat{n} = \big(n(X), X \subseteq \mathfrak{X}\big) \in S^{2^N}$ из $2^N$-вершинного симплекса $S^{2^N}$, то есть таким, что

$$n = \sum_{X \subseteq \mathfrak{X}} n(X),$$

но для которых выполнены еще $N$ равенств

$$n_x = \sum_{x \in X} n(X), \quad x \in \mathfrak{X}.$$

## 1.1. Биномиальное одномерное распределение

При $N = 1$ (когда *порождающее множество* $\mathfrak{X} = \{x\}$ — моноплет событий) биномиальное одномерное распределение случайной величины $\xi_x$ совпадает с классическим биномиальным распределением с параметрами $(n; p_x)$. Иначе говоря, вероятности *биномиального одномерного распределения* имеют классический вид:

$$b_{n_x}(n; p_x) = \mathbf{P}(\xi_x = n_x) = C_n^{n_x} p_x^{n_x} (1 - p_x)^{n - n_x},\ 0 \leqslant n_x \leqslant n.$$

## 1.2. Биномиальное двумерное распределение

При $N = 2$ (когда *порождающее множество* $\mathfrak{X} = \{x, y\}$ — дуплет событий) биномиальное двумерное распределение случайного вектора $\hat{\xi} = (\xi_x, \xi_y) = (\xi_x, x \in \mathfrak{X})$ определяется четырьмя параметрами $\big(n; p(x), p(y), p(xy)\big)$, где[3]

$$p(x) = \mathbf{P}(x \cap y^c),\ p(y) = \mathbf{P}(x^c \cap y),\ p(xy) = \mathbf{P}(x \cap y).$$

Вероятности *биномиального двумерного распределения* вычисляются для любого целочисленного вектора $\hat{n} = (n_x, n_y) \in [0, n]^2$ по формуле

[2] В эвентологии вообще и в данном контексте в частности понятие «вектор» используется в расширенном смысле: как *неупорядоченное* конечное множество или *неупорядоченный* конечный набор неких элементов.

[3] Очевидно, что $p(\varnothing) = 1 - p(x) - p(y) - p(xy)$. Ниже используются обозначения: $p_x = \mathbf{P}(x) = p(x) + p(xy)$, $p_y = \mathbf{P}(y) = p(y) + p(xy)$, $\mathrm{Kov}_{xy} = p(xy) - p_x p_y$, $\sigma_x^2 = p_x(1 - p_x)$, $\sigma_y^2 = p_y(1 - p_y)$.

$$b_{\hat{n}}\big(n; p(x), p(y), p(xy)\big) = \mathbf{P}(\hat{\xi} = \hat{n}) = \mathbf{P}(\xi_x = n_x,\ \xi_y = n_y) =$$
$$= \sum_{n(xy)=\max\{0, n_x+n_y-n\}}^{\min\{n_x, n_y\}} m_{\check{n}}\big(n; p(\varnothing), p(x), p(y), p(xy)\big),$$

где

$$m_{\check{n}}\big(n; p(\varnothing), p(x), p(y), p(xy)\big) = \mathbf{P}(\check{\xi} = \check{n}) =$$
$$= \mathbf{P}\Big(\big(\xi(\varnothing), \xi(x), \xi(y), \xi(xy)\big) = \big(n(\varnothing), n(x), n(y), n(xy)\big)\Big) =$$
$$= \frac{n!}{n(\varnothing)!n(x)!n(y)!n(xy)!}[p(\varnothing)]^{n(\varnothing)}[p(x)]^{n(x)}[p(y)]^{n(y)}[p(xy)]^{n(xy)}$$

— вероятности 4-мерного полиномиального распределения случайного вектора $\check{\xi} = \big(\xi(\varnothing), \xi(x), \xi(y), \xi(xy)\big)$ с параметрами $\big(n; p(\varnothing), p(x), p(y), p(xy)\big)$, а суммирование проводится по всем наборам $\check{n} = \big(n(\varnothing), n(x), n(y), n(xy)\big)$ таким, что $n = n(\varnothing) + n(x) + n(y) + n(xy)$, для которых выполнены ещё два равенства, $n_x = n(x) + n(x,y)$, $n_y = n(y) + n(x,y)$, и может быть сведено к суммированию по одному параметру $n(x,y)$ в границах Фреше, поскольку при фиксированных $n_x$ и $n_y$ все величины $n(\varnothing), n(x), n(y), n(xy)$ можно выразить через один параметр, например, $n(xy)$:

$$n(x) = n_x - n(xy),\ n(y) = n_y - n(xy),\ n(\varnothing) = n - n_x - n_y + n(xy),$$

