УДК 519.21

С. А. Матвейчук, Ю. И. Петунин

Обобщение схемы Бернулли, возникающее в вариационной статистике. I

Пусть $X = (x_1, \ldots, x_n)$ и $Y = (y_1, \ldots, y_m)$ — независимые выборки из генеральных совокупностей G_x и G_y, $x^{(1)} \leqslant \ldots \leqslant x^{(n)}$ — вариационный ряд, построенный по выборке X. Рассматривается схема испытаний, связанная с появлением зависимых событий $A_k = \{y_k \in (x^{(i)}, x^{(j)})\}$, $i < j$, $k = 1, 2, \ldots m$, где $x^{(i)}$, $x^{(j)}$ — порядковые статистики, которая представляет обобщение схемы испытаний Бернулли. Изучаются распределение частоты появления событий A_k и предельные теоремы, описывающие асимптотические свойства этих частот.

Нехай $X = (x_1, \ldots, x_n)$ та $Y = (y_1, \ldots, y_m)$ — незалежні вибірки з генеральних сукупностей G_x і G_y, $x^{(1)} \leqslant \ldots \leqslant x^{(n)}$ — варіаційний ряд, побудований по вибірці X. Розглядається схема випробувань, пов'язана з появою залежних випадкових подій $A_k = \{y_k \in (x^{(i)}, x^{(j)})\}$, $i < j$, $k = 1, 2, \ldots, m$, де $x^{(i)}$, $x^{(j)}$ — порядкові статистики, яка представляє узагальнення схеми випробувань Бернуллі. Вивчаються розподіл частоти появи подій A_k та граничні теореми, що описують асимптотичні властивості цих частот.

Введение. В настоящей статье изучаются обобщения схемы независимых испытаний Бернулли, которые возникают при построении критерия для идентификации генеральной совокупности в случае конечного класса альтернативных гипотез [1]. Кроме того, эта схема может служить для построения целого класса подобных критериев. Результаты, полученные в статье, необходимы для построения статистических критериев для проверки гипотез о равенстве функций распределения, которые используются при дифференциальной диагностике онкологических заболеваний.

Пусть $\bar{x}=(x_1, x_2, \dots, x_n)$ — выборка, полученная путем простого случайного выбора из генеральной совокупности G_x. Построим по ней вариационный ряд $x^{(0)} < x^{(1)} \leqslant \dots \leqslant x^{(n)} < x^{(n+1)}$ (условно приняты обозначения $x^{(0)} = -\infty$, $x^{(n+1)} = \infty$) и рассмотрим случайный интервал $J_{i,q} = (x^{(i)}, x^{(i+q)})$, i, q — фиксированные числа, $1 \leqslant q \leqslant n$, $0 \leqslant i \leqslant n - q + 1$.

Пусть $\bar{y} = (y_1, y_2, \dots, y_n)$ — выборка, полученная аналогичным образом из генеральной совокупности G_y. Построим на основании $\bar{y}$ и $J_{i,q}$ следующую схему испытаний: на k-м шаге, $k=1, 2, \dots, m$, проверяется, будет ли y_k принадлежать интервалу $J_{i,q}$. Имеем группу событий $A_k = \{y_k \in J_{i,q}\}$, каждое из которых может произойти с определенной вероятностью $\mathbf{P}(A_k) = p_k$. Будет показано, что эта вероятность не зависит от номера k, и тем самым все события A_k имеют одинаковую вероятность, но в отличие от схемы испытаний Бернулли события A_k зависимы между собой. Таким образом, исследуемая последовательность испытаний является в некотором роде обобщением схемы испытаний Бернулли.

Описанную выше схему испытаний в случае различных генеральных совокупностей G_x и G_y условимся называть модифицированной схемой испытаний Бернулли, а в случае совпадения генеральных совокупностей ($G_x = G_y$) — обобщенной схемой испытаний Бернулли. Обобщенная схема испытаний является частным случаем модифицированной схемы испытаний, поэтому все дальнейшие исследования посвящены именно модифицированной схеме испытаний.

Предположим, что генеральные совокупности G_x и G_y имеют непрерывные строго возрастающие функции распределения, которые обозначим соответственно через $F_x(u)$ и $F_y(u)$. Введем функцию

$$G(v) = F_y[F_x^{-1}(v)],\ v \in [0, 1], \tag{1}$$

и назовем ее идентификатором генеральных совокупностей G_x и G_y. Заметим, что в силу указанных предположений функции $G(v)$ и $G^{-1}(v)$ являются монотонными и непрерывными, поскольку обратные функции $F_x^{-1}(v)$ и $F_y^{-1}(v)$ будут непрерывными. Функция $G(v)$ играет фундаментальную роль в описанных схемах, в частности, в обобщенной схеме испытаний Бернулли

$$G(v) = v, \quad v \in [0, 1]. \tag{2}$$

На основании этого тождества все результаты, полученные для модифицированной схемы испытаний, непосредственно переносятся и на обобщенную схему испытаний Бернулли.

Введем следующий линейный функционал:

$$\mathfrak{F}_{n,i,q}(f) = \frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!} \int_0^1 v_1^{i-1} \times \left\{ \int_{v_1}^1 (v_2 - v_1)^{q-1} (1 - v_2)^{n-i-q} f(v_1, v_2)\, dv_2 \right\} dv_1, \tag{3}$$

где n, i, q — натуральные числа, $1 \leqslant q \leqslant n$, $0 \leqslant i \leqslant n - q + 1$, $f(v_1, v_2)$ — непрерывная функция, определенная на квадрате $K = \{(v_1, v_2): 0 \leqslant v_1, v_2 \leqslant 1\}$.

