О максимальном расхождении между двумя эмпирическими распределениями

Е. Л. Рвачева

Введение

Приложения математической статистики в промышленной и научной практике часто приводят к необходимости оценки правильности гипотезы, что две серии наблюдений получены в результате испытаний над случайными величинами с одним и тем же распределением вероятностей.

Н. В. Смирновым в работе [1] была доказана в этом направлении следующая важная теорема.

Пусть ξ — случайная величина с непрерывной функцией распределения $F(x) = P\{\xi < x\}$. Пусть

$$x_1, \ldots, x_{n_1};\ \ y_1, \ldots, y_{n_2} \tag{1}$$

результаты двух серий независимых наблюдений над величиною ξ. Обозначим через $S_{n_1}(x)$ и $T_{n_2}(x)$ — эмпирические функции распределения рядов (1)

$$S_{n_1}(x) = \frac{k_1(x)}{n_1}; \quad T_{n_2}(x) = \frac{k_2(x)}{n_2};$$

$k_1(x)$ — число x_i меньших x; $k_2(x)$ — число y_i меньших x. Введем в рассмотрение величины $D^+_{n_1n_2}$, $D_{n_1n_2}$:

$$D^+_{n_1n_2} = \sup_x \{S_{n_1}(x) - T_{n_2}(x)\},$$

$$D_{n_1n_2} = \sup_x |S_{n_1}(x) - T_{n_2}(x)|.$$

Тогда при условии $\lim\limits_{n_1\to\infty} \frac{n_2}{n_1} = \tau$ $(0 < \tau < \infty)$ имеют место соотношения

$$\lim_{n_1\to\infty} P\left\{\sqrt{\frac{n_1n_2}{n_1+n_2}}\, D^+_{n_1n_2} < z\right\} = 1 - e^{-2z^2} \quad \text{при} \quad z > 0, \tag{2}$$

$$\lim_{n_1\to\infty} P\left\{\sqrt{\frac{n_1n_2}{n_1+n_2}}\, D_{n_1n_2} < z\right\} = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2z^2}, \quad \text{при} \quad z > 0. \tag{3}$$

Эти результаты Н. В. Смирнова могут быть использованы для оценки правильности интересующей нас гипотезы тождественности распределений вероятностей случайных величин, в результате испытаний над которыми были получены ряды (1).

Для практического применения указанные результаты, однако, недостаточны, так как они, с одной стороны, носят лишь асимптотический характер и не улавливают влияния числа наблюдений, которое может быть весьма значительным, если числа n_1 и n_2 невелики. С другой стороны, меры расхождения эмпирических функций распределения — величины $D^+_{n_1n_2}$, $D_{n_1n_2}$, предложенные Н. В. Смирновым, ставят в равные условия произвольные участки изменения аргумента x, тогда как может оказаться целесообразным максимум расхождения между эмпирическими функциями распределения рассматривать лишь на отдельных участках, представляющих по той или иной причине наибольший интерес для исследователя, или на различных участках по-разному производить этот учет.

Таким образом, возникают две задачи, связанные с настоящим кругом вопросов: 1) определение распределения соответствующих вероятностей величин $D^+_{n_1n_2}$, $D_{n_1n_2}$ для конечных n_1 и n_2; 2) разыскание для конечных n_1 и n_2, а также асимптотических вероятностей попадания величин $D^+_{n_1n_2}$, $D_{n_1n_2}$ в полосу некоторой переменной ширины (например, суживающейся к концам интервала возможных значений случайной величины).

К этому же кругу вопросов относится и задача 3 — установления критериев, использующих более полно информацию материала, чем это позволяет сделать теорема Н. В. Смирнова и непосредственные ее обобщения в указанных двух направлениях.

Первые шаги к решению задачи 1 сделаны в работе [2], где дано распределение величин $D^+_{n_1n_2}$, $D_{n_1n_2}$. В работе [3] дано совместное распределение величин $D^+_{n_1n_2}$ и $D^-_{n_1n_2} = -\inf\limits_x \{S_{n_1}(x) - T_{n_2}(x)\}$ при частном предположении $n_1 = n_2$ и асимптотическое представление этого совместного распределения.

В работе [4] рассмотрен частный случай задачи 2: изучено асимптотическое распределение величины $D^+_{n_1n_2}$ и $D_{n_1n_2}$ на интервале (α, β) $(0 < \alpha < \beta < 1)$. В работе [5] в том же направлении получено асимптотическое распределение величин $D^+_{n_1n_2}$ и $D_{n_1n_2}$ вне некоторого интервала (α, β) $(0 < \alpha < \beta < 1)$.

В работе [6] наряду с другими результатами приведены некоторые новые предельные теоремы, относящиеся к рассматриваемому кругу вопросов и основанные на более полном использовании статистического материала, чем в ранее известных теоремах.

Использовав метод, изложенный в работах [2], [3], в настоящей работе для простейшего и практически наиболее важного случая одинаковой численности наблюдений в обеих сериях (1) мы даем не только предельные результаты работ [4], [5], [6], но и допредельные их формулировки; в § 2 и 3 в теоремах 3, 4, 7, 8 дано решение задачи 2 для случая разбиения интервала возможных значений на конечное число интервалов, в каждом из которых по-разному учитывается уклонение между эмпирическими функциями распределения. В теоремах 1, 2, 5, 6 рассмотрена та же задача разбиения интервала возможных значений случайной величины на конечное число интервалов, когда учитывается дополнительная информация — значения уклонений одной эмпирической кривой от другой на концах этих интервалов. Параграф 4 посвящен изучению уклонения между эмпирическими кривыми в данной точке.

§ 1. Обозначения и вспомогательные предложения

Пусть

$$x_1,\ x_2,\ \ldots,\ x_n;\ y_1,\ y_2,\ \ldots,\ y_n \tag{4}$$

результаты двух серий независимых наблюдений над случайными величинами ξ_1, ξ_2 с непрерывной функцией распределения $F(x)$. Обозначим через $S_n(x)$ и $T_n(x)$ соответствующие эмпирические функции распределения рядов (4), т. е. $S_n(x) = \frac{k_1(x)}{n}$, $k_1(x)$ — число x_i меньших x; $T_n(x) = \frac{k_2(x)}{n}$, $k_2(x)$ — число y_i меньших x.