который меняется в пределах границ Фреше:

$$\max\{0, n_x + n_y - n\} \leqslant n(xy) \leqslant \min\{n_x, n_y\}.$$

Формулу можно также записать в виде

$$b_{\hat{n}}\big(n; p(x), p(y), p(xy)\big) = \mathbf{P}(\hat{\xi} = \hat{n}) =$$
$$= [p(\varnothing)]^n\,[\tau(x)]^{n_x}\,[\tau(y)]^{n_y} \sum_{n(xy)=\max\{0, n_x+n_y-n\}}^{\min\{n_x, n_y\}} \mathcal{C}_n^{n(x,y)}(\hat{n})\,[\tau(x,y)]^{n(xy)},$$

где

$$\mathcal{C}_n^{n(x,y)}(\hat{n}) = \frac{n!}{(n - n_x - n_y + n(xy))!(n_x - n(xy))!(n_y - n(xy))!n(xy)!}$$

— *двумерный биномиальный коэффициент*, а

$$\tau(x) = \frac{p(x)}{p(\varnothing)},\ \tau(y) = \frac{p(y)}{p(\varnothing)},\ \tau(x,y) = \frac{p(\varnothing)p(xy)}{p(x)p(y)}$$

— *мультковариации* первого и второго порядка событий $x$ и $y$.

Вектор математических ожиданий биномиального двумерного случайного вектора $(\xi_x, \xi_y)$ равен $(\mathbf{E}\xi_x, \mathbf{E}\xi_x) = (np_x, np_y)$, а его ковариационная матрица выражается через ковариационную матрицу случайного вектора $(\mathbf{1}_x, \mathbf{1}_y)$ индикаторов событий из *порождающего множества* $\mathfrak{X} = \{x, y\}$ и имеет вид

$$\begin{pmatrix} np_x(1-p_x) & n\mathrm{Kov}_{xy} \\ n\mathrm{Kov}_{xy} & np_y(1-p_y) \end{pmatrix} = n\begin{pmatrix} p_x(1-p_x) & \mathrm{Kov}_{xy} \\ \mathrm{Kov}_{xy} & p_y(1-p_y) \end{pmatrix}$$

Ковариационная матрица центрированного и нормированного биномиального двумерного случайного вектора

$$\left(\frac{\xi_x - np_x}{\sigma_x}, \frac{\xi_y - np_y}{\sigma_y}\right)$$

выражается через ковариационную матрицу случайного вектора

$$\left(\frac{\mathbf{1}_x - p_x}{\sigma_x}, \frac{\mathbf{1}_y - p_y}{\sigma_y}\right)$$

центрированных и нормированных индикаторов событий из $\mathfrak{X} = \{x, y\}$ и имеет вид

$$\begin{pmatrix} n & n\rho_{xy} \\ n\rho_{xy} & n \end{pmatrix} = n \begin{pmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{pmatrix},$$

где $\rho_{xy} = \frac{\mathrm{Cov}_{xy}}{\sigma_x \sigma_y}$ — коэффициент корреляции случайных величин $\mathbf{1}_x$ и $\mathbf{1}_y$ — индикаторов событий из $\mathfrak{X} = \{x, y\}$.

## 1.3. Характеристики биномиального многомерного распределения

Вектор математических ожиданий биномиального многомерного случайного вектора $(\xi_x, x \in \mathfrak{X})$ равен $(\mathbf{E}\xi_x, x \in \mathfrak{X}) = (np_x, x \in \mathfrak{X})$, а его ковариационная матрица выражается через ковариационную матрицу случайного вектора $(\mathbf{1}_x, x \in \mathfrak{X})$ индикаторов событий из *порождающего множества* $\mathfrak{X}$ и имеет вид

$$\begin{pmatrix} n\sigma_x^2 & \dots & n\mathrm{Kov}_{xy} \\ \dots & \dots & \dots \\ n\mathrm{Kov}_{xy} & \dots & n\sigma_y^2 \end{pmatrix} = n \begin{pmatrix} \sigma_x^2 & \dots & \mathrm{Kov}_{xy} \\ \dots & \dots & \dots \\ \mathrm{Kov}_{xy} & \dots & \sigma_y^2 \end{pmatrix},$$

где $\sigma_x^2 = p_x(1 - p_x)$, а $\mathrm{Kov}_{xy} = -p_x p_y$ при $x \neq y$.

Ковариационная матрица порожденного разбиением центрированного и нормированного биномиального многомерного случайного вектора

$$\left(\frac{\xi_x - np_x}{\sigma_x}, x \in \mathfrak{X}\right)$$

выражается через ковариационную матрицу случайного вектора

$$\left(\frac{(\mathbf{1}_x - p_x)}{\sigma_x},\ x \in \mathfrak{X}\right)$$

центрированных и нормированных индикаторов событий из $\mathfrak{X}$ и имеет вид

$$\begin{pmatrix} n\sigma_x^2 & \dots & n\rho_{xy} \\ \dots & \dots & \dots \\ n\rho_{xy} & \dots & n\sigma_y^2 \end{pmatrix} = n \begin{pmatrix} \sigma_x^2 & \dots & \rho_{xy} \\ \dots & \dots & \dots \\ \rho_{xy} & \dots & \sigma_y^2 \end{pmatrix},$$

где $\rho_{xy} = \dfrac{\mathrm{Cov}_{xy}}{\sigma_x \sigma_y} = -\dfrac{p_x p_y}{\sigma_x \sigma_y}$ — коэффициент корреляции случайных величин $\mathbf{1}_x$ и $\mathbf{1}_y$ — индикаторов событий из $\mathfrak{X}$.

