Теория вероятностей и случайные процессы

УДК 519.2

Двухмерное биномиальное распределение

Ю. С. Хохлов, И. П. Шестаков

Кафедра теории вероятностей и математической статистики
Российский университет дружбы народов
ул. Миклухо–Маклая, д. 6, Москва, 117198, Россия

Рассматривается модель двухмерного случайного вектора с зависимыми, биномиально распределёнными маргиналами, определяется закон распределения построенной модели.

Ключевые слова: натуральное экспонентное семейство распределений, двухмерное биномиальное распределение, производящая функция кумулянтов, дисперсионная функция.

1. Введение

Известно, что биномиальное распределение относится к семейству натуральных экспонентных распределений на $\mathbb{R}^1$ с квадратичной относительно математического ожидания функцией дисперсии [1]. Одномерное распределение хорошо изучено, чего нельзя сказать о случайных векторах, маргиналами которых являются одномерные случайные величины, распределённые согласно биномиальному закону. Изучение осложняется характером зависимости между компонентами вектора, который может быть различным при одних и тех же натуральных параметрах координат.

В данной работе мы рассматриваем особое параметрическое семейство распределений $\mathcal{F} \in \mathbb{R}^2$. Это семейство обладает следующими свойствами:

1) $\mathcal{F}$ — натуральное экспонентное семейство распределений на $\mathbb{R}^2$;
2) распределения маргиналов принадлежат натуральному экспонентному семейству на $\mathbb{R}^1$.

Для начала опишем натуральное экспонентное семейство распределений (НЭС) $\mathcal{F}$ на основе производящей функции кумулянтов (ПФК). Далее формулируется теорема об удобном виде представления ПФК семейства распределений $\mathcal{F}$, приводятся следствия, позволяющие восстанавливать её на основе ПФК маргиналов и функции, характеризующей зависимость между ними. Затем строится двухмерная модель, полезная с точки зрения прикладных задач, на основе биномиально распределённых зависимых случайных величин, определяется их совместное распределение вероятностей.

2. Предварительные сведения

Пусть $X = (X_1, X_2, \ldots, X_n)$ — случайный вектор с множеством значений на $\chi \subset \mathbb{R}^n$ и $\vartheta = (\vartheta_1, \vartheta_2, \ldots, \vartheta_n) \in \Theta \subset \mathbb{R}^n$ — n-мерный параметр.

Введём скалярное произведение на $\mathbb{R}^n$, как $\langle x, \vartheta \rangle = \sum\limits_{j=1}^{n} x_j \vartheta_j$. Для положительной меры μ на χ определим: преобразование Лапласа меры μ при заданном значении параметра ϑ:

$$L_\mu(\vartheta) = \int \exp \langle \vartheta, x \rangle \, \mu(dx) \leqslant +\infty,$$

Статья поступила в редакцию 26 мая 2011 г.

производящую функцию кумулянтов $k_\mu(\vartheta) = \ln L_\mu(\vartheta)$ на $\Theta(\mu)$, где параметрическое множество (естественное параметрическое пространство семейства) есть $\Theta_\mu = (\vartheta \in \Theta : L_\mu(\vartheta) < +\infty)$.

Определим натуральное экспонентное семейство (НЭС), порождённое положительной мерой μ на χ, как распределение вероятностей

$$\mathcal{F}(\mu) = \{P(\vartheta,\mu)(dx) = \exp\{\langle\vartheta, x\rangle - k_\mu(\vartheta)\}\,\mu(dx)\,;\vartheta \in \Theta(\mu)\}\,.$$

Для таких распределений можем определить вектор математических ожиданий и дисперсионную функцию [2] как

$$E_\vartheta X_i = m_i = \frac{\partial k_\mu(\vartheta)}{\partial\vartheta_i}, Var_\vartheta X = (\sigma_{ij}) = \left(\frac{\partial^2 k_\mu(\vartheta)}{\partial\vartheta_i\partial\vartheta_j}\right), i, j = 1,\ldots,n.$$

3. Постановка задачи

Рассмотрим последовательность $\{\varepsilon_j, j \geqslant 1\}$независимых одинаково распределённых случайных величин со значениями во множестве двухмерных индексов I [3], распределение которых задаётся по правилу:

$$P(\varepsilon_j = (1,0)) = p_1, \quad P(\varepsilon_j = (0,1)) = p_2, \quad P(\varepsilon_j = (1,1)) = 1 - p_1 - p_2,$$

где $p_1 \geqslant 0$, $p_2 \geqslant 0$, $p_1 + p_2 < 1$.