Расположим результаты наблюдений (4) обеих серий в одну последовательность в порядке возрастания $z_1 < z_2 < \ldots < z_{2n}$. Пусть $a_i = \frac{i}{2n}$, $i = 0,\ 1, \ldots,\ 2n$.

Введем обозначения

$$\begin{aligned}
&\omega_n(a_j) = S_n(z_j) - T_n(z_j), \\
&D_n^+(a_p, a_q) = \max_{p \leqslant j \leqslant q} \omega_n(a_j), \\
&D_n^-(a_p, a_q) = \min_{p \leqslant j \leqslant q} \omega_n(a_j), \qquad (0 \leqslant p \leqslant q \leqslant 2n)\\
&D_n(a_p, a_q) = \max_{p \leqslant j \leqslant q} |\omega_n(a_j)|.
\end{aligned}$$

Пусть $p_1,\ p_2, \ldots,\ p_k$ — целые, $0 = p_0 < p_1 < \ldots < p_k < p_{k+1} = 2n$, $\omega_n(a_{p_i}) = \frac{r_i}{2n}$, $i = 1,\ 2, \ldots,\ k$; $a_{p_{i+1}} - a_{p_i} = a_i$. Числа p_i и r_i имеют попарно одинаковую четность, поэтому $na_i \pm \frac{r_i \pm r_{i+1}}{2}$ всегда целые числа. Обозначим

$$\varphi_{a_i}(c;\ r_i,\ r_{i+1}) = \begin{cases} 0, & \text{если } c \leqslant \max(r_i, r_{i+1}), \\ C_{2na_i}^{na_i - \frac{r_i - r_{i+1}}{2}} - C_{2na_i}^{na_i - \frac{r_i + r_{i+1}}{2} + c}, & c > \max(r_i, r_{i+1}). \end{cases}$$

Здесь и в дальнейшем считаем, что соответствующие биномиальные коэфициенты равны нулю, если верхний индекс оказывается отрицательным или большим нижнего. Таким образом, при $c > na_i + \frac{r_{i+1} + r_i}{2}$

$$\varphi_{a_i}(c;\ r_i,\ r_{i+1}) = C_{2na_i}^{na_i - \frac{r_i - r_{i+1}}{2}};$$

при $|r_i - r_{i+1}| > 2na_i$ $\varphi_{a_i}(c;\ r_i,\ r_{i+1}) = 0$. Заметим также, что

$$\varphi_{a_i}(c;\ r_i,\ r_{i+1}) = \varphi_{a_i}(c;\ r_{i+1},\ r_i).$$

Обозначим далее

$$\Phi_{a_i}(c,\ d;\ r_i,\ r_{i+1}) = \begin{cases} 0, \text{ если } c \leqslant \max(r_i, r_{i+1}) \text{ или } d \geqslant \min(r_i, r_{i+1}); \\ \sum\limits_j C_{2na_i}^{na_i + j(c-d) - \frac{r_i - r_{i+1}}{2}} - \sum\limits_j \Big(C_{2na_i}^{na_i + j(c-d) + c - \frac{r_i + r_{i+1}}{2}} + \\ + C_{2na_i}^{na_i - j(c-d) + d - \frac{r_i + r_{i+1}}{2}} \Big), \quad \begin{cases} c > \max(r_i, r_{i+1}), \\ d < \min(r_i, r_{i+1}). \end{cases} \end{cases}$$

Первая сумма распространена по всем j, при которых верхний индекс биномиальных коэфициентов, оставаясь неотрицательным, не превосходит нижний, а вторая — по всем таким неотрицательным значениям j. Очевидно также

$$\Phi_{a_i}(c, d; r_i, r_{i+1}) = \Phi_{a_i}(c, d; r_{i+1}, r_i)$$

и при $d < -na_i + \frac{r_i + r_{i+1}}{2}$

$$\Phi_{a_i}(c, d; r_i, r_{i+1}) = \varphi_{a_i}(c; r_i, r_{i+1}).$$

Действительно, при $d < -na_i + \frac{r_i + r_{i+1}}{2}$ отличными от нуля окажутся только соответствующие $j = 0$ слагаемое первой суммы и положительное слагаемое второй суммы. Например, при $j = 1$ верхний индекс соответствующего слагаемого первой суммы, который может быть записан как

$$na_i - j(c-d) + \frac{r_i - r_{i+1}}{2},$$

будет меньше нуля:

$$na_i - j(c-d) + \frac{r_i - r_{i+1}}{2} < na_i - c - na_i + \frac{r_i + r_{i+1}}{2} + \frac{r_i - r_{i+1}}{2} = -c + r_i < 0.$$

При $j = 0$ верхний индекс отрицательного слагаемого второй суммы также меньше нуля:

$$na_i + d - j(c-d) - \frac{r_i + r_{i+1}}{2} < na_i - na_i + \frac{r_i + r_{i+1}}{2} - \frac{r_i + r_{i+1}}{2} = 0.$$

При $j = 1$ верхний индекс положительного слагаемого второй суммы в силу условия $c > \max(r_i, r_{i+1})$ оказывается большим нижнего:

$$na_i + j(c-d) + c - \frac{r_i + r_{i+1}}{2} > na_i + 2c + na_i - \frac{r_i + r_{i+1}}{2} - \frac{r_i + r_{i+1}}{2} =$$

$$= 2na_i + 2c - r_i - r_{i+1} > 2na_i$$

и т. д.

Доказательство всех результатов этой работы основано на использовании схемы блуждания частицы, описанной в работах [2], [3].