## 1.4. Полиномиальное распределение — частный случай биномиального многомерного распределения, когда последнее порождается разбиением пространства элементарных событий

Когда *порождающее $N$-множество* $\mathfrak{X}$ составлено из событий, образующих разбиение $\Omega = \sum\limits_{x \in \mathfrak{X}} x$, *биномиальное многомерное распределение* случайного вектора $\hat{\xi} = (\xi_x, x \in \mathfrak{X})$ определяется $N$ параметрами[4] $(n; p_x, x \in \mathfrak{X})$, где $p_x = \mathbf{P}(x)$, $\sum\limits_{x \in \mathfrak{X}} p_x = 1$, и представляет из себя полиномиальное распределение с данными параметрами (см. Рис. 1).

[4] Поскольку $\sum\limits_{x \in \mathfrak{X}} p_x = 1$, то среди $N$ вероятностей только $N - 1$ можно выбирать независимо.

Отсюда вероятности *биномиального многомерного распределения, порожденного разбиением* $\Omega$, определяются для любого целочисленного вектора $\hat{n} = (n_x, x \in \mathfrak{X})$ из симплекса $\mathcal{S}^N$ (так как $\sum_{x\in\mathfrak{X}} n_x = n$) такой же формулой, что и вероятности соответствующего полиномиального распределения

$$b_{\hat{n}}(n; p_x, x \in \mathfrak{X}) = \mathbf{P}(\hat{\xi} = \hat{n}) = \mathbf{P}(\xi_x = n_x, x \in \mathfrak{X}) = \frac{n!}{\prod\limits_{x\in\mathfrak{X}} n_x!} \prod_{x\in\mathfrak{X}} [p_x]^{n_x}.$$

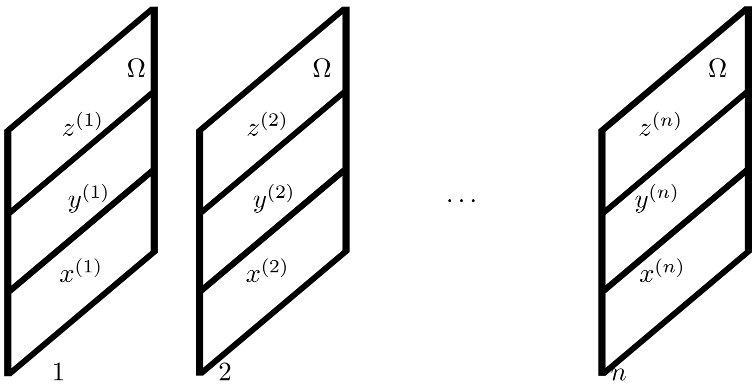


Рис. 1. Полиномиальная схема испытаний Бернулли, определяющая полиномиальное распределение с параметрами $(n;\ p_x, p_y, p_z)$, порождаемое триплетом событий $\mathfrak{X} = \{x, y, z\}$

## 1.5. Биномиальное $N$-мерное распределение, порожденное множеством $\mathfrak{X}$, определяет полиномиальное $2^N$-мерное распределение, порожденное множеством событий-террасок $\{\mathrm{ter}(X),\ X \subseteq \mathfrak{X}\}$, но не наоборот

*Многомерная ($N$-мерная) схема $n$ испытаний Бернулли с порождающим множеством событий* $\mathfrak{X}$, которое подчиняется эвентологическому распределению $\{p(X),\ X \subseteq \mathfrak{X}\}$, определяет $N$ случайных величин

$$\xi_x(\omega) = \sum_{i=1}^{n} \mathbf{1}_{x^{(i)}}(\omega),\ x^{(i)} \in \mathfrak{X}^{(i)},\ x \in \mathfrak{X},$$

каждая из которых имеет биномиальное распределение с параметрами $\big(n; p_x = \mathbf{P}(x)\big)$, а все вместе образуют $N$-мерный случайный вектор $\hat{\xi} = (\xi_x,\ x \in \mathfrak{X})$, распределенный по *биномиальному многомерному ($N$-мерному) закону* с $2^N$ параметрами $\big(n; \{p(X), \varnothing \neq X \subseteq \mathfrak{X}\}\big)$, в состав которых входит число испытаний $n$ и $2^N - 1$ вероятность из эвентологического распределения порождающего множества событий $\mathfrak{X}$ (иначе говоря, все $2^N$ вероятностей $p(X)$ за исключением одной: $p(\varnothing)$) (см. Рис. 2).