Необходимо найти вид совместного распределения вектора $Y = \sum\limits_{j=1}^{n}\varepsilon_j$, т.е.:

1) доказать, что распределение Yпринадлежит НЭС $\mathcal{F}$;
2) найти функцию, характеризующую зависимость между компонентами вектора Y;
3) найти ПФК распределения вектора Y;
4) найти дисперсионную функцию.

Для решения этой задачи необходимо сформулировать и доказать теорему об удобном виде представления ПФК семейства распределений $\mathcal{F}$, позволяющего восстанавливать её на основе ПФК маргиналов и функции, характеризующей зависимость между ними.

4. Получение распределения вероятностей вектора Y

Теорема 1. *Производящая функция кумулянтов двухмерного распределения $\mathcal{F}$ с натуральным параметром $(\vartheta_1, \vartheta_2) \in \mathbb{R}^2$ может быть записана в следующем виде: $k_\mu(\vartheta_1, \vartheta_2) = k_1(\vartheta_1 + \beta_1(\vartheta_2)) + f_1(\vartheta_2) = k_2(\vartheta_2 + \beta_2(\vartheta_1)) + f_2(\vartheta_1)$, где k_i, $i = 1, 2$ — маргинальная производящая функции кумулянтов; β_i, f_i, $i = 1, 2$ — некоторые функции.*

Доказательство. Так как вектор $X = (X_1, X_2)$ принадлежит классу двухмерных НЭС, который описывается с помощью закона

$$P_{\vartheta_1,\vartheta_2}(x_1, x_2) = \exp\{x_1\vartheta_1 + x_2\vartheta_2 - k_\mu(\vartheta_1, \vartheta_2)\}\,\mu(x_1, x_2)\,,$$

где $\mu(x_1, x_2)$ — порождающая мера для двумерного НЭС, то распределение случайной величины X_1, как компоненты вектора X, имеет вид

$$P_1(x_1) = \int P_{\vartheta_1,\vartheta_2}(x_1, x_2)\,dx_2 = \exp\{x_1\vartheta_1 - k_\mu(\vartheta_1, \vartheta_2)\}\int \exp\{x_2\vartheta_2\}\,\mu(x_1, x_2)\,dx_2. \tag{1}$$

С другой стороны, X_1 принадлежит классу одномерных натуральных семейств распределений, порождённых мерой μ_1, т.е.,

$$P_1(x_1) = \exp\{x_1\alpha_1 - k_1(\alpha_1)\}\,\mu_1(x_1)\,, \tag{2}$$

где $\alpha_1 \in \mathbb{R}^1$ — одномерный натуральный параметр.

Сравним правые части соотношений (1) и (2). Интеграл в соотношении (1) не зависит ни от значения x_2(переменной интегрирования), ни от значения параметра ϑ_1, следовательно, учитывая (2), имеем:

$$\int \exp\{x_2\vartheta_2\}\,\mu(x_1, x_2)\,dx_2 = \exp\{x_1\beta_1(\vartheta_2) + f_1(\vartheta_2)\}\,\mu_1(x_1)\,,$$

где $\beta_1(\vartheta_2)$ и $f_1(\vartheta_2)$ — некоторые функции, зависящие только от ϑ_2, а значит, имеют место следующие соотношения между одномерными и двумерными натуральными параметрами и производящими функциями кумулянтов:

$$\begin{cases} \vartheta_1 + \beta_1(\vartheta_2) = \alpha_1(\vartheta_1, \vartheta_2)\,, \\ k_\mu(\vartheta_1, \vartheta_2) - f_1(\vartheta_2) = k_1(\alpha_1)\,. \end{cases} \tag{3}$$

Проводя аналогичные рассуждения относительно компоненты X_2 вектора X, запишем

$$\begin{cases} \vartheta_2 + \beta_2(\vartheta_1) = \alpha_2(\vartheta_1, \vartheta_2)\,, \\ k_\mu(\vartheta_1, \vartheta_2) - f_2(\vartheta_1) = k_2(\alpha_2)\,, \end{cases} \tag{4}$$

где аналогично рассмотренному случаю с компонентой X_1, $\beta_2(\vartheta_1)$ и $f_2(\vartheta_1)$ — функции, зависящие только от ϑ_1. Эти функции характеризуют зависимость между компонентами случайного вектора X.