Приведем без доказательства некоторые свойства функционала (3), которые необходимы для изложения дальнейшего материала.

Предложение 1. *Справедливы равенства а)* $\mathfrak{F}_{n,i,q}(1) = 1$; *б)* $\mathfrak{F}_{n,i,q}(v_1) = i/(n+1)$; *в)* $\mathfrak{F}_{n,i,q}(v_2) = (i+q)/(n+1)$.

Предложение 2. *Пусть $f(v)$ и $\varphi(v)$ — дифференцируемые функции, определенные на отрезке $[0,1]$, причем $f(v)$ — возрастающая и принимает значения из $[0,1]$, если при $n\to\infty$ величины i и q неограниченно возрастают таким образом, что $i/(n+1)\to p^*$, $q/(n+1)\to p_0$, $0<p^*, p_0<$ <1. Тогда* 1)

$$\mathfrak{F}_{n,i,q}\{\varphi[f(v_2)-f(v_1)]\}=\varphi\left[f\left(\frac{i+q}{n+1}\right)-f\left(\frac{i}{n+1}\right)\right]+O(1/n), \quad (4)$$

2)

$$\lim_{n\to\infty}\mathfrak{F}_{n,i,q}\{\varphi[f(v_2)-f(v_1)]\}=\varphi[f(p^*+p_0)-f(p^*)]. \quad (5)$$

1. Модифицированная схема Бернулли. Перейдем к изучению модифицированной схемы испытаний Бернулли. Определим вероятность события A_k

$$\mathbf{P}(A_k)=\iint\limits_{u_1\leqslant u_2}[F_y(u_2)-F_y(u_1)]\,dF_{i,i+q}(u_1,u_2), \quad (6)$$

где $F_{i,i+q}(u_1,u_2)$ — совместная функция распределения порядковых статистик $x^{(i)}$ и $x^{(i+q)}$, Как известно [2],

$$dF_{i,i+q}(u_1,u_2)=\frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!}\,F_x^{i-1}(u_1)\,dF_x(u_1)\times$$

$$\times[F_x(u_2)-F_x(u_1)]^{q-1}[1-F_x(u_2)]^{n-i-q}dF_x(u_2). \quad (7)$$

Подставляя это выражение в (6) и переходя от двойного интеграла к повторному, получаем

$$\mathbf{P}(A_k)=\frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!}\int\limits_{-\infty}^{\infty}F_x^{i-1}(u_1)\left\{\int\limits_{u_1}^{\infty}[F_x(u_2)-F_x(u_1)]^{q-1}\times\right.$$

$$\left.\times[1-F_x(u_2)]^{n-i-q}[F_y(u_2)-F_y(u_1)]\,dF_x(u_2)\right\}dF_x(u_1).$$

Производя замену переменных $v_1=F_x(u_1)$, $v_2=F_x(u_2)$ и используя (3), переписываем это равенство следующим образом:

$$\mathbf{P}(A_k)=\mathfrak{F}_{n,i,q}[G(v_2)-G(v_1)], \quad (8)$$

где $G(v)$ определяется формулой (1).

Отсюда следует, что в модифицированной схеме испытаний вероятности всех событий совпадают между собой. Обозначим эту вероятность через $p_{n,1}$,

$$p_{n,1}=\mathbf{P}(A_k),\quad k=1,2,\dots,m. \quad (9)$$

Аналогично можно получить вероятность совместного наступления событий A_k и A_l при $k\neq l$

$$\mathbf{P}(A_kA_l)=\mathfrak{F}_{n,i,q}\{[G(v_2)-G(v_1)]^2\}, \quad (10)$$

где $G(v)$ определяется (1). Очевидно, эти вероятности равны между собой. Обозначим вероятность (10) через $p_{n,2}$

$$p_{n,2}=\mathbf{P}(A_kA_l),\quad k\neq l,\quad k,l=1,2,\dots,m. \quad (11)$$

Введем случайную величину $\varkappa_{i,q}$, равную числу событий A_k, появившихся в серии из m испытаний; $\varkappa_{i,q}$ принимает $m+1$ значение от нуля до m с вероятностями $p_l=\mathbf{P}(\varkappa_{i,q}=l)$. Справедлива следующая теорема.

Теорема 1. *В модифицированной схеме испытаний Бернулли случайная величина $\varkappa_{i,q}$ имеет следующее распределение вероятностей:*

$$\mathbf{P}(\varkappa_{i,q}=l)=\mathfrak{F}_{n,i,q}\{C_m^l[G(v_2)-G(v_1)]^l[1-G(v_2)+G(v_1)]^{m-l}\},$$

$$l=0,1,\dots,m, \quad (12)$$

где m, n — объемы выборок $\bar{y}$, $\bar{x}$ соответственно, i, q — фиксированные числа, $1 \leqslant q \leqslant n$, $0 \leqslant i \leqslant n-q+1$, C_m^l — число сочетаний из m по l.

Доказательство. Покажем, что числа (12) действительно образуют распределение вероятностей, т. е. $\sum_{l=0}^{m} \mathbf{P}(\varkappa_{i,q}=l)=1$.