Та же многомерная схема $n$ испытаний Бернулли определяет $2^N$ случайных величин

$$\xi(X)(\omega) = \sum_{i=1}^{n} \mathbf{1}_{\mathrm{ter}(X^{(i)})}(\omega),\ X^{(i)} \subseteq \mathfrak{X}^{(i)},\ X \in \mathfrak{X},$$

каждая из которых имеет биномиальное распределение с параметрами $\big(n; p(X) = \mathbf{P}(\mathrm{ter}(X))\big)$, а все вместе образуют $2^N$-мерный случайный вектор $\check{\xi} = \big(\xi(X),\ X \subseteq \mathfrak{X}\big)$, распределенный по *полиномиальному многомерному ($2^N$-мерному) закону*, порожденному множеством событий-террасок $\{\mathrm{ter}(X), X \subseteq \mathfrak{X}\}$ и определяемому $2^N + 1$ параметром $(n; \{p(X),\ X \subseteq \mathfrak{X}\}$, в состав которых входит число испытаний $n$ и все $2^N$ вероятностей $p(X)$ из эвентологического распределения порождающего множества событий $\mathfrak{X}$.

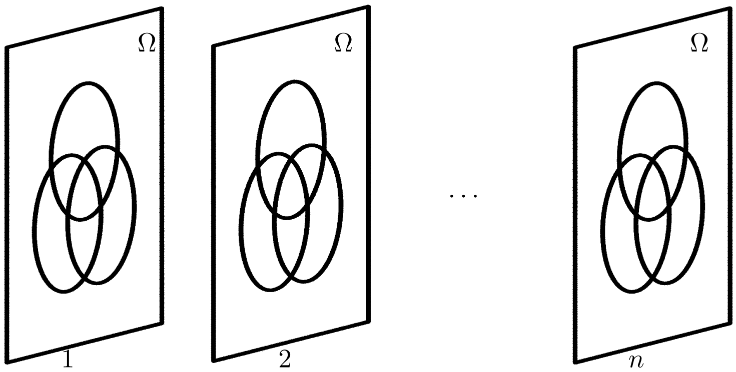


Рис. 2. Трехмерная схема $n$ испытаний Бернулли, порожденная триплетом событий $\mathfrak{X} = \{x, y, z\}$, определяющая 3-мерный случайный вектор $\hat{\xi} = (\xi_x, x \in \mathfrak{X})$, распределенный по биномиальному 3-мерному закону с параметрами $\big(n;\ p(x), p(y), p(z), p(xy), p(xz), p(yz), p(xyz)\big)$; а также 8-мерный случайный вектор $\check{\xi} = \big(\xi(\varnothing), \xi(x), \xi(y), \xi(z), \xi(xy), \xi(xz), \xi(yz), \xi(xyz)\big)$, распределенный по полиномиальному 8-мерному закону с теми же параметрами, что и у биномиального 3-мерного распределения

Вероятности данных *биномиального и полиномиального многомерных распределений* связаны для любых $N$-мерных наборов неотрицательных чисел $\hat{n} = \{n_x, x \in \mathfrak{X}\} \in [0, n]^N$ формулой $\mathbf{P}(\hat{\xi} = \hat{n}) = \sum_{\check{n}} \mathbf{P}(\check{\xi} = \check{n})$, где суммирование проводится по всем $2^N$-мерным наборам неотрицательных чисел $\check{n} = \big(n(X),\ X \subseteq \mathfrak{X}\big) \in \mathcal{S}^{2^N}$ таким, что $n = \sum_{X \subseteq \mathfrak{X}} n(X)$, для которых также выполнены $N$ равенств $n_x = \sum_{x \in X} n(X),\ x \in \mathfrak{X}$.

**Замечание 1.** *Любому биномиальному $N$-мерному распределению, порожденному множеством событий $\mathfrak{X}$, соответствует* единственное *полиномиальное $2^N$-мерное распределение, порожденное множеством соответствующих событий-террасок $\{\mathrm{ter}(X), X \subseteq \mathfrak{X}\}$. Обратное не верно: произвольному полиномиальному $2^N$-мерному распределению, порожденному $2^N$-множеством событий, образующих разбиение $\Omega$, соответствуют, вообще говоря, $(2^N)!$ биномиальных $N$-мерных распределений, порожденных $N$-множествами событий $\mathfrak{X}$, составленных из событий разбиения как из событий-террасок, что существенно зависит от способов отнесения событий разбиения к подмножествам $X \subseteq \mathfrak{X}$ (общее число таких способов как раз и равно $(2^N)!$).*

## 2. Многомерное распределение Пуассона

Многомерное распределение Пуассона — дискретное распределение вероятностей случайного вектора $\hat{\xi} = (\xi_x,\ x \in \mathfrak{X})$, принимающего значения $\hat{n} = (n_x,\ x \in \mathfrak{X})$ с вероятностями

$$\mathbf{P}(\hat{\xi} = \hat{n}) = \mathbf{P}(\xi_x = n_x,\ x \in \mathfrak{X}) = \pi_{\hat{n}}\big(\lambda(X),\ \varnothing \neq X \subseteq \mathfrak{X}\big) =$$

$$= e^{-\lambda} \sum_{\check{n}} \prod_{X \neq \varnothing} \frac{[\lambda(X)]^{n(X)}}{n(X)!},$$