Рассмотрим последовательность вспомогательных случайных величин

$$\xi_1, \xi_2, \ldots, \xi_{2n},$$

где

$$\xi_k = \begin{cases} +1, & \text{если } z_k \text{ принадлежит I серии наблюдений,} \\ -1, & \text{если } z_k \text{ принадлежит II серии наблюдений.} \end{cases}$$

Обозначим

$$s_k = \xi_1 + \xi_2 + \ldots + \xi_k.$$

Очевидно,

$$n\omega_n(\alpha_j) = s_j \qquad (j = 1, 2, \ldots, 2n).$$

Рассмотрим частицу, находящуюся в момент $t=0$ в положении $x=0$ и подверженную случайным толчкам в моменты $t=1, 2, \ldots, 2n$, в результате которых она может сдвигаться на $+1$ или -1. В плоскости (x, t) путь частицы при каждом толчке изобразится перемещением на единицу „вверх" и на единицу „вправо" или „влево".

Докажем следующие вспомогательные предложения.

Лемма 1. Число возможных путей блуждающей частицы из точки (r_i, p_i) в точку (r_{i+1}, p_{i+1}), проходящих „левее" прямой $x=c$, равно

$$\varphi_{a_i}(c;\ r_i,\ r_{i+1})$$

(употребляемые здесь обозначения имеют приданный им ранее смысл).

Лемма 2. Число возможных путей из точки (r_i, p_i), в точку (r_{i+1}, p_{i+1}), проходящих левее прямой (a) $x=c$ и правее прямой (b) $x=d$, равно Φ_{a_i} $(c, d;\ x_i, x_{i+1})$.

Доказательство леммы 1 использует прием, примененный в [2].

Легко видеть, что для того чтобы частица, вышедшая из точки (r_i, p_i), через $p_{i+1}-p_i=2na_i$ шагов попала в точку (r_{i+1}, p_{i+1}), необходимо, чтобы среди этих шагов „левых" оказалось на $r_{i+1}-r_i$ больше, чем „правых". Таким образом, число всех возможных путей от точки (r_i, p_i) к точке (r_{i+1}, p_{i+1}) равно

$$C_{2na_i}^{na_i-\frac{r_{i+1}-r_i}{2}}.$$

Поставим теперь в соответствие каждому пути, приводящему из точки (r_i, p_i) в точку (r_{i+1}, p_{i+1}) и достигающему прямую $x=c$, путь, расположенный симметрично относительно этой прямой $x=c$, начиная с момента этой встречи, и приводящий в точку, симметричную точке (r_{i+1}, p_{i+1}) относительно прямой $x=c$, т. е. в точку $(2c-r_{i+1}, p_{i+1})$. Не трудно подсчитать, что число таких путей равно

$$C_{2na_i}^{na_i-\frac{r_{i+1}+r_i}{2}+c}$$

(разность между числом „правых" и „левых" шагов равна $2c-r_{i+1}-r_i$).

Таким образом, число путей, исходящих из точки (r_i, p_i), приводящих в точку (r_{i+1}, p_{i+1}) и не достигающих прямой $x=c$, равно

$$C_{2na_i}^{na_i-\frac{r_{i+1}-r_i}{2}}-C_{2na_i}^{na_i-\frac{r_{i+1}+r_i}{2}+c},$$

если $c>\max(r_i, r_{i+1})$. Очевидно, что если $c\leqslant\max(r_i, r_{i+1})$, то не существует ни одной такой траектории.

Доказательство леммы 2 основано на приеме, использованном в работе [3].

Очевидно, как и в случае, рассмотренном в лемме 1, число всех путей, приводящих частицу из точки (r_i, p_i) в точку (r_{i+1}, p_{i+1}), равно

$$C_{2na_i}^{na_i-\frac{r_{i+1}-r_i}{2}}.$$

Разобьем множество всех путей $\mathfrak{A}$ на следующие непересекающиеся множества: $\mathfrak{M}_0$ — траектории, не пересекающие ни (a), ни (b); $\mathfrak{M}_1$ —

траектории, пересекающие (a), но не пересекающие (b); $\mathfrak{N}_1$ — траектории, пересекающие (b), но не пересекающие (a); $\mathfrak{M}_2$ — траектории, сначала пересекающие (a), затем (b) и более не пересекающие (a); $\mathfrak{N}_2$ — траектории, пересекающие сначала (b), затем (a) и более не пересекающие (b); $\mathfrak{M}_3$ — траектории, сначала пересекающие (a), затем (b), затем снова (a) и более не пересекающие (b) и т. д.

Очевидно, что, начиная с некоторого момента, так образуемые множества траекторий окажутся пустыми. Кроме того, очевидно, что

$$\mathfrak{A}=\mathfrak{M}_0+\Sigma(\mathfrak{M}_i+\mathfrak{N}_i).$$

Наряду с только что построенными множествами, образуем еще следующие множества траекторий: M_1 — пересекающих (a) по меньшей мере один раз; N_1 — пересекающих (b) по меньшей мере один раз; M_2 — пересекающих по меньшей мере один раз (a), а затем (b); N_2 — пересекающих по меньшей мере один раз (b), а затем (a); M_3 — пересекающих по меньшей мере один раз прямые (a) и (b) в порядке aba. Этот процесс продолжим до естественного исчерпания, так как имеют место равенства

$$M_1=\mathfrak{M}_1+\sum_{i=2}(\mathfrak{M}_i+\mathfrak{N}_i);\quad N_1=\mathfrak{N}_1+\sum_{i=2}(\mathfrak{M}_i+\mathfrak{N}_i);$$

$$M_2=\mathfrak{M}_2+\sum_{i=3}(\mathfrak{M}_i+\mathfrak{N}_i);\quad N_2=\mathfrak{N}_2+\sum_{i=3}(\mathfrak{M}_i+\mathfrak{N}_i)$$

и т. д., то при любом $i\geqslant 1$

$$(M_{2i-1}-M_{2i})+(N_{2i-1}-N_{2i})=\mathfrak{M}_{2i-1}+\mathfrak{N}_{2i-1}+\mathfrak{M}_{2i}+\mathfrak{N}_{2i},$$