В самом деле,

$$\sum_{l=0}^{m} \mathbf{P}(\varkappa_{i,q}=l)=\mathfrak{F}_{n,i,q}\Big\{\sum_{l=0}^{m} C_m^l\,[G(v_2)-G(v_1)]^l\,[1-G(v_2)+G(v_1)]^{m-l}\Big\}=$$

$$=\mathfrak{F}_{n,i,q}\{[G(v_2)-G(v_1)+1-G(v_2)+G(v_1)]^m\}=1.$$

Покажем теперь, что $\mathbf{P}(\varkappa_{i,q}=l)$ вычисляется по формуле (12). Нетрудно видеть, что справедливо соотношение

$$\mathbf{P}(\varkappa_{i,q}=l)=\sum_{(i_1,\ldots,i_m)} \mathbf{P}(A_{i_1}\ldots A_{i_l}\bar{A}_{i_{l+1}}\ldots\bar{A}_{i_m}),$$

где суммирование производится по всем наборам индексов $(i_1,\ldots,i_m)$, $i_k=1,2,\ldots,m$, $i_k\neq i_s$, если $k\neq s$; в этой сумме C_m^l слагаемых, а $\bar{A}_{i_k}$—противоположенное событие к A_{i_k}, т. е. $\bar{A}_{i_k}=\{y_{i_k}\bar{\in} J_{i,q}\}$. Обозначим вероятность $\mathbf{P}(A_{i_1}\ldots A_{i_l}\bar{A}_{i_{l+1}}\ldots\bar{A}_{i_m})$ через $p_{i_1\ldots i_m}$ и определим ее значение

$$p_{i_1\ldots i_m}=\int\underset{\Omega}{\cdots}\int dF(y_{i_1},\ldots,y_{i_m},x_i,x_{i+q}), \tag{13}$$

где

$$\Omega=\{(y_{i_1},\ldots,y_{i_m},x_i,x_{i+q}):-\infty<x_i\leqslant x_{i+q}<\infty,$$

$$x_i<y_{i_k}\leqslant x_{i+q},\ k=1,2,\ldots,l,\ y_{i_j}\bar{\in}(x_i,x_{i+q}],\ j=l+1,\ l+2,\ldots,m\},$$

а $F(y_{i_1},\ldots,y_{i_m},x_i,x_{i+q})$ — совместная функция распределения случайных величин $y_1,\ldots,y_m,x^{(i)},x^{(i+q)}$. В силу независимости выборок $\bar{x}$ и $\bar{y}$, а также независимости выборочных значений y_i между собой имеем

$$dF(y_{i_1},\ldots,y_{i_m},x_i,x_{i+q})=dF(y_{i_1})\ldots dF(y_{i_m})\,dF_{i,i+q}(x_i,x_{i+q}),$$

причем $F_{i,i+q}(x_i,x_{i+q})$ определяется по формуле (7).

Перейдем к вычислению интеграла (13). Имеем

$$p_{i_1\ldots i_m}=\frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!}\int_{-\infty}^{\infty}F_x^{i-1}(x_i)\,dF_x(x_i)\int_{x_i}^{\infty}[F_x(x_{i+q})-$$

$$-F_x(x_i)]^{q-1}\,[1-F_x(x_{i+q})]^{n-i-q}dF_x(x_{i+q})\int_{x_i}^{x_{i+q}}dF_y(y_{i_1})\ldots\int_{x_i}^{x_{i+q}}dF_y(y_{i_l})\times$$

$$\times\Big[1-\int_{x_i}^{x_{i+q}}dF_y(y_{i_{l+1}})\Big]\ldots\Big[1-\int_{x_i}^{x_{i+q}}dF_y(y_{i_m})\Big].$$

Проинтегрировав и сделав замену переменных $v_1=F_x(x_i)$, $v_2=F_x(x_{i+q})$, получим $p_{i_1\ldots i_m}=\mathfrak{F}_{n,i,q}\{[G(v_2)-G(v_1)]^l\,[1-G(v_2)+G(v_1)]^{m-l}\}$, где $G(v)$ соответствует (1), а $\mathfrak{F}_{n,i,q}[\cdot]$ — (3).

Отсюда следует, что $p_{i_1\ldots i_m}$ не зависит от $i_1,\ldots,i_m$, а значит, $\mathbf{P}(\varkappa_{i,q}=l)=C_m^l p_{i_1\ldots i_m}$. Теорема доказана.

В качестве следствия из теоремы 1 нетрудно получить выражение для математического ожидания и дисперсии случайной величины $\varkappa_{i,q}$. Эти величины соответственно равны

$$\mathbf{E}(\varkappa_{i,q}) = m\mathfrak{F}_{n,i,q}[G(v_2) - G(v_1)], \tag{14}$$

$$\mathbf{D}(\varkappa_{i,q}) = m^2(\mathfrak{F}_{n,i,q}\{[G(v_2) - G(v_1)]^2\} - [\mathfrak{F}_{n,i,q}\{[G(v_2) - G(v_1)]\}\,_2 - \\ - m[\mathfrak{F}_{n,i,q}\{[G(v_2) - G(v_1)]^2\} - \mathfrak{F}_{n,i,q}\{[G(v_2) - G(v_1)]\}]. \tag{15}$$

Если учесть обозначения (8) — (11), то выражения (14) и (15) принимают более компактный вид

$$\mathbf{E}(\varkappa_{i,q}) = mp_{n,1}, \tag{16}$$

$$\mathbf{D}(\varkappa_{i,q}) = m^2(p_{n,2} - p_{n,1}^2) - m(p_{n,2} - p_{n,1}). \tag{17}$$