где суммирование распространяется на такие наборы неотрицательных целых чисел $\check{n} = \big(n(X), \varnothing \neq X \subseteq \mathfrak{X}\big)$, для которых справедливы $N$ равенств $n_x = \sum_{x \in X} n(X),\ x \in \mathfrak{X}$, а $\{\lambda(X),\ \varnothing \neq X \subseteq \mathfrak{X}\}$ — параметры: $\lambda(X)$ — среднее число наступления события-терраски

$$ter(X) = \bigcap_{x \in X} x \bigcap_{x \in X^c} x^c,$$

иначе говоря, среднее число наступления всех событий из $X$ и ни одного из $X^c$.

$$\lambda = \sum_{X \neq \varnothing} \lambda(X)$$

— среднее число наступления хотя бы одного события из $\mathfrak{X}$, иначе говоря, среднее число наступления события $\bigcup\limits_{x\in\mathfrak{X}} x$ (объединения всех событий из $\mathfrak{X}$).

Например, при $\hat{n} = (0,\dots,0)$

$$\mathbf{P}\big(\hat{\xi} = (0,\dots,0)\big) = \mathbf{P}(\xi_x = 0,\ x \in \mathfrak{X}) = e^{-\lambda},$$

при $\hat{n} = (0,\dots,0,n_x,0,\dots,0),\ x \in \mathfrak{X}$,

$$\mathbf{P}\big(\hat{\xi} = (0,\dots,0,n_x,0,\dots,0)\big) = \mathbf{P}(\xi_x = n_x,\ \xi_y = 0, y \neq x) = e^{-\lambda}\frac{[\lambda(x)]^{n(x)}}{n(x)!}.$$

Если же в векторе $\hat{n}$ одна компонента $n_x$ фиксирована, а остальные произвольны: $\hat{n} = (\cdot,\dots,\cdot,n_x,\cdot,\dots,\cdot),\ x \in \mathfrak{X}$, то получаем

$$\mathbf{P}\big(\hat{\xi} = (\cdot,\dots,\cdot,n_x,\cdot,\dots,\cdot)\big) = \mathbf{P}(\xi_x = n_x) = e^{-\lambda_x}\frac{[\lambda_x]^{n_x}}{n_x!}$$

— формулу одномерного распределения Пуассона с параметром $\lambda_x$ случайной величины $\xi_x$, где $\lambda_x = \sum_{x\in X}\lambda(X)$ для каждого $x \in \mathfrak{X}$ определяется параметром многомерного распределения Пуассона.

## 2.1. Эвентологическая интерпретация

Проводится счетное множество независимых экспериментов, в результате каждого $n$-го эксперимента наступают события из множества $\mathfrak{X}$. Вероятности $p_x = \mathbf{P}(x)$ событий $x \in \mathfrak{X}$ малы, то есть малы и вероятности $p(X),\ \varnothing \neq X \subseteq \mathfrak{X}$ порождаемых ими событий-террасок $\mathrm{ter}(X),\ \varnothing \neq X \subseteq \mathfrak{X}$, причем при $n \to \infty$ имеется $np(X) \to \lambda(X),\ \varnothing \neq X \subseteq \mathfrak{X}$. Тогда случайный вектор

$$\hat{\xi} = (\xi_x,\ x \in \mathfrak{X}) = \sum_{n=1}^{\infty} \mathbf{1}_{x^{(n)}},\ x \in \mathfrak{X}$$

подчиняется многомерному ($N$-мерному) распределению Пуассона с параметрами
$\{\lambda(X),\ \varnothing \neq X \subseteq \mathfrak{X}\}$ (см. Рис. 3,4).

**Замечание 2.** *Не верно представлять, что вероятности стремятся к нулю так, что только в одном $n$-м испытании $np(X) = \lambda(X),\quad X \subseteq \mathfrak{X}$. На самом деле, правильнее считать, что стремление вероятностей к нулю таково, что данное соотношение выполнено* во всех первых $n$ испытаниях. *Таким образом, случайный эксперимент заключается в том, что проводится* последовательность $n$-серий независимых испытаний (серий из $n$ испытаний)*, и данное соотношение выполнено для всех испытаний из $n$-серии. Тогда $n$-серия определяет биномиальное многомерное ($N$-мерное) распределение с параметрами $\big(n; p(X), \varnothing \neq X \subseteq \mathfrak{X}\big)$, которое при $n \to \infty$ стремится к многомерному ($N$-мерному) пуассоновскому с параметрами $\big(\lambda(X), \varnothing \neq X \subseteq \mathfrak{X}\big)$.*