и, следовательно,

$$\mathfrak{M}_0=\mathfrak{M}-\sum_{i=1}\{M_{2i-1}+N_{2i-1}-M_{2i}-N_{2i}\}.$$

Теперь нам остается подсчитать число путей в каждом из множеств

$$M_{2i-1},\ M_{2i},\ N_{2i-1},\ N_{2i}\qquad (i=1,2,\ldots).$$

Подсчитаем, например, число путей в множествах M_1 и M_2. Каждой траектории, исходящей из точки (r_i, p_i) и достигающей прямой $x=c$, поставим в соответствие, как это мы делали при доказательстве леммы 1, траекторию, исходящую из точки (r_i, p_i), идущую до прямой (a) по первоначальному пути, а от точки пересечения представляющую собой зеркальное отображение первоначального пути относительно прямой (a), и заканчивающуюся в точке $(2c-r_{i+1}, p_{i+1})$. Число таких траекторий равно

$$C_{2na_i}^{na_i-\frac{r_{i+1}+r_i}{2}+c}.$$

Если траектория достигает прямой (a), а затем прямой (b), то построенная только что указанным приемом отраженная траектория достигает прямой $x=2c-d$. Для подсчета таких траекторий отразим отраженную траекторию еще раз, начиная от точки пересечения с прямой $x=2c-d$, около этой прямой $x=2c-d$. Таким образом отраженная дважды траектория, начавшись в точке (r_i, p_i), закончится в точке

$(2c-2d+r_{i+1},\ p_{i+1})$. Число таких траекторий (достигающих (a), а затем (b)) равно

$$C_{2na_i}^{na_i+(c-d)+\frac{r_{i+1}-r_i}{2}}$$

$\left(\text{число положительных шагов равно } na_i+(c-d)+\frac{r_{i+1}-r_i}{2},\right.$ число отрицательных шагов равно $\left. na_i-(c-d)-\frac{r_{i+1}-r_i}{2}\right)$. Подобными же рассуждениями находим, что число траекторий в множествах M_{2j-1}, N_{2j-1}, M_{2j}, N_{2j} соответственно равно

$$C_{2na_i}^{na_i+c+j(c-d)-\frac{r_{i+1}+r_i}{2}};\quad C_{2na_i}^{na_i+d-j(c-d)-\frac{r_{i+1}+r_i}{2}};$$

$$C_{2na_i}^{na_i+j(c-d)+\frac{r_{i+1}-r_i}{2}};\quad C_{2na_i}^{na_i-j(c-d)+\frac{r_{i+1}-r_i}{2}}.$$

Этим лемма 2 доказана.

Лемма 3. Положим $c=[z\sqrt{2n}]$, $r_i=[x_i\sqrt{2n}]$; $a_i=a+o(1)$. В указанных ранее предположениях и обозначениях имеет место следующее предельное соотношение

$$f_a(z;\ x_i,\ x_{i+1})=\lim_{n\to\infty}\frac{\varphi_{a_i}(c;\ r_i,\ r_{i+1})}{C_{2na_i}^{na_i-\frac{r_i-r_{i+1}}{2}}}=$$

$$=\begin{cases}0, & z\leqslant\max(x_i,\ x_{i+1});\\ 1-e^{-\frac{2}{a}(z-x_i)(z-x_{i+1})}, & z>\max(x_i,\ x_{i+1}).\end{cases}$$

Доказательство. Рассмотрим отношение

$$I=\frac{C_{2na_i}^{na_i-\frac{r_i+r_{i+1}}{2}+c}}{C_{2na_i}^{na_i-\frac{r_i-r_{i+1}}{2}}}=\frac{\left(na_i-\frac{r_i-r_{i+1}}{2}\right)!\left(na_i+\frac{r_i-r_{i+1}}{2}\right)!}{\left(na_i-\frac{r_i+r_{i+1}}{2}+c\right)!\left(na_i+\frac{r_i+r_{i+1}}{2}-c\right)!}.$$

Используя формулу Стирлинга, находим

$$I=\frac{\left(1-\frac{r_i-r_{i+1}}{2na_i}\right)^{na_i-\frac{r_i-r_{i+1}}{2}}\left(1+\frac{r_i-r_{i+1}}{2na_i}\right)^{na_i+\frac{r_i-r_{i+1}}{2}}}{\left(1-\frac{r_i+r_{i+1}-2c}{2na_i}\right)^{na_i-\frac{r_i+r_{i+1}}{2}+c}\left(1+\frac{r_i+r_{i+1}-2c}{2na_i}\right)^{na_i+\frac{r_i+r_{i+1}}{2}-c}}(1+o(1));$$

разлагая в ряд Маклорена и имея в виду равенства $c=[z\sqrt{2n}]$, $r_i=[x_i\sqrt{2n}]$, получаем

$$\lg I=\frac{(x_i-x_{i+1})^2}{2a_i}-\frac{(x_i+x_{i+1}-2z)^2}{2a_i}+o(1),$$

откуда

$$I=e^{-\frac{2}{a_i}(z-x_i)(z-x_{i+1})}(1+o(1)),$$

что и доказывает наше предложение.

Лемма 4. Пусть $c=[z\sqrt{2n}]$, $d=[t\sqrt{2n}]$, $r_i=[x_i\sqrt{2n}]$, $a_i=a+o(1)$. В принятых обозначениях имеет место следующее предельное соотношение

$$F_a(z, t; x_i, x_{i+1}) = \lim_{n\to\infty} \frac{\Phi_{a_i}(c, d; r_i, r_{i+1})}{C_{2na_i}^{na_i-\frac{r_i-r_{i+1}}{2}}} =$$

$$= \begin{cases} 0, \text{ если } t \geqslant \min(x_i, x_{i+1}) \text{ или } z \leqslant \max(x_i, x_{i+1}); \\ e^{\frac{(x_i-x_{i+1})^2}{2a}} \sum\limits_{j=-\infty}^{\infty} \left(e^{-\frac{(x_i-x_{i+1}-2j(z-t))^2}{2a}} - e^{-\frac{(x_i+x_{i+1}-2z-2j(z-t))^2}{2a}}\right), \end{cases}$$

в остальных случаях.

Доказательство этого утверждения легко получаем, применив рассуждения, использованные при доказательстве леммы 3.