Таким образом, из (3) и (4) следует, что

$$k_\mu(\vartheta_1, \vartheta_2) = k_1(\vartheta_1 + \beta_1(\vartheta_2)) + f_1(\vartheta_2) = k_2(\vartheta_2 + \beta_2(\vartheta_1)) + f_2(\vartheta_1)\,.$$

Доказательство теоремы завершено. □

Следствие 2. Для того, чтобы по известным распределениям маргиналов восстановить производящую функцию кумулянтов их совместного распределения, необходимо только знать характер зависимости компонент или, что тоже самое, — функции $\beta_1(\vartheta_2)$, $\beta_2(\vartheta_1)$, которыми выражается эта зависимость.

Доказательство. Из теоремы 1 следует, что

$$k_\mu(\vartheta_1, \vartheta_2) = k_1(\vartheta_1 + \beta_1(\vartheta_2)) + f_1(\vartheta_2) = k_2(\vartheta_2 + \beta_2(\vartheta_1)) + f_2(\vartheta_1). \tag{5}$$

Тогда, используя свойство производящей функции кумулянтов НЭС, продифференцируем (5) по ϑ_1 и ϑ_2:

$$m_1(\vartheta_1 + \beta_1(\vartheta_2)) = m_2(\vartheta_2 + \beta_2(\vartheta_1))\,\beta_2'(\vartheta_1) + f_2'(\vartheta_1)\,, \tag{6}$$

$$m_2(\vartheta_2 + \beta_2(\vartheta_1)) = m_1(\vartheta_1 + \beta_1(\vartheta_2))\,\beta_1'(\vartheta_2) + f_1'(\vartheta_2)\,, \tag{7}$$

где $m_i(\alpha_i) = E_{\alpha_i}X_i$, $i = 1, 2$. Поэтому для того чтобы определить $f_1(\vartheta_2)$и $f_2(\vartheta_1)$, нам необходимо знать только функции $\beta_1(\vartheta_2)$, $\beta_2(\vartheta_1)$.

Таким образом, двухмерное распределения $\mathcal{F}$ с натуральным параметром $(\vartheta_1, \vartheta_2) \in \mathbb{R}^2$ определяется функциями $\beta_1(\vartheta_2)$ и $\beta_2(\vartheta_1)$, так как они определяют натуральный параметр и производящую функцию кумулянтов совместного распределения, а последняя является преобразованием Лапласа меры $\mu(x_1, x_2)$.□

Следствие 3. Производящая функция кумулянтов двухмерного распределения $\mathcal{F}$ с натуральным параметром $(\vartheta_1, \vartheta_2) \in \mathbb{R}^2$ может быть записана в следующем виде:

$$k_\mu(\vartheta_1, \vartheta_2) =$$
$$= k_1(\vartheta_1 + \beta_1(\vartheta_2)) + \int m_2(\vartheta_2 + \beta_2(\vartheta_1))\, d\vartheta_2 - \int m_1(\vartheta_1 + \beta_1(\vartheta_2))\, d(\beta_1(\vartheta_2)) =$$
$$k_2(\vartheta_2 + \beta_2(\vartheta_1)) + \int m_1(\vartheta_1 + \beta_1(\vartheta_2))\, d\vartheta_1 - \int m_2(\vartheta_2 + \beta_2(\vartheta_1))\, d(\beta_2(\vartheta_1)).$$

Следствие 4. Если компоненты вектора имеют однотипные распределения, симметричные с точностью до значений натуральных параметров, т.е. $\beta_1(\vartheta) = \beta_2(\vartheta) = \beta(\vartheta)$, $f_1(\vartheta) = f_2(\vartheta) = f(\vartheta)$, $k_1(\vartheta) = k_2(\vartheta) = k(\vartheta)$, то производящая функция кумулянтов двухмерного распределения $\mathcal{F}$ с натуральным параметром $(\vartheta_1, \vartheta_2) \in \mathbb{R}^2$ может быть записана в следующем виде:

$$k_\mu(\vartheta_1, \vartheta_2) = k(\vartheta_1 + \beta(\vartheta_2)) + f(\vartheta_2) = k(\vartheta_2 + \beta(\vartheta_1)) + f(\vartheta_1).$$