## 2.2. Характеристики многомерного распределения Пуассона

Вектор математических ожиданий распределения Пуассона $(\mathbf{E}\xi_x,\ x \in \mathfrak{X}) = (\lambda_x,\ x \in \mathfrak{X})$, где $\lambda_x = \sum\limits_{x\in X}\lambda(X),\ x \in \mathfrak{X}$. Поскольку $\mathrm{Cov}(\xi_x, \xi_y) = \lambda_{xy}$, где $\lambda_{xy} = \sum\limits_{\{x,y\}\subseteq X}\lambda(X)$,
$\{x, y\} \subseteq \mathfrak{X}$, то ковариационная матрица имеет вид

$$\begin{pmatrix} \lambda_x & \dots & \lambda_{xy} \\ \dots & \dots & \dots \\ \lambda_{xy} & \dots & \lambda_y \end{pmatrix}.$$

В двумерном варианте, когда $\mathfrak{X} = \{x, y\}$, суммирование происходит по одному параметру $n(xy) = n(\{x, y\})$, который меняется в пределах границ Фреше:

$$\mathbf{P}(\hat{\xi} = \hat{n}) = \mathbf{P}(\xi_x = n_x, \xi_y = n_y) = \pi_{\hat{n}}\big(\lambda(x), \lambda(y), \lambda(xy)\big) =$$

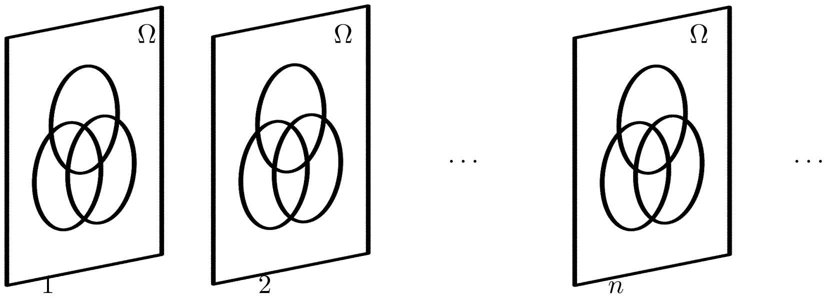


Рис. 3. Схема счетного ряда испытаний Бернулли, порожденная триплетом событий $\mathfrak{X} = \{x, y, z\}$, определяющая 3-мерное распределение Пуассона с параметрами $\big(\lambda(x), \lambda(y), \lambda(z), \lambda(xy), \lambda(xz), \lambda(yz), \lambda(xyz)\big)$

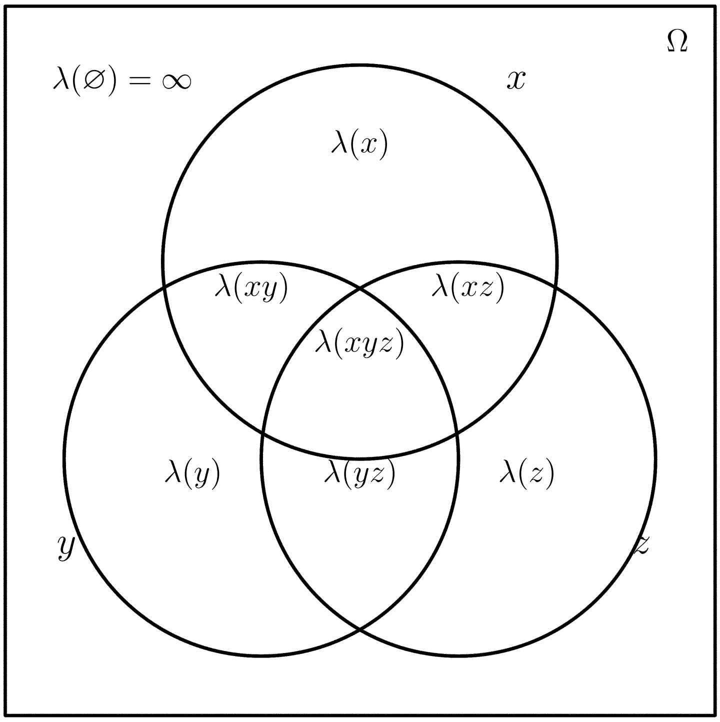


Рис. 4. Параметры $\lambda(x), \lambda(y), \lambda(z), \lambda(xy), \lambda(xz), \lambda(yz), \lambda(xyz)$ 3-мерного распределения Пуассона, порожденного триплетом событий $\mathfrak{X} = \{x, y, z\}$, имеют смысл среднего числа наступления соответствующих событий-террасок $\mathrm{ter}(x), \mathrm{ter}(y), \mathrm{ter}(z), \mathrm{ter}(xy), \mathrm{ter}(xz), \mathrm{ter}(yz), \mathrm{ter}(xyz)$. Величина $\lambda(\varnothing)$, хотя и не считается параметром, по определению полагается равной бесконечности, как предельное число испытаний, в которых не наступило ни одно из событий

$$= e^{-\lambda} \sum_{n(xy)=0}^{\min\{n_x, n_y\}} \frac{[\lambda(x)]^{n(x)}}{n(x)!} \frac{[\lambda(y)]^{n(y)}}{n(y)!} \frac{[\lambda(xy)]^{n(xy)}}{n(xy)!},$$

где $n(x) = n_x - n(xy)$, $n(y) = n_y - n(xy)$, а $\lambda = \lambda(x) + \lambda(y) + \lambda(xy)$.