§ 2. Совместное распределение односторонних уклонений на различных вероятностных интервалах

Теорема 1. Пусть $c_i=[z_i\sqrt{2n}]$, $r_i=[x_i\sqrt{2n}]$, $i=1, 2, \ldots, k$. В принятых ранее обозначениях вероятность совместного распределения. величин $D_n^+(0, \alpha_1)$, $D_n^+(\alpha_1, \alpha_2), \ldots, D_n^+(\alpha_k, 1)$ при условии $\omega_n(\alpha_1)=\frac{r_1}{n}$, $\omega_n(\alpha_2)=\frac{r_2}{n}, \ldots, \omega_n(\alpha_k)=\frac{r_k}{n}$ равна

$$P\left\{\sqrt{\frac{n}{2}}D_n^+(0, \alpha_1)<z_0, \ldots, \sqrt{\frac{n}{2}}D_n^+(\alpha_k, 1)<z_k \Big/ \omega_n(\alpha_1)=\frac{r_1}{n}, \ldots, \omega_n(\alpha_k)=\frac{r_k}{n}\right\} =$$

$$= \frac{\varphi_{a_0}(c_0; 0, r_1)\,\varphi_{a_1}(c_1; r_1, r_2)\ldots\varphi_{a_k}(c_k; r_k, 0)}{C_{2na_0}^{na_0-\frac{r_1}{2}}\, C_{2na_1}^{na_1-\frac{r_2-r_1}{2}} \ldots C_{2na_k}^{na_k-\frac{r_k}{2}}}.$$

Доказательство. Действительно, в нашей схеме блуждания частицы условия $\omega_n(\alpha_1)=\frac{r_1}{n}$, $\omega_n(\alpha_2)=\frac{r_2}{n}, \ldots, \omega_n(\alpha_k)=\frac{r_k}{n}$ означают, что блуждающая частица, выходя из точки $(0,0)$ и приходя в точку $(0,2n)$, в моменты времени $t=p_1, \ldots, t=p_k$, соответственно, находилась в точках $(r_1, p_1), \ldots, (r_k, p_k)$ и в промежутки времени $p_1=2na_0$, $p_2-p_1=2na_1, \ldots, p_k-p_{k-1}=a_{k-1}$, $2n-p_k=a_k$ находилась левее соответствующих прямых $x=z_0, \ldots, x=z_k$. А так как число всех возможных путей от точки (x_i, p_i) до точки (x_{i+1}, p_{i+1}) $(i=0, \ldots, k)$ равно

$$C_{2na_i}^{na_i-\frac{r_i-r_{i+1}}{2}},$$

то наше утверждение легко следует из леммы 1.

Приведем некоторые частные результаты, вытекающие из теоремы 1, которые могут иметь непосредственное практическое применение в качестве критериев согласия.

а) Пусть $\alpha=\frac{p_1}{2n}$, $\beta=\frac{p_2}{2n}$; p_1, p_2 — целые, $0<p_1<p_2<2n$, $\gamma=\beta-\alpha$; $r=[x\sqrt{2n}]$, $s=[y\sqrt{2n}]$, $c=[z\sqrt{2n}]$. Для урезанного вероятностного интервала (α, β) условная вероятность односторонних уклонений $D_n^+(\alpha, \beta)$ одной эмпирической кривой от другой при условии $\omega_n(\alpha) = \frac{r}{n}$, $\omega_n(\beta) = \frac{s}{n}$ равна

$$P\left\{ \sqrt{\frac{n}{2}} D_n^+(\alpha, \beta) < z \Big/ \omega_n(\alpha) = \frac{r}{n};\ \omega_n(\beta) = \frac{s}{n} \right\} =$$

$$= \frac{\varphi_\gamma(c; r, s)}{C_{2n\gamma}^{n\gamma - \frac{r-s}{2}}} = \begin{cases} 0, & \text{если } c \leqslant \max(r, s); \\ 1 - \dfrac{C_{2n\gamma}^{n\gamma - \frac{r+s}{2} + c}}{C_{2n\gamma}^{n\gamma - \frac{r-s}{2}}}, & \text{если } c > \max(r, s). \end{cases}$$

Для доказательства достаточно положить в теореме 1:

$$k=2, \quad c_0 > na_0 + \frac{r}{2}, \quad c_2 > na_2 + \frac{s}{2}, \quad c_1 = c.$$

б) Для вероятностного интервала $\overline{(\alpha, \beta)}$, дополнительного к интервалу (α, β), та же вероятность равна

$$P\left\{ \sqrt{\frac{n}{2}} D_n^+\overline{(\alpha, \beta)} < z \Big/ \omega_n(\alpha) = \frac{r}{n};\ \omega_n(\beta) = \frac{s}{n} \right\} = \varphi_\alpha(c; 0, r) \cdot \varphi_{1-\beta}(c; s, 0) =$$

$$= \begin{cases} 0, & \text{если } c \leqslant \max(0, r, s); \\ \left(1 - \dfrac{C_{2n\alpha}^{n\alpha - \frac{r}{2} + c}}{C_{2n\alpha}^{n\alpha - \frac{r}{2}}}\right)\left(1 - \dfrac{C_{2n(1-\beta)}^{n(1-\beta) - \frac{s}{2} + c}}{C_{2n(1-\beta)}^{n(1-\beta) - \frac{s}{2}}}\right), & \text{если } c > \max(0, r, s). \end{cases}$$

Для получения этого равенства полагаем в теореме 1:

$$k=2, \quad c_1 > n\gamma + \frac{r+s}{2}, \quad c_0 = c_2 = c.$$

в) При $\alpha = \beta$ $(\gamma = 0,\ r = s)$ из б) получаем условное распределение величины $D_n^+(0,1)$ при условии, что $\omega_n(\alpha) = \frac{r}{n}$

$$P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, 1) < z \Big/ \omega_n(\alpha) = \frac{r}{n} \right\} = \frac{\varphi_\alpha(c; 0, r)\, \varphi_{1-\alpha}(c; r, 0)}{C_{2n\alpha}^{n\alpha - \frac{r}{2}}\, C_{2n(1-\alpha)}^{n(1-\alpha) - \frac{r}{2}}} =$$

$$= \frac{\varphi_\alpha(c; 0, r)}{C_{2n\alpha}^{n\alpha - \frac{r}{2}}} \frac{\varphi_{1-\alpha}(c; 0, r)}{C_{2n(1-\alpha)}^{n(1-\alpha) - \frac{r}{2}}}.$$