Аналогично можно получить выражения для характеристической функции случайной величины $\varkappa_{i,q}$, которую обозначим через $\varphi_{m,n}(t)$

$$\varphi_{m,n}(t) = \mathfrak{F}_{n,i,q}\{[1 + (e^{it} - 1)(G(v_2) - G(v_1))]^m\}. \tag{18}$$

Перейдем теперь к изучению асимптотических свойств статистики $\varkappa_{i,q}$. Необходимо исследовать поведение 1) при фиксированном n и $m \to \infty$; 2) при фиксированном m и $n \to \infty$, если $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$; 3) при условии $m \to \infty$, $n \to \infty$, если $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$; 4) если $m \to \infty$, $n \to \infty$, $i/(n+1) \to p^*$, $0 < p^* < 1$, q — фиксированное.

Вместо статистики $\varkappa_{i,q}$ будем рассматривать статистику $\xi_{m,n}^{(i,q)} = (\varkappa_{i,q} - \mathbf{E}(\varkappa_{i,q}))/\sqrt{\mathbf{D}(\varkappa_{i,q})}$. Ее характеристическая функция равна

$$\bar{\varphi}_{m,n}(t) = \exp\{-it\mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})}\}\,\varphi_{m,n}(t/\sqrt{\mathbf{D}(\varkappa_{i,q})}), \tag{19}$$

где $\varphi_{m,n}(t)$ — характеристическая функция статистики $\varkappa_{i,q}$, определяемая формулой (18). Через $\xi_n^{(i,q)}, \xi_m$ обозначим предельные величины, к которым стремится статистика $\xi_{m,n}^{(i,q)}$ при $m \to \infty$ и $n \to \infty$ соответственно, а через $\bar{\varphi}_n(t)$, $F_n(x)$ и $\bar{\varphi}_m(t)$, $F_m(x)$ — характеристическую функцию и функцию распределения случайных величин $\xi_n^{(i,q)}$ и ξ_m соответственно. Всюду в дальнейшем сходимость последовательности случайных величин ξ_n к ξ понимается в смысле слабой сходимости функций распределения (ССФР): $F_{\xi_n}(x) \to F_\xi(x)$, $n \to \infty$, на множестве $C(F_\xi)$ точек непрерывности функции $F_\xi(x)$. Наконец, через ξ^* будем обозначать предельную величину, к которой стремится статистика $\xi_{m,n}^{(i,q)}$ при $m \to \infty$ и $n \to \infty$ одновременно; пусть $\bar{\varphi}(t)$ и $F(x)$ — характеристическая функция и функция распределения ξ^*. Имеет место следующая теорема.

Теорема 2. *В модифицированной схеме Бернулли при $m \to \infty$ и при каждом n имеет место равномерная по $t \in R^1$ сходимость*

$$\bar{\varphi}_{n,m}(t) \to \bar{\varphi}_n(t) = \mathfrak{F}_{n,i,q}\{e^{it[G(v_2) - G(v_1) - p_{n,1}]/\sqrt{p_{n,2} - p_{n,1}^2}}\}, \tag{20}$$

где $p_{n,1}$ и $p_{n,2}$ определяются соотношениями (8) — (11).

Доказательство. Обозначим через $\psi_m(t)$ следующее выражение:

$$\psi_m(t) = e^{-it\mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})}}\{1 + [e^{it/\sqrt{\mathbf{D}(\varkappa_{i,q})}} - 1][G(v_2) - G(v_1)]\}^m, \tag{21}$$

тогда $\bar{\varphi}_{m,n}(t) = \mathfrak{F}_{n,i,q}[\psi_m(t)]$.

Рассмотрим функцию $\ln\psi_m(t)$, разложив предварительно $\exp\{it/\sqrt{\mathbf{D}(\varkappa_{i,q})}$ в ряд Тейлора в окрестности единицы

$$\ln\psi_m(t) = -it\mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})} + m\ln\{1 + [it/\sqrt{\mathbf{D}(\varkappa_{i,q})} + \\ + O(1/\sqrt{\mathbf{D}(\varkappa_{i,q})})][G(v_2) - G(v_1)]\}.$$

Учитывая, что $\mathbf{D}(\varkappa_{i,q}) = O(m^2)$ и раскладывая логарифмы в ряд Тейлора в окрестности единицы, в силу формулы (15) имеем

$$\ln \psi_m(t) = -\frac{it\mathbf{E}(\varkappa_{i,q})}{\sqrt{\mathbf{D}(\varkappa_{i,q})}} + \frac{itm[G(v_2) - G(v_1)]}{\sqrt{\mathbf{D}(\varkappa_{i,q})}} + O(1/m).$$

Учитывая выражения (16) и (17), получаем

$$\psi_m(t) = \exp\left\{\frac{it[G(v_2) - G(v_1) - p_{n,1}]}{\sqrt{p_{n,2} - p_{n,1}^2 - (p_{n,2} - p_{n,1})/m}}\right\} + O(1/m),$$

тогда

$$\overline{\varphi}_{m,n}(t) = \mathfrak{F}_{n,i,q}\left[\exp\left\{\frac{it[G(v_2) - G(v_1) - p_{n,1}]}{\sqrt{p_{n,2} - p_{n,1}^2 - (p_{n,2} - p_{n,1})/m}}\right\}\right] + O(1/m).$$

Переходя к пределу при $m \to \infty$, завершаем доказательство теоремы.