Теорема 2. *При сформулированных выше условиях для модели двухмерного случайного вектора* $Y = \sum_{j=1}^{n} \varepsilon_j$ *верны следующие утверждения:*

1) *маргиналы случайного вектора* Y *имеют распределения:* $Y_1 \sim Bi(n; 1 - p_2)$, $Y_2 \sim Bi(n; 1 - p_1)$;
2) *распределение вектора* Y *принадлежит классу НЭС, единственно и имеет следующий вид:*

$$P_Y(y_1, y_2) = \frac{n!}{(n - y_2)!\,(n - y_1)!\,(y_1 + y_2 - n)!} \times$$
$$\times \exp\left\{\vartheta_1' y_1 + \vartheta_2' y_2 - n \cdot \ln\left(e^{\vartheta_1'} + e^{\vartheta_2'} + e^{\vartheta_1' + \vartheta_2'}\right)\right\}$$

для $(\vartheta_1', \vartheta_2') \in \mathbb{R}^2$ *и любых натуральных* y_1 *и* y_2 *таких, что* $y_1 \leqslant n$, $y_2 \leqslant n$, $y_1 + y_2 \geqslant n$, *где* $\vartheta_1' = \ln \frac{1 - p_1 - p_2}{p_2}$, $\vartheta_2' = \ln \frac{1 - p_1 - p_2}{p_1}$, $k_{\mu(y_1, y_2)}(\vartheta_1', \vartheta_2') = n \cdot \ln\left(e^{\vartheta_1'} + e^{\vartheta_2'} + e^{\vartheta_1' + \vartheta_2'}\right)$ *— производящая функция кумулянтов распределения вектора* Y, $\mu(y_1, y_2) = \frac{n!}{(n - y_2)!\,(n - y_1)!\,(y_1 + y_2 - n)!}$ *— порождающая мера;*

3) *распределение вектора* Y *характеризуется при помощи любого из 3 эквивалентных способов:*
   3.1) *«бэта-функция»* $\beta(\vartheta) = \ln\left(1 + e^{-\vartheta}\right)$;
   3.2) *производящая функция кумулянтов* $k_\mu(\vartheta_1; \vartheta_2) = n \ln\left(e^{\vartheta_1} + e^{\vartheta_2} + e^{\vartheta_1 + \vartheta_2}\right)$;
   3.3) *дисперсионная функция* $Var(m) = \begin{pmatrix} m_1 - \frac{m_1^2}{n} & \sigma_{12} \\ \sigma_{21} & m_2 - \frac{m_2^2}{n} \end{pmatrix}$, *где* $m_i = n \cdot \frac{e^{\vartheta_i} + e^{\vartheta_1 + \vartheta_2}}{e^{\vartheta_1} + e^{\vartheta_2} + e^{\vartheta_1 + \vartheta_2}}$, $i = 1, 2$; $\sigma_{12} = \sigma_{21} = -\frac{n m_1 - m_1^2}{n(1 + e^{\vartheta_2})} = -\frac{n m_2 - m_2^2}{n(1 + e^{\vartheta_1})}$.

Доказательство. Рассмотрим распределение вероятностей случайных величин ε_j со значениями во множестве двухмерных индексов I для $p_1 + p_2 \leqslant 1$ (табл. 1).

Первое утверждение теоремы очевидно:

Таблица 1

Распределение вероятностей случайных величин ε_j

№ исхода	ε_j	Вероятность
1	(1 0)	p_1
2	(0 1)	p_2
3	(1 1)	$1-p_1-p_2$

1) Y_1 представляет собой количество исходов № 1 и № 3 в n независимых испытаниях, так как $\{\varepsilon_j, j \geqslant 1\}$ — последовательность независимых одинаково распределённых случайных величин. Таким образом $Y_1 \sim Bi\,(n; 1-p_2)$.
2) Y_2 представляет собой количество исходов № 2 и № 3 в n независимых испытаниях, так как $\{\varepsilon_j, j \geqslant 1\}$ — последовательность независимых одинаково распределённых случайных величин. Таким образом $Y_2 \sim Bi\,(n; 1-p_1)$.