В частности, когда известно уклонение в средине вероятностного интервала $\left(\alpha = \frac{1}{2}\right)$, получаем

$$P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, 1) < z \Big/ \omega_n\left(\frac{1}{2}\right) = \frac{r}{n} \right\} = \left(\frac{\varphi_{\frac{1}{2}}(c; 0, r)}{C_n^{\frac{n-r}{2}}}\right)^2.$$

Теорема. 2. Пусть $a_i = \bar{a}_i + o(1)$. В указанных ранее предположениях имеет место следующее предельное соотношение

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, \alpha_1) < z_0, \ldots, \sqrt{\frac{n}{2}} D_n^+(\alpha_k, 1) < z_k \Big/ \omega_n(\alpha_1) = \frac{r_1}{n}; \ldots; \omega_n(\alpha_k) = \frac{r_k}{n} \right\} =$$
$$= f_{\bar{a}_0}(z_0; 0, x_1) f_{\bar{a}_1}(z_1; x_1, x_2) \ldots f_{\bar{a}_k}(z_k; x_k\ 0).$$

Доказательство этой теоремы следует непосредственно из теоремы 1 и леммы 3.

Приведем некоторые предельные результаты, легко следующие из теоремы 2.

а) В обозначениях, принятых ранее, при $\alpha = \bar{\alpha} + o(1)$, $\beta = \bar{\beta} + o(1)$, $\bar{\beta} - \bar{\alpha} = \gamma$

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}} D_n^+(\alpha, \beta) < z \Big/ \sqrt{n}\, \omega_n(\alpha) = \frac{[x\sqrt{2n}]}{\sqrt{n}}; \sqrt{n}\, \omega_n(\beta) = \frac{[y\sqrt{2n}]}{\sqrt{n}} \right\} = f_\gamma(z; x, y).$$

б) Для вероятностного интервала $\overline{(\alpha, \beta)}$, дополнительного к интервалу (α, β)

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}} D_n^+\overline{(\alpha, \beta)} < z \Big/ \sqrt{n}\, \omega_n(\alpha) = \frac{[x\sqrt{2n}]}{\sqrt{n}};\ \sqrt{n}\, \omega_n(\beta) = \frac{[y\sqrt{2n}]}{\sqrt{n}} \right\} =$$
$$= f_{\bar{\alpha}}(z; 0, x) f_{1-\bar{\beta}}(z; 0, y).$$

в) Для условного распределения величины $D_n^+(0, 1)$ при условии $\omega_n(\alpha) = \frac{r}{n}$ имеет место предельное соотношение

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, 1) < z \Big/ \sqrt{n}\, \omega_n(\alpha) = \frac{[x\sqrt{2n}]}{\sqrt{n}} \right\} = f_{\bar{\alpha}}(z; 0, x) f_{1-\bar{\alpha}}(z; 0, x),$$

в частности, при $\alpha = \frac{1}{2}$ получаем

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, 1) \Big/ \sqrt{n}\, \omega_n\left(\frac{1}{2}\right) = \frac{[x\sqrt{2n}]}{\sqrt{n}} \right\} = (f_{\frac{1}{2}}(z; 0, x))^2.$$

Теорема 3. Вероятность совместного распределения величин $D_n^+(0, \alpha_1)$, $D_n^+(\alpha_1, \alpha_2), \ldots, D_n^+(\alpha_k, 1)$ равна

$$P\left\{ \sqrt{\frac{n}{2}} D_n^+(0, \alpha_1) < z_0;\ \sqrt{\frac{n}{2}} D_n^+(\alpha_1, \alpha_2) < z_1; \ldots;\ \sqrt{\frac{n}{2}} D_n^+(\alpha_k, 1) < z_k \right\} =$$
$$= \frac{1}{C_{2n}^n} \sum_{(r_1, r_2, \ldots, r_k)} \varphi_{a_0}(c_0; 0, r_1) \cdot \varphi_{a_1}(c_1; r_1, r_2) \ldots \varphi_{a_k}(c_k; r_k\ 0),$$

где суммирование проводится по всем тем значениям $r_1, r_2, \ldots, r_k$, при которых соответствующие биномиальные коэфициенты не теряют смысла.

Доказательство этого предложения непосредственно следует из теоремы 1 и теоремы сложения вероятностей.

Из теоремы 3, в частности, имеем.

а) Вероятность односторонних уклонений двух эмпирических распределений на урезанном вероятностном интервале (α, β) $(\beta - \alpha = \gamma)$, равна

$$P\left\{\sqrt{\frac{n}{2}} D_n^+(\alpha, \beta) < z\right\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r, s)} \varphi_\alpha \left(n\alpha + \frac{r}{2} + 1; 0, r\right) \varphi_\gamma(c; r, s)\, \varphi_{1-\beta}\left(n\alpha + \frac{s}{2} + 1; s, 0\right) =$$

$$= \begin{cases} 0, \text{ если } c \leqslant \max(-\varrho, -\tau) \\ \dfrac{1}{C_{2n}^n} \sum\limits_{\substack{s \geq -\tau \\ (s \equiv 2n\beta \,(\mathrm{mod}\, 2))}}^{\min(\tau, c-1)} \sum\limits_{\substack{r \geq \max(-\varrho;\, s-2n\gamma) \\ r \equiv 2n\alpha \,(\mathrm{mod}\, 2)}}^{\min(c-1;\, \varrho;\, s+2n\gamma)} C_{2n\alpha}^{n\alpha - \frac{r}{2}} C_{2n(1-\beta)}^{n(1-\beta) - \frac{s}{2}} \left(C_{2n\gamma}^{n\gamma - \frac{r-s}{2}} - C_{2n\gamma}^{n\gamma - \frac{r+s}{2} + c}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad \max(-\varrho, -\tau) < c, \end{cases}$$