Пусть теперь $F_{m,n}(x)$ — функция распределения статистики $\xi_{m,n}^{(i,q)}$, где $x = [l - \mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})}]$, $l = \overline{0, m}$. Предположим, что $G(v)$ является дифференцируемой функцией, тогда существует ее производная $G'(v)$. Имеет место следующая теорема.

Теорема 3. *В модифицированной схеме испытаний Бернулли при $m \to \infty$, каждом фиксированном n и $x \in [-p_{n,1}/\sqrt{p_{n,2} - p_{n,1}^2}, (1 - p_{n,1})/ /\sqrt{p_{n,2} - p_{n,1}^2}]$ $F_{m,n}(x) \to F_n(x)$, где $F_n(x)$ — непрерывная функция распределения, имеющая плотность*

$$f_n(x) = \frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!} \int_0^{G^{-1}(1-y)} w^{i-1}\{G^{-1}[G(w+y)] - w\}^{q-1} \times$$

$$\times \{1 - G^{-1}[G(w+y)]\}^{n-i-q}] \frac{\sqrt{p_{n,2} - p_{n,1}^2}}{G'\{G^{-1}[G(w) + y]\}}\, dw \tag{22}$$

при $y = p_{n,1} + x\sqrt{p_{n,2} - p_{n,1}^2}$.

Доказательство. Рассмотрим предельную функцию

$$\overline{\varphi}_n(t) = \mathfrak{F}_{n,i,q}[\exp\{it[G(v_2) - G(v_1) - p_{n,1}]/\sqrt{p_{n,2} - p_{n,1}^2}\}].$$

Докажем, что $\overline{\varphi}_n(t)$ непрерывна в точке $t = 0$. Выберем некоторое $\delta > 0$ такое, что $|t| \leqslant \delta$. Учитывая, что $\mathfrak{F}_{n,i,q}[1] = 1$, получаем

$$|\overline{\varphi}_n(t) - \overline{\varphi}_n(0)| \leqslant \mathfrak{F}_{n,i,q}\left[\left|\exp\left\{\frac{it[G(v_2) - G(v_1) - p_{n,1}]}{\sqrt{p_{n,2} - p_{n,1}^2}}\right\} - 1\right|\right].$$

Воспользовавшись для экспоненты теоремой Лагранжа о среднем значении [3], перепишем это неравенство следующим образом:

$$|\overline{\varphi}_n(t) - \overline{\varphi}_n(0)| \leqslant \mathfrak{F}_{n,i,q}\left[\left|\frac{it[G(v_2) - G(v_1) - p_{n,1}]\theta}{\sqrt{p_{n,2} - p_{n,1}^2}}\right|\right],$$

где $0 \leqslant \theta \leqslant 1$.

Далее, $|\theta[p_{n,1} - G(v_2) + G(v_1)]| \leqslant \max(p_{n,1}, 1 - p_{n,1}) \leqslant 1$, значит $|\overline{\varphi}_n(t) - \overline{\varphi}_n(0)| \leqslant |t|/\sqrt{p_{n,2} - p_{n,1}^2} \leqslant \delta/\sqrt{p_{n,2} - p_{n,1}^2}$. Используя теперь предельную теорему Леви — Крамера для характеристических функций [4] и учитывая теорему 2, получаем $F_{m,n}(x) \to F_n(x)$, $m \to \infty$, равномерно по $x \in [-p_{n,1}/\sqrt{p_{n,2}^2 - p_{n,1}^2}, (1 - p_{n,1})/\sqrt{p_{n,2} - p_{n,1}^2}]$ и имеет $\overline{\varphi}_n(t)$ своей характеристической функцией. Но тогда по определению характеристической функции

$$\overline{\varphi}_n(t) = \int_a^b e^{itx} dF_n(x); \qquad b = \frac{1 - p_{n,1}}{\sqrt{p_{n,2} - p_{n,1}^2}}, \qquad a = -\frac{p_{n,1}}{\sqrt{p_{n,2} - p_{n,1}^2}}. \tag{23}$$

Но, с другой стороны, проделав в (20) замену переменных $w = v_1$, $x = [G(v_2) - G(v_1) - p_{n,1}]/\sqrt{p_{n,2} - p_{n,1}^2}$, после ряда необходимых преобразований будем иметь

$$\bar{\varphi}_n(t) = \int_a^b e^{itx} f_n(x)\,dx, \tag{24}$$

где $f_n(x)$ определяется выражением (22). Из соотношений (23) и (24) следует, что $f_n(x)$ является плотностью функции распределения $F_n(x)$. Теорема доказана.

В качестве следствия теорем 2 и 3 справедлива теорема.

Теорема 4. *В модифицированной схеме испытаний Бернулли при $m \to \infty$ случайная величина $\xi_{m,n}^{(i,q)}$ сходится в смысле ССФР к непрерывной случайной величине $\xi_n^{(i,q)}$, плотность распределения которой определяется выражением* (22).

Исследуем теперь поведение $\xi_{m,n}^{(i,q)}$ при фиксированном m и $n \to \infty$. Имеет место следующая теорема.