Из [1] следует, что трёхмерное полиномиальное распределение может быть описано с помощью производящей функции кумулянтов вида:

$$k_\mu\left(\vartheta_1;\vartheta_2\right) = n \cdot \ln\left(1+e^{\vartheta_1}+e^{\vartheta_2}\right),$$

где $\vartheta_1 = \ln\dfrac{p_1}{(1-p_1-p_2)}$, $\vartheta_2 = \ln\dfrac{p_2}{(1-p_1-p_2)}$. Отсюда запишем функцию распределения случайного вектора $Y = \sum\limits_{j=1}^{n}\varepsilon_j$ для любых натуральных y_1 и y_2, таких что $y_1 \leqslant n$, $y_2 \leqslant n$, $y_1+y_2 \geqslant n$:

$$\begin{gathered}
P_y\left(y_1,y_2\right) = P_X\left(n-y_2, n-y_1\right) = \\
= \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}\exp\left\{(n-y_2)\,\vartheta_1 + (n-y_1)\,\vartheta_2 - n\cdot\ln\left(1+e^{\vartheta_1}+e^{\vartheta_2}\right)\right\} = \\
= \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}\exp\left\{n\left(\vartheta_1+\vartheta_2\right) - \vartheta_1 y_2 - \vartheta_2 y_1 - n\cdot\ln\left(1+e^{\vartheta_1}+e^{\vartheta_2}\right)\right\} = \\
= \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}\exp\left\{y_1\left(-\vartheta_2\right) + y_2\left(-\vartheta_1\right) - n\cdot\ln\left(\frac{1+e^{\vartheta_1}+e^{\vartheta_2}}{e^{\vartheta_1+\vartheta_2}}\right)\right\} = \\
= \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}\exp\left\{y_1\left(-\vartheta_2\right) + y_2\left(-\vartheta_1\right) - \right. \\
\left. -n\cdot\ln\left(e^{(-\vartheta_1)}+e^{(-\vartheta_2)}+e^{(-\vartheta_1)+(-\vartheta_2)}\right)\right\} = \\
= \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}\exp\left\{\vartheta_1' y_1 + \vartheta_2' y_2 - n\cdot\ln\left(e^{\vartheta_1'}+e^{\vartheta_2'}+e^{\vartheta_1'+\vartheta_2'}\right)\right\},
\end{gathered}$$

где $\vartheta_1' = -\vartheta_2 = \ln\dfrac{1-p_1-p_2}{p_2}$, $\vartheta_2' = -\vartheta_1 = \ln\dfrac{1-p_1-p_2}{p_1}$.

Покажем, что полученное распределение принадлежит классу двухмерных натуральных экспонентных семейств на $\mathbb{R}^2$ с производящей функцией кумулянтов

$$k_\mu\left(\vartheta_1;\vartheta_2\right) = n\cdot\ln\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right).$$

Из следствий 3 теоремы 1 и свойства производящей функции кумулянтов

$$\frac{\partial^2 k_\mu\left(\vartheta_1,\vartheta_2\right)}{\partial\vartheta_1\vartheta_2} = \sigma_{12} = \frac{\partial^2 k_\mu\left(\vartheta_1,\vartheta_2\right)}{\partial\vartheta_2\vartheta_1} = \sigma_{21}$$

следует, что $\sigma_{11}(\vartheta_1+\beta(\vartheta_2))\beta'(\vartheta_2)=\sigma_{12}$, $\sigma_{22}(\vartheta_2+\beta(\vartheta_1))\beta'(\vartheta_1)=\sigma_{21}$.

Таким образом, если распределение вектора Y принадлежит классу НЭС $\mathcal{F}$ и $n\cdot\ln\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right)$ — его ПФК, то

$$\beta'(\vartheta_2)=-n\frac{e^{\vartheta_1+\vartheta_2}}{\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right)^2}\cdot\frac{\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right)^2}{n\left(e^{\vartheta_1+\vartheta_2}+e^{\vartheta_1+2\vartheta_2}\right)}=-\frac{1}{1+e^{\vartheta_2}},$$

$$\beta'(\vartheta_1)=-n\frac{e^{\vartheta_1+\vartheta_2}}{\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right)^2}\cdot\frac{\left(e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2}\right)^2}{n\left(e^{\vartheta_1+\vartheta_2}+e^{\vartheta_2+2\vartheta_1}\right)}=-\frac{1}{1+e^{\vartheta_1}},$$

т.е. $\beta(\vartheta)=\ln\left(1+e^{\vartheta}\right)-\ln e^{\vartheta}+\ln c=\ln\left(c(1+e^{-\vartheta})\right)$, где c — положительная константа.