Вектор математических ожиданий двумерного распределения Пуассона $(\mathbf{E}\xi_x,\ \mathbf{E}\xi_y) = (\lambda_x,\ \lambda_y)$, где $\lambda_x = \lambda(x) + \lambda(xy)$ и $\lambda_y = \lambda(y) + \lambda(xy)$, а ковариационная матрица имеет вид

$$\begin{pmatrix} \lambda_x & \lambda(xy) \\ \lambda(xy) & \lambda_y \end{pmatrix},$$

так как в двумерном варианте $\lambda_{xy} = \lambda(xy)$.

## 2.3. Многомерное пуассоновское приближение

Если количество независимых экспериментов $n$ велико, а вероятности $p_x = \mathbf{P}(x)$ событий $x \in \mathfrak{X}$ малы, то есть малы и вероятности $p(X),\ \varnothing \neq X \subseteq \mathfrak{X}$, порождаемых ими событий-террасок $\mathrm{ter}(X), \varnothing \neq X \subseteq \mathfrak{X}$, то для любого набора целых чисел $\hat{n} = (n_x,\ x \in \mathfrak{X}) \in [0,n]^N$ биномиальные вероятности приближенно выражаются в терминах *многомерного распределения Пуассона*:

$$b_{\hat{n}}\big(n; p(X),\ \varnothing \neq X \subseteq \mathfrak{X}\big) \approx e^{-n\sum_{X\neq\varnothing} p(X)} \sum \prod_{X\neq\varnothing} \frac{[np(X)]^{n(X)}}{n(X)!},$$

где суммирование распространяется на такие наборы $\big(n(X), \varnothing \neq X \subseteq \mathfrak{X}\big)$, для которых $n \geqslant \sum\limits_{X\subseteq\mathfrak{X}} n(X)$, и справедливы $N$ равенств $n_x = \sum\limits_{x\in X} n(X),\ x \in \mathfrak{X}$.

В двумерном варианте, когда $\mathfrak{X} = \{x, y\}$, суммирование происходит по одному параметру $n(xy) = n(\{x,y\})$, который меняется в пределах так называемых границ Фреше:

$$b_{\hat{n}}\big(n; p(x), p(y), p(xy)\big) \approx$$

$$\approx e^{-n(p(x)+p(y)+p(xy))} \sum_{n(xy)=0}^{\min\{n_x,n_y\}} \frac{[np(x)]^{n(x)}}{n(x)!} \frac{[np(y)]^{n(y)}}{n(y)!} \frac{[np(xy)]^{n(xy)}}{n(xy)!},$$

где $n(x) = n_x - n(xy), n(y) = n_y - n(xy)$.

**Теорема Пуассона (многомерный вариант).** *Пусть $p_x \to 0,\ x \in \mathfrak{X}$, при $n \to \infty$, причем $np(X) \to \lambda(X)$ для всех непустых подмножеств $\varnothing \neq X \subseteq \mathfrak{X}$. Тогда для любого набора целых чисел $\hat{n} = (n_x, x \in \mathfrak{X}) \in [0,n]^N$ при $n \to \infty$*

$$b_{\hat{n}}\big(n; p(X), \varnothing \neq X \subseteq \mathfrak{X}\big) \to \pi_{\hat{n}}\big(\lambda(X), \varnothing \neq X \subseteq \mathfrak{X}\big),$$

*где*

$$\pi_{\hat{n}}\big(\lambda(X), \varnothing \neq X \subseteq \mathfrak{X}\big) = e^{-\sum_{X\neq\varnothing}\lambda(X)} \sum_{\check{n}} \prod_{X\neq\varnothing} \frac{[\lambda(X)]^{n(X)}}{n(X)!}$$

— многомерная пуассоновская вероятность*, а суммирование распространяется на такие наборы $\check{n}$, для которых $n_x = \sum_{x\in X} n(X),\ x \in \mathfrak{X}$.*

Доказательство. Так как при больших n условие $n \geqslant \sum_{X\subseteq\mathfrak{X}} n(X)$ выполнено при любых фиксированных $n(X), X \subseteq \mathfrak{X}$, то суммирование в формулах биномиальной

$$b_{\hat{n}}\big(n; p(X), \varnothing \neq X \subseteq \mathfrak{X}\big) = \sum_{\check{n}} m_{\check{n}}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big)$$

и пуассоновской

$$\pi_{\hat{n}}\big(\lambda(X), \varnothing \neq X \subseteq \mathfrak{X}\big) = e^{-\sum_{X\neq\varnothing}\lambda(X)} \sum_{\check{n}} \prod_{X\neq\varnothing} \frac{[\lambda(X)]^{n(X)}}{n(X)!}$$

вероятностей распространяется на одинаковые наборы $\check{n}$, для которых $n_x = \sum_{x\in X} n(X),\ x \in \mathfrak{X}$.