Теорема 5. *В модифицированной схеме испытаний Бернулли при $n \to \infty$ и $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, при каждом фиксированном m имеет место равномерная по $t \in R^1$ сходимость*

$$\bar{\varphi}_{m,n}(t) \to \bar{\varphi}_m(t) = \exp\{-itE_m/\sqrt{D_m}\}(1 + [\exp\{it/\sqrt{D_m}\} - 1] \times$$
$$\times [G(p^* + p_0) - G(p^*)])^m, \tag{25}$$

где

$$E_m = m[G(p^* + p_0) - G(p^*)], \quad D_m = m[G(p^* + p_0) - G(p^*)] \times$$
$$\times [1 - G(p^* + p_0) + G(p^*)].$$

Доказательство. Функция

$$\psi_m(t) = \exp\{-it\mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})}\}\{1 + [\exp\{it/\sqrt{\mathbf{D}(\varkappa_{i,q})}\} - 1] \times$$
$$\times [G(v_2) - G(v_1)]\}^m$$

является дифференцируемой функцией по переменной $u = G(v_2) - G(v_1)$ при фиксированном t, функция $G(v)$ — дифференцируемая и возрастающая в силу допущенных предположений, поэтому можно применить предложение 2. Следовательно,

$$\bar{\varphi}_{m,n}(t) = \mathfrak{F}_{n,i,q}[\psi_m(t)] = \exp\{-it\mathbf{E}(\varkappa_{i,q})/\sqrt{\mathbf{D}(\varkappa_{i,q})}\} \times$$
$$\times [1 + [\exp\{it/\sqrt{\mathbf{D}(\varkappa_{i,q})}\} - 1]\left[G\left(\frac{i+q}{n+1}\right) - G\left(\frac{i}{n+1}\right)\right]\Bigg]^m + O(1/n). \tag{26}$$

Применяя к выражениям (14) и (15) предложение 2, получаем

$$\lim_{n\to\infty} \mathbf{E}(\varkappa_{i,q}) = m[G(p^* + p_0) - G(p^*)],$$

$$\lim_{n\to\infty} \mathbf{D}(\varkappa_{i,q}) = m[G(p^* + p_0) - G(p^*)][1 - G(p^* + p_0) + G(p^*)].$$

Переходя теперь в (26) к пределу при $n \to \infty$ и учитывая, что $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, завершаем доказательство теоремы.

Исходя из предельной теоремы Леви — Крамера для характеристических функций [4], учитывая, что $\bar{\varphi}_m(t)$ является характеристической функцией биномиального распределения $B_m(mp_1 + x\sqrt{mp_1(1-p_1)}, p_1)$, где $p_1 = G(p^* + p_0) - G(p^*)$ и ссылаясь на теорему 5, можем доказать такую теорему.

Теорема 6. *В модифицированной схеме испытаний Бернулли при $n \to \infty$, $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, равномерно по m и*

$x \in [-mp_1/\sqrt{mp_1(1-p_1)},\ m(1-p_1)/\sqrt{mp_1(1-p_1)}]$,

$$F_{m,n}(x) \to F_m(x) = B_m(mp_1 + x\sqrt{mp_1(1-p_1)},\ p_1), \tag{27}$$

где $p_1 = G(p^* + p_0) - G(p^*)$.

Из теорем 5 и 6 следует теорема.

Теорема 7. *В модифицированной схеме испытаний Бернулли при* $n \to \infty$, $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, *случайная величина* $\xi_{m,n}^{(i,q)}$ *сходится в смысле ССФР к случайной величине* ξ_m, *имеющей биномиальный закон распределения.*

Теперь рассмотрим поведение $\xi_{m,n}^{(i,q)}$ при $m \to \infty$ и $n \to \infty$ одновременно.

Теорема 8. *В модифицированной схеме испытаний Бернулли при* $m \to \infty$, $n \to \infty$ *и* $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, *при каждом* $t \in R^1$ $m/n \to r$ $(r \in R^1)$,

$$\bar{\varphi}_{m,n}(t) \to \varphi(t) = \exp(-t^2/2). \tag{28}$$

Доказательство. Преобразуем функцию $\psi_m(t)$ (см. (21)) так же, как и при доказательстве теоремы 2, используя интегральное представление функционала $\mathfrak{F}_{n,i,q}[\cdot]$ (см. (3)), применяя для факториалов формулу Стирлинга (см., например, [5]) и делая замену переменных $w_1 = v_1\sqrt{n-2} - \dfrac{i-1}{\sqrt{n-2}}$, $w_2 = v_2\sqrt{n-2} - \dfrac{i+q-1}{\sqrt{n-2}}$, получаем следующее выражение для $\bar{\varphi}_{m,n}(t)$:

$$\bar{\varphi}_{m,n}(t) = \frac{O(1)}{2\pi}\sqrt{\frac{(i-1)(q-1)(i+q-1)}{(n-2)^3}}\int\limits_{-a_n}^{b_n} dw_1 \int\limits_{c_n}^{d_n} dw_2 \exp\left\{-\frac{w_1^2(n-2)}{2(i-1)} - \right.$$

$$-\frac{(w_2-w_1)^2(n-2)}{2(q-1)} - \frac{w_2^2(n-2)}{2(n-i-q)} - \frac{mit}{\sqrt{D(\varkappa_{i,q})}}[p_{1,n} - p_n(w_1, w_2)] -$$

$$\left. -\frac{mt^2}{2} p_n(w_1, w_2)[1 - p_n(w_1, w_2)] + O((n-2)^{-1/2} + mD^{-3/2}(\varkappa_{i,q}))\right\},$$