Из доказательства следствия 1 теоремы 1 следует, что

$$m(\vartheta_1+\beta(\vartheta_2))=m(\vartheta_2+\beta(\vartheta_1))\beta'(\vartheta_1)+f'(\vartheta_1), \tag{8}$$

$$m(\vartheta_2+\beta(\vartheta_1))=m(\vartheta_1+\beta(\vartheta_2))\beta'(\vartheta_2)+f'(\vartheta_2), \tag{9}$$

где для биномиально распределённых маргиналов $m(\vartheta)=\dfrac{ne^{\vartheta}}{1+e^{\vartheta}}$.

Таким образом,

$$\frac{ne^{\vartheta_1+\beta(\vartheta_2)}}{1+e^{\vartheta_1+\beta(\vartheta_2)}}=\frac{ne^{\vartheta_2+\beta(\vartheta_1)}}{1+e^{\vartheta_2+\beta(\vartheta_1)}}\cdot\left(-\frac{1}{1+e^{\vartheta_1}}\right)+f'(\vartheta_1),$$

$$\begin{aligned}f'(\vartheta_1)&=\frac{ne^{\vartheta_1}(1+e^{-\vartheta_2})}{1+e^{\vartheta_1}(1+e^{-\vartheta_2})}-\frac{ne^{\vartheta_2}(1+e^{-\vartheta_1})}{1+e^{\vartheta_2}(1+e^{-\vartheta_1})}\cdot\left(-\frac{1}{1+e^{\vartheta_1}}\right)=\frac{n(e^{\vartheta_1}+e^{\vartheta_1-\vartheta_2})}{1+e^{\vartheta_1}+e^{\vartheta_1-\vartheta_2}}+\\&+\frac{n(e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1})}{1+2\cdot e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1}+e^{\vartheta_1}+e^{\vartheta_1+\vartheta_2}}=n-\frac{n(e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1})}{(1+e^{\vartheta_1}+e^{\vartheta_1-\vartheta_2})(e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1})}-\\&-\frac{n(e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1})}{1+2\cdot e^{\vartheta_2}+e^{\vartheta_2-\vartheta_1}+e^{\vartheta_1}+e^{\vartheta_1+\vartheta_2}}=n,\end{aligned}$$

$$f(\vartheta_1)=n\vartheta_1+c_1,\quad f(\vartheta_2)=n\vartheta_2+c_1,$$

где c_1 — неотрицательная константа. Отсюда

$$\begin{aligned}k_\mu(\vartheta_1,\vartheta_2)&=k(\vartheta_1+\beta(\vartheta_2))+f(\vartheta_2)=\\&=n\cdot\ln\left(1+\exp\left\{\vartheta_1+\ln\left(c\cdot\left(1+e^{-\vartheta_2}\right)\right)\right\}\right)+n\vartheta_2+c_1=\\&=n\cdot\ln\left(e^{\vartheta_2}\left(1+e^{\vartheta_1}c\cdot\left(1+e^{-\vartheta_2}\right)\right)\right)+c_1=n\cdot\ln\left(c\cdot e^{\vartheta_1}+e^{\vartheta_2}+c\cdot e^{\vartheta_1+\vartheta_2}\right)+c_1.\end{aligned}$$

Пусть имеется 2 производящие функции кумулянтов:

1) $k_{\mu_1(x_1,x_2)}(\vartheta_1,\vartheta_2)$ и $k_{\mu_2(x_1,x_2)}(\vartheta_1,\vartheta_2)=k_{\mu_1(x_1,x_2)}(\vartheta_1,\vartheta_2)+c_1$, тогда $\mu_2(x_1,x_2)=\dfrac{\mu_1(x_1,x_2)}{\exp(c_1)}$, что следует из определения производящей функции кумулянтов
2) $k_{\mu(x_1,x_2)}(\vartheta_1,\vartheta_2)=\ln\left(\displaystyle\int\exp\{\vartheta_1x_1+\vartheta_2x_2\}\,\mu(d(x_1,x_2))\right)$ и закона двухмерного распределения из класса $\mathcal{F}$,

$$P_{\vartheta_1,\vartheta_2}(x_1,x_2)=\exp\{x_1\vartheta_1+x_2\vartheta_2-k_\mu(\vartheta_1,\vartheta_2)\}\,\mu(x_1,x_2).$$