Покажем пуассоновское приближение для полиномиальных вероятностей:

$$m_{\check{n}}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big) = \frac{n!}{\prod\limits_{X\subseteq\mathfrak{X}} n(X)!} \prod_{X\subseteq\mathfrak{X}} [p(X)]^{n(X)}. \tag{1}$$

Заметим, что для любых фиксированных $n(X), X \subseteq \mathfrak{X}$, и достаточно больших $n$ выполнены следующие соотношения:

$$\frac{m_{(n(\varnothing),n(Z),\{n(X),Z\neq X\subseteq\mathfrak{X}\})}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big)}{m_{(n(\varnothing)+1,n(Z)-1,\{n(X),Z\neq X\subseteq\mathfrak{X}\})}\big(n; \{p(X), X \subseteq \mathfrak{X}\}\big)} = \frac{p(Z)\big(n(\varnothing)+1\big)}{n(Z)p(\varnothing)},$$

где $Z \subseteq \mathfrak{X}$. Умножив и разделив числитель и знаменатель на $n$ и учитывая, что $\frac{n(\varnothing)+1}{n} \approx 1$ и $p(\varnothing) \approx 1$, где символ $\approx$ означает приближенное равенство с точностью до членов порядка $n^{-1}$, получаем

$$\frac{p(Z)\big(n(\varnothing)+1\big)}{n(Z)p(\varnothing)}\frac{n}{n} = \frac{np(Z)}{n(Z)} \cdot \frac{n(\varnothing)+1}{n} \cdot \frac{1}{p(\varnothing)} \approx \frac{np(Z)}{n(Z)}.$$

По условию теоремы $np(Z) \to \lambda(Z)$, поэтому

$$\frac{m_{(n(\varnothing),n(Z),\{n(X),Z\neq X\subseteq\mathfrak{X}\})}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big)}{m_{(n(\varnothing)+1,n(Z)-1,\{n(X),Z\neq X\subseteq\mathfrak{X}\})}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big)} \approx \frac{\lambda(Z)}{n(Z)}. \tag{2}$$

При $n(X)=0, \varnothing \neq X \subseteq \mathfrak{X}$,

$$m_{(n(\varnothing),0,\dots,0)}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big) = [p(\varnothing)]^n = \left(1-\sum_{X\subseteq\mathfrak{X}}\frac{\lambda(X)}{n}\right)^n = \left(1-\frac{\lambda}{n}\right)^n,$$

где $\lambda = \sum_{X\subseteq\mathfrak{X}}\lambda(X)$. Логарифмируя обе части уравнения и раскладывая в ряд Маклорена[5], имеем

$$\ln\left[m_{(n(\varnothing),0,\dots,0)}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big)\right] = n\cdot\ln\left(1-\frac{\lambda}{n}\right) = -\lambda-\frac{\lambda^2}{2n}-\dots$$

При больших $n$ заключаем, что

$$m_{(n(\varnothing),0,\dots,0)}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big) \approx e^{-\lambda}. \tag{3}$$

Последовательно применяя соотношение 2 к приближению 3, приходим к

$$m_{(n(\varnothing),\{n(X),X\subseteq\mathfrak{X}\})}\big(n;\{p(X),X\subseteq\mathfrak{X}\}\big) \approx e^{-\lambda}\prod_{\varnothing\neq X\subseteq\mathfrak{X}}\frac{[\lambda(X)]^{n(X)}}{n(X)!}$$

— пуассоновскому приближению полиномиальной вероятности 1, из которого следует утверждение теоремы. $\square$

## Список литературы

[1] О.Ю.Воробьёв, Эвентология, Красноярск, Сибирский федеральный университет, 2007.

## Discrete Multidimensional Distributions

**Oleg Yu.Vorob'ov**
**Lavrenty S.Golovkov**

*Polynomial distribution, which is used at present as a generalization for Binomial, in fact, doesn't take into account the specific notion from the probability theory namely independence of events, random variables, tests. Such a generalization can be used in a special case, that is to say when events don't intersect. Eventology deals with various structures of events' dependencies, therefore it is reasonable that in case of arbitrary structure dependency structure question emerges about more harmonious generalization of the Binomial distribution.*

*Besides there is cite about approximation Binomial multivariate distribution with another new one — multivariate analogue of the Poisson distribution. The article brings in definitions for those new objects and also their main characteristics and properties.*

Key words: *multidimensional polynomial distribution, multidimensional binomial distribution.*

[5] $\ln(1+x) = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{x^k}{k}$.