где

$$p_n(w_1, w_2) = G\left(\frac{i+q-1}{n-2} + \frac{w_2}{\sqrt{n-2}}\right) - G\left(\frac{i-1}{n-2} + \frac{w_1}{\sqrt{n-2}}\right),$$

$$a_n = (i-1)\sqrt{n-2},\ b_n = (n-i-q)\sqrt{n-2},\quad c_n = w_1\sqrt{n-2} -$$

$$-(i-1)/\sqrt{n-2},\quad d_n = (n-i-q)/\sqrt{n-2}.$$

Перейдем в этом выражении к пределу при $m, n \to \infty$. Учитывая условия теоремы и $\lim\limits_{m,n\to\infty} p_{1,n} = p_1$,

$$\lim_{m,n\to\infty} D(\varkappa_{i,q})/m = r\alpha(p^*, p_0) + p_1(1-p_1),$$

где

$$p_1 = G(p^* + p_0) - G(p^*),\quad \alpha(p^*, p_0) = [G'(p^*)]^2 p^*(1-p^*) +$$

$$+[G'(p^*+p_0)]^2(p^*+p_0)(1-p^*-p_0) - 2G(p^*+p_0)G(p^*)p^*(1-p^*-p_0),$$

после ряда преобразований получаем искомый результат. Теорема доказана.

Аналогично изложенному выше могут быть доказаны предельные теоремы для функций распределения $\{F_{m,n}(x)\}$ и статистик $\{\xi_{m,n}^{(i,q)}\}$. Сформулируем их.

Теорема 9. *В модифицированной схеме испытаний Бернулли при* $n \to \infty$, $m \to \infty$, $m/n \to r$ и $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, *равномерно по* $x \in [a, b] \subset R^1$ *справедливо предельное соотношение* $F_{m,n}(x) \to F(x)$, *где* $F(x)$ — *функция распределения стандартного нормального закона распределения.*

Теорема 10. *В модифицированной схеме испытаний Бернулли при $m \to \infty$ и $n \to \infty$ если $i/(n+1) \to p^*$, $q/(n+1) \to p_0$, $0 < p^*$, $p_0 < 1$, то последовательность случайных величин $\{\xi_{m,n}^{(i,q)}\}$ сходится в смысле ССФР к случайной величине ξ^*, имеющей стандартное нормальное распределение.*

Теоремы 8 — 10 описывают асимптотическое поведение статистики $\xi_{m,n}^{(i,q)}$ при условии, что $\lim\limits_{n\to\infty} q/(n+1) = p_0$. Это означает, что число порядковых статистик q, составляющих случайный интервал $J_{i,q} = (x^{(i)}, x^{(i+q)})$, неограниченно возрастает при увеличении n. Поэтому необходимо исследовать поведение $\varkappa_{i,q}$ в том случае, когда величина q не зависит от n и с ростом n остается постоянной.

Обозначим через ξ_q^* предельную случайную величину, к которой сходится по вероятности статистика $\varkappa_{i,q}$ при фиксированном q, m, $n \to \infty$ и $i/(n+1) \to p^*$, а через $F_q(x)$ и $\varphi_q(t)$ — функцию распределения и характеристическую функцию статистики ξ_q^* соответственно. Справедлива следующая теорема.

Теорема 11. *В модифицированной схеме испытаний Бернулли при $m \to \infty$ и $n \to \infty$ если $m/(n-2) \to r$, $r \in R^+$, $i/(n+1) \to p^*$, $0 < p^* < 1$, q — фиксированное и $G(v)$ является дифференцируемой функцией, то при каждом t*

$$\varphi_{m,n}(t) \to \varphi_q(t) = \{1 + rG'(p^*)(1 - e^{it})\}^{-q}, \tag{29}$$

где $G'(v)$ — производная функции $G(v)$, $\varphi_{m,n}(t)$ — характеристическая функция $\varkappa_{i,q}$.

Доказательство. Воспользовавшись формулами (3) и (18), запишем

$$\varphi_{m,n}(t) = \frac{n!}{(i-1)!\,(q-1)!\,(n-i-q)!} \int\limits_0^1 v_1^{i-1} dv_1 \int\limits_{v_1}^1 (v_2 - v_1)^{q-1} (1 - v_2)^{n-i-q} \times$$
$$\times [1 + (e^{it} - 1)\{G(v_2) - G(v_1)\}]^m dv_2. \tag{30}$$

Сделаем следующую замену переменных: $w_1 = v_1\sqrt{n-2} - (i-1)/\sqrt{n-2}$, $w_2 = (n-2)(v_2 - v_1)$. Якобиан этого преобразования равен $J(w_1, w_2) = 1/[(n-2)\sqrt{n-2}]$. Учитывая, что с ростом n величины i и $n-i-q$ неограниченно возрастают, используем формулу Стирлинга [5] для $(n-2)!$, $(i-1)!$, $(n-i-q)!$, переписываем выражение (30) в виде

$$\varphi_{m,n}(t) = \frac{e^{\theta_n - q + 1}}{(q-1)!\,\sqrt{2\pi}} \sqrt{\frac{(n-1)^2}{(i-1)(n-i-q)}}\, \frac{n}{n-2} \times$$
$$\times \int\limits_a^b dw_1 \int\limits_0^c w_2^{q-1} B_n(w_1, w_2)\, f_m(w_1, w_2)\, dw_2, \tag{31}$$