Зададим меру $\mu(x_1, x_2)$, для которой $c_1 = 0$. Тогда при $\beta(\vartheta) = \ln\left(c\left(1+e^{-\vartheta}\right)\right)$ мы получаем производящую функцию кумулянтов

$$k_\mu(\vartheta_1, \vartheta_2) = n \cdot \ln\left(c \cdot e^{\vartheta_1} + e^{\vartheta_2} + c \cdot e^{\vartheta_1+\vartheta_2}\right).$$

Определим $(\vartheta_1, \vartheta_2)$ натуральный параметр распределения из класса $\mathcal{F}$ соответствующего функции $\beta(\vartheta) = \ln\left(c\left(1+e^{-\vartheta}\right)\right)$ и порождающую меру $\mu(x_1, x_2)$.

Из доказательства теоремы 1 следует, что

$$\begin{cases} \vartheta_1 = \alpha_1 - \beta(\vartheta_2), \\ \vartheta_2 = \alpha_2 - \beta(\vartheta_1), \end{cases} \quad \begin{cases} \vartheta_1 = \alpha_1 - \ln\left(c\left(1+e^{-\vartheta_2}\right)\right), \\ \vartheta_2 = \alpha_2 - \ln\left(c\left(1+e^{-\vartheta_1}\right)\right), \end{cases}$$

$$\begin{cases} \vartheta_2 = \ln\left(\dfrac{1}{c^{-1} \cdot e^{\alpha_1-\vartheta_1} - 1}\right), \\ \vartheta_1 = \ln\left(\dfrac{c^{-2} \cdot e^{\alpha_1+\alpha_2} - 1}{c^{-1} \cdot e^{\alpha_2} + 1}\right), \end{cases} \quad \begin{cases} \vartheta_2 = \ln\left(\dfrac{c^{-2} \cdot e^{\alpha_1+\alpha_2} - 1}{c^{-1} \cdot e^{\alpha_1} + 1}\right), \\ \vartheta_1 = \ln\left(\dfrac{c^{-2} \cdot e^{\alpha_1+\alpha_2} - 1}{c^{-1} \cdot e^{\alpha_2} + 1}\right). \end{cases}$$

Из [1] следует, что $\alpha_1 = \ln\left(\dfrac{1-p_2}{p_2}\right)$, $\alpha_2 = \ln\left(\dfrac{1-p_1}{p_1}\right)$, тогда $\vartheta_1 = \ln\dfrac{1-p_1-p_2}{c \cdot p_2}$, $\vartheta_2 = \ln\dfrac{1-p_1-p_2}{c \cdot p_1}$.

Таким образом, полученное распределение вероятностей

$$P_Y(y_1, y_2) = \frac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!} \times \\ \times \exp\left\{\vartheta_1' y_1 + \vartheta_2' y_2 - n \cdot \ln\left(e^{\vartheta_1'} + e^{\vartheta_2'} + e^{\vartheta_1'+\vartheta_2'}\right)\right\}$$

единственно и принадлежит классу двухмерных натуральных экспонентных семейств $\mathcal{F}$.

Действительно для $c = 1$ получим:

$$\beta(\vartheta) = \ln\left(1+e^{-\vartheta}\right), \quad \vartheta_1 = \ln\frac{1-p_1-p_2}{p_2}, \quad \vartheta_2 = \ln\frac{1-p_1-p_2}{p_1},$$
$$k_\mu(\vartheta_1, \vartheta_2) = n \cdot \ln\left(e^{\vartheta_1} + e^{\vartheta_2} + e^{\vartheta_1+\vartheta_2}\right),$$

$\mu(y_1, y_2) = \dfrac{n!}{(n-y_2)!\,(n-y_1)!\,(y_1+y_2-n)!}$ — порождающая мера, так как $P_y(y_1, y_2)$ есть вероятность, следовательно, выполняется свойство нормировки вероятности $\sum\limits_{\substack{y_1 \leqslant n,\, y_2 \leqslant n \\ y_1+y_2 \geqslant n}} P_y(y_1, y_2) = 1$.