где

$$\varrho = \begin{cases} 2n\alpha, & \text{если } \alpha \leqslant \dfrac{1}{2}; \\ 2n(1-\alpha), & \text{если } \alpha > \dfrac{1}{2}; \end{cases} \qquad \tau = \begin{cases} 2n\beta, & \beta \leqslant \dfrac{1}{2}; \\ 2n(1-\beta), & \beta > \dfrac{1}{2}. \end{cases} \tag{5}$$

б) Как следствие а) получаем функцию распределения величины $\omega_n(\alpha)$. Полагаем $\alpha = \beta$, получаем

$$D_n^+(\alpha, \beta) = \omega_n(\alpha)$$

и, следовательно,

$$P\left\{\sqrt{\frac{n}{2}}\, \omega_n(\alpha) < z\right\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{r=-\tau}^{c-1} \varphi_\alpha\left(n\alpha + \frac{r}{2} + 1; 0, r\right) \varphi_{1-\alpha}\left(n\alpha + \frac{r}{2} + 1; r, 0\right) =$$

$$= \begin{cases} 0, \quad \text{если } c \leqslant -\tau; \\ \sum\limits_{\substack{s \geq -\tau \\ (s \equiv 2n\alpha \,(\mathrm{mod}\, 2))}}^{c-1} \dfrac{C_{2n\alpha}^{n\alpha - \frac{s}{2}} C_{2n(1-\alpha)}^{n(1-\alpha) - \frac{s}{2}}}{C_{2n}^n}, \quad -\tau < c. \end{cases}$$

В частности, при любом n и α $\left(\alpha = \dfrac{p}{2n},\ p \text{ — целое, } 0 \leqslant p \leqslant 2n\right)$

$$P\left\{\sqrt{\frac{n}{2}}\, \omega_n(\alpha) < 0\right\} = \frac{1}{2}.$$

При $\alpha = \dfrac{1}{2}$ $(c = [z\sqrt{2n}])$

$$P\left\{\sqrt{\frac{n}{2}}\, \omega_n\left(\frac{1}{2}\right) < z\right\} = \begin{cases} 0, \quad \text{если } c \leqslant \tau; \\ \dfrac{1}{C_{2n}^n} \sum\limits_{\substack{s \geq -n \\ (s \equiv n \,(\mathrm{mod}\, 2))}}^{c-1} \left(C_n^{\frac{n-s}{2}}\right)^2. \end{cases}$$

в) При $\alpha=0,\ \beta=1\ (\gamma=1)$ из а) получаем распределение величины $D_n^+=D_n^+(0,1)$ — результат, содержащийся в работе [2].

$$P\left\{\sqrt{\frac{n}{2}}D_n^+<z\right\}=\begin{cases}0, & c\leqslant 0;\\ 1-\dfrac{C_{2n}^{n-c}}{C_{2n}^{n}}, & c>0.\end{cases}$$

г) Для вероятностного интервала $\overline{(\alpha,\beta)}$, дополнительного к интервалу (α,β),

$$P\left\{\sqrt{\frac{n}{2}}D_n^+\overline{(\alpha,\beta)}<z\right\}=$$

$$=\frac{1}{C_{2n}^n}\sum_{(r,s)}\varphi_\alpha(c;\ 0,\ r)\,\varphi_\gamma\left(n\gamma+\frac{r+s}{2}+1;\ r,\ s\right)\varphi_{1-\beta}(c;\ s,\ 0)=$$

$$=\begin{cases}0, \quad \text{если}\quad c\leqslant 0;\\ \displaystyle\sum_{\substack{s\geq-\tau\\(s\equiv 2n\beta\,(\mathrm{mod}\ 2))}}^{\min(\tau,\,c-1)}\ \sum_{\substack{r\geq\max(-\varrho,\,s-2n\gamma)\\(r\equiv 2n\alpha\,(\mathrm{mod}\ 2))}}^{\min(\varrho,\,c-1,\,s+2n\gamma)}\frac{C_{2n\gamma}^{n\gamma-\frac{r-s}{2}}}{C_{2n}^n}\left(C_{2n\alpha}^{n\alpha-\frac{r}{2}}-C_{2n\alpha}^{n\alpha+c-\frac{r}{2}}\right)\times\\ \qquad\times\left(C_{2n(1-\beta)}^{n(1-\beta)-\frac{s}{2}}-C_{2n(1-\beta)}^{n(1-\beta)+c-\frac{s}{2}}\right), \quad c>0;\end{cases}$$

где ϱ и τ определены как в (5).

Теорема 4. Пусть $c_i=[z_i\sqrt{2n}]$, $\alpha_i=\overline{\alpha}_i+o(1)$, $i=0,1,\ldots,k$; $a_i=\overline{\alpha}_{i+1}-\overline{\alpha}_i$. Для совместного распределения величин $D_n^+(0,\alpha_1)$, $D_n^+(\alpha_1,\alpha_2),\ldots,D_n^+(\alpha_k,1)$ имеет место следующее предельное соотношение

$$\lim_{n\to\infty}P\left\{\sqrt{\frac{n}{2}}D_n^+(0,\alpha_1)<z_0;\ldots;\sqrt{\frac{n}{2}}D_n^+(\alpha_k,1)<z_k\right\}=$$

$$=\frac{1}{\sqrt{(2\pi)^k a_0a_1\ldots a_k}}\int_{-\infty}^{b_k}\int_{-\infty}^{b_{k-1}}\cdots\int_{-\infty}^{b_1}e^{-\frac{1}{2}s(x_1,\ldots,x_k)}f_{a_0}(z_0;\ 0,\ x_1)\times$$

$$\times f_{a_1}(z_1;\ x_1,\ x_2)\ldots f_{a_k}(z_k;\ x_k,\ 0)\,dx_1\ldots dx_k,$$