где

$$a = -(i-1)/\sqrt{n-2}, \quad b = (n-i-1)/\sqrt{n-2}, \quad c = (n-2) \times$$
$$\times \left[\frac{n-i-1}{n-2} - w_1 \frac{1}{\sqrt{n-2}}\right], \quad \theta_n = \bar\theta/12n, \quad |\bar\theta| \leqslant 1,$$

$$B_n(w_1, w_2) = \left(1 + \frac{w_1\sqrt{n-2}}{i-1}\right)^{i-1} \left(1 + \frac{q - 1 - w_1\sqrt{n-2} - w_2}{n-i-q}\right)^{n-i-q}, \tag{32}$$

$$f_m(w_1, w_2) = \left\{1 + (e^{it} - 1)\left[G\left(\frac{w_2}{n-2} + \frac{w_1}{\sqrt{n-2}} + \frac{i-1}{n-2}\right) - \right.\right.$$
$$\left.\left. - G\left(\frac{w_1}{\sqrt{n-2}} + \frac{i-1}{n-2}\right)\right]\right\}^m.$$

Воспользуемся для функции $G(v)$ теоремой Лагранжа о среднем значении [3]

$$G\left(\frac{w_2}{n-2}+\frac{w_1}{\sqrt{n-2}}+\frac{i-1}{n-2}\right)-G\left(\frac{w_1}{\sqrt{n-2}}+\frac{i-1}{n-2}\right)=G'(\alpha_n)\frac{w_2}{n-2}, \tag{33}$$

где

$$\frac{w_1}{\sqrt{n-2}}+\frac{i-1}{n-2}\leqslant\alpha_n\leqslant\frac{w_2}{n-2}+\frac{w_1}{\sqrt{n-2}}+\frac{i-1}{n-2}.$$

С учетом этих результатов рассмотрим теперь выражение

$$\ln[B_n(w_1,w_2)f_m(w_1,w_2)]=(i-1)\ln[1+w_1\sqrt{n-2}/(i-1)]+(n-i-q)\times$$
$$\times\ln[1+(q-1-w_1\sqrt{n-2}-w_2)/(n-i-q)]+m\ln[1+(e^{it}-1)\times$$
$$\times G'(\alpha_n)w_2/(n-2)].$$

Разлагая каждый из логарифмов в ряд Тейлора и проделывая необходимые преобразования, записываем

$$\ln[B_n(w_1,w_2)f_m(w_1,w_2)]=q-1-w_2-w_1^2(n-2)^2/[2(i-1)(n-i-q)]+$$
$$+[m/(n-2)](e^{it}-1)G'(\alpha_n)w_2+O(1/\sqrt{n})+O(m/(n-2)^2). \tag{34}$$

Переходя в (31) к пределу при $m,n\to\infty$ и учитывая условия теоремы, а также соотношения (32) — (34), находим

$$\lim_{\substack{n\to\infty\\ m\to\infty}}\varphi_{m,n}(t)=\frac{1}{\sqrt{2\pi p^*(1-p^*)}}\int_{-\infty}^{\infty}\exp\left\{-\frac{w_1^2}{2p^*(1-p^*)}\right\}dw_1\times$$
$$\times\frac{1}{(q-1)!}\int_0^{\infty}w_2^{q-1}\exp\{-w_2[1+r(e^{it}-1)G'(p^*)]\}dw_2.$$

Вычисляя полученины нтегралы, завершаем доказательство теоремы.

Из формулы (29) следует [6], что $\varphi_q(t)$ является характеристической функцией отрицательного биномиального распределения. Можно показать, что имеют место теоремы, аналогичные теоремам 9 и 10. Пусть $\overline{F}_{m,n}(k)$ — функция распределения статистики $\varkappa_{i,q}$.

Теорема 12. *В модифицированной схеме испытаний Бернулли при $m,n\to\infty$ если $\frac{m}{n-2}\to r$, $r\in R^+$, $\frac{i}{n+1}\to p^*$, q — фиксированное и $G(v)$ является дифференцируемой функцией, то равномерно по k*

$$\overline{F}_{m,n}(k)\to F_q(k)=\sum_{l=0}^{k}C_{q+l-1}^{l}[rG'(p^*)]^l/[1+rG'(p^*)]^{q+l}, \tag{35}$$

где $k\leqslant N$, k,N — целые положительные числа.

Теорема 13. *В модифицированной схеме испытаний Бернулли при $m,n\to\infty$ и если $\frac{m}{n-2}\to r$, $r\in R^+$, $\frac{i}{n+1}\to p^*$, $0<p^*<1$, q — фиксированное и $G(v)$ является дифференцируемой функцией, то последовательность величин $\{\varkappa_{i,q}\}$ сходится в смысле ССФР к случайной величине ξ_q^*, имеющей отрицательное биномиальное распределение.*

1. *Петунин Ю. И., Тимошенко Я. Г., Петунина М. Ю.* Критерий для идентификации генеральных совокупностей в случае конечного класса альтернативных гипотез // Докл. АН УССР. Сер. А.— 1984.— № 6.— С. 28—31.
2. *Дэйвид Г.* Порядковые статистики.— М. : Наука, 1979.— 355 с.

3. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления: В 3-:
М. : Наука, 1966.— Т. 1.— 800 с.
4. *Крамер Г.* Математические методы статистики.— М. : Мир, 1975.— 648 с.
5. *Фихтенгольц Г. М.* Курс дифференциального и интегрального исчисления: В 3-:
М. : Наука, 1966.— Т. 11.— 800 с.
6. *Кендалл М. Дж., Стьюарт А.* Теория распределений.— М. : Наука, 1966.— {

Киев. ун-т Получено 10