Запишем для полученного распределения дисперсионную функцию:

$$Var(m) = \begin{pmatrix} m_1 - \dfrac{m_1^2}{n} & \sigma_{12} \\ \sigma_{21} & m_2 - \dfrac{m_2^2}{n} \end{pmatrix},$$

где $m_i = n \cdot \dfrac{e^{\vartheta_i} + e^{\vartheta_1+\vartheta_2}}{e^{\vartheta_1} + e^{\vartheta_2} + e^{\vartheta_1+\vartheta_2}}$, $i = 1, 2$; $\sigma_{12} = \sigma_{21} = -\dfrac{nm_1 - m_1^2}{n\left(1+e^{\vartheta_2}\right)} = -\dfrac{nm_2 - m_2^2}{n\left(1+e^{\vartheta_1}\right)}$.

Определим параметрическое множество, опираясь на свойство положительной определённости ПФК [4]. Для этого найдём матрицу $D^2 k_\mu$ вторых частных производных $k_\mu(\vartheta_1,\vartheta_2)$:

$$D^2 k_\mu = \begin{pmatrix} \sigma_{11}(\vartheta_1+\beta(\vartheta_2)) & \sigma_{11}(\vartheta_1+\beta(\vartheta_2))\,\beta'(\vartheta_2) \\ \sigma_{22}(\vartheta_2+\beta(\vartheta_1))\,\beta'(\vartheta_1) & \sigma_{22}(\vartheta_2+\beta(\vartheta_1)) \end{pmatrix}$$

$$\left|D^2 k_\mu\right| = \sigma_{11}(\vartheta_1+\beta(\vartheta_2))\,\sigma_{22}(\vartheta_2+\beta(\vartheta_1))\,[1-\beta'(\vartheta_1)\,\beta'(\vartheta_2)]$$

Таким образом, для выполнения условия положительной определённости ПФК $k_\mu(\vartheta_1,\vartheta_2)$ необходимо и достаточно выполнение условия $\beta'(\vartheta_1)\,\beta'(\vartheta_2) < 1$. Найдём ϑ_1 и ϑ_2, удовлетворяющие этому условию:

$$-\frac{1}{1+e^{\vartheta_1}}\cdot\left(-\frac{1}{1+e^{\vartheta_2}}\right) < 1, \quad 1+e^{\vartheta_1}+e^{\vartheta_2}+e^{\vartheta_1+\vartheta_2} > 1.$$

Значит $(\vartheta_1,\vartheta_2)\in\mathbb{R}^2$.

Доказательство теоремы завершено. □

5. Заключение

Построенная в данной работе модель может быть использована в прикладных целях, таких как: оценка стоимости опционов, определение стоимости страхового контракта с одинаковыми страховыми суммами по пересекающимся рискам, разработка алгоритмов маршрутизации пакетов в телекоммуникационных системах.

Литература

1. *Morris C.* Natural Exponential Families with Quadratic Variance-Functions // Annals of Statistics. — 1982. — Vol. 10, No 1. — Pp. 65–80.
2. *Ivanova N. L.* The Reconsruction of Natural Exponential Families by Their Marginals // Journal of Mathematical Sciences. — 2001. — Vol. 106, No 1. — Pp. 2672–2681.
3. *Иванова Н. Л., Хохлов Ю. С.* Многомерная модель коллективного риска // Вестник МГУ. — 2005. — Т. 15, № 3. — С. 22–30. [*Ivanova N. L., Khokhlov Yu. S.* Mnogomernaya modelj kollektivnogo riska // Vestnik MGU. — 2005. — T. 15, No 3. — S. 22–30.]
4. *Иванова Н. Л., Хохлов Ю. С.* О восстановлении многомерного распределения по его компонентам // Вестник МГУ. — 2001. — Т. 15, № 1. — С. 32–37. [*Ivanova N. L., Khokhlov Yu. S.* O vosstanovlenii mnogomernogo raspredeleniya po ego komponentam // Vestnik MGU. — 2001. — T. 15, No 1. — S. 32–37.]

UDC 519.2

Two-Dimensional Binomial Distribution

Y. S. Khokhlov, I. P. Shestakov

Department of Probability Theory and Mathematical Statistics
Peoples' Friendship University of Russia
Miklukho-Maklaya str., 6, Moscow, 117198, Russia

We consider a model of two-dimensional random vector with dependent binomial marginals. The distribution for constructed model is determined.

Key words and phrases: natural exponential families, two-dimensional binomial distribution, cumulant generating function, variance function.