где

$$s(x_1,\ldots,x_k)=\frac{x_1^2}{a_0}+\frac{(x_2-x_1)^2}{a_1}+\ldots+\frac{(x_k-x_{k-1})^2}{a_{k-1}}+\frac{x_k^2}{a_k};$$

$$b_1=\min(z_0,\ z_1);\ b_2=\min(z_1,\ z_2);\ \ldots;\ b_k=\min(z_{k-1};\ z_k).$$

Доказательство. В силу теоремы 3 нам достаточно рассмотреть предел при $n\to\infty$ величины

$$\frac{1}{C_{2n}^n}\sum_{(r_1,r_2,\ldots,r_k)}\varphi_{a_0}(c_0;\ 0,\ r_1)\,\varphi_{a_1}(c_1;\ r_1,\ r_2)\ldots\varphi_{a_k}(c_k;\ r_k,\ 0)=$$

$$=\frac{1}{C_{2n}^n}\sum_{(r_1,r_2,\ldots,r_k)}\frac{\varphi_{a_0}(c_0;0,r_1)}{C_{2na_0}^{na_0-\frac{r_1}{2}}}\cdot\frac{\varphi_{a_1}(c_1;r_1,r_2)}{C_{2na_1}^{na_1-\frac{r_2-r_1}{2}}}\cdots\frac{\varphi_{a_k}(c_k;r_k,0)}{C_{2na_k}^{na_k-\frac{r_k}{2}}}\times$$

$$\times C_{2na_0}^{na_0-\frac{r_1}{2}}\ldots C_{2na_k}^{na_k-\frac{r_k}{2}}.$$

Рассмотрим отношение

$$I = \frac{C_{2na_0}^{na_0-\frac{r_1}{2}} C_{2na_1}^{na_1-\frac{r_2-r_1}{2}} \dots C_{2na_k}^{na_k-\frac{r_k}{2}}}{C_{2n}^{a}} =$$

$$= \frac{(2na_0)!\,(2na_1)!\dots(2na_k)!\,(n!)^2}{\left(na_0-\frac{r_1}{2}\right)!\left(na_0+\frac{r_1}{2}\right)!\dots\left(na_k-\frac{r_k}{2}\right)!\left(na_k+\frac{r_k}{2}\right)!\,(2n)!}.$$

По формуле Стирлинга находим

$$I = \frac{1}{\sqrt{(2n)^k a_0 a_1 \dots a_k}}\left(\sqrt{\frac{2}{n}}\right)^k \times$$

$$\times \frac{1+o(1)}{\left(1-\frac{r_1}{2na_0}\right)^{na_0-\frac{r_1}{2}}\left(1+\frac{r_1}{2na_0}\right)^{na_0+\frac{r_1}{2}}\left(1-\frac{r_2-r_1}{2na_1}\right)^{na_1-\frac{r_2-r_1}{2}}\dots\left(1+\frac{r_k}{2na_k}\right)^{na_k+\frac{r_k}{2}}}. \tag{6}$$

Разложив в ряд Маклорена знаменатель I и приняв во внимание равенства $c_i=[z_i\sqrt{2n}]$, $r_i=[x_i\sqrt{2n}]$, получаем

$$\lg\left[\left(1-\frac{r_1}{2na_0}\right)^{na_0-\frac{r_1}{2}}\dots\left(1+\frac{r_k}{2na_k}\right)^{na_k+\frac{r_k}{2}}\right] =$$

$$= \left(\frac{x_1^2}{2a_0}+\frac{(x_2-x_1)^2}{2a_1}+\dots+\frac{x_k^2}{2a_k}\right)(1+o(1)). \tag{7}$$

Положим теперь $\Delta x_1=\Delta x_2=\dots=\Delta x_k=\sqrt{\frac{2}{n}}$ пределы изменения x_i будут определяться неравенством $-n\leqslant[x_i\sqrt{2n}]\leqslant[z_i\sqrt{2n}]$; примем во внимание, что в силу леммы 3

$$\lim_{n\to\infty}\frac{\varphi_{a_i}(c_i;\,r_i,\,r_{i+1})}{C_{2na_i}^{na_i-\frac{r_{i+1}-r_i}{2}}} = f_{a_i}^{*}(z_i;\,x_i,\,x_{i+1}). \tag{8}$$

Сопоставляя соотношение (6), (7), (8), получаем доказательство нашей теоремы.

Из теоремы 4 легко получить следующие результаты.

а) Для урезанного вероятностного интервала (α,β), $(\alpha=\bar{\alpha}+o(1)$, $\beta=\bar{\beta}+o(1))$

$$P\left\{\sqrt{\frac{n}{2}}\,D_n^+(\alpha,\ \beta)\leqslant z\right\}\xrightarrow[n\to\infty]{}\Phi^+(\alpha,\,\beta;\,z),$$

где

$$\Phi^+(\alpha,\,\beta;\,z) = \frac{1}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{a}\int_{-\infty}^{b} e^{-\frac{1}{2}\bar{Q}(z_1,\,z_2)}\,dz_1\,dz_2 -$$

$$-\frac{e^{-2z^2}}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{a'}\int_{-\infty}^{b'} e^{-\frac{1}{2}Q(z_1,\,z_2)}\,dz_1\,dz_2,$$

$$a=\frac{z}{\sqrt{\overline{\alpha}(1-\overline{\alpha})}};\quad b=\frac{z}{\sqrt{\overline{\beta}(1-\overline{\beta})}};\quad R=\sqrt{\frac{\overline{\alpha}(1-\overline{\beta})}{\overline{\beta}(1-\overline{\alpha})}};\quad a'=\frac{z-2z\overline{\alpha}}{\sqrt{\overline{\alpha}(1-\overline{\alpha})}};$$

$$b'=\frac{z-2z(1-\overline{\beta})}{\sqrt{\overline{\beta}(1-\overline{\beta})}},\quad Q(z_1, z_2)=$$

$$=\frac{1}{1-R^2}[z_1^2+2Rz_1z_2+z_2^2];\quad \overline{Q}(z_1, z_2)=Q(z_1, -z_2). \tag{9}$$

Это утверждение содержится в общем случаем $n_1\neq n_2$ в работе [4].

б) Из следствия б) теоремы 3 легко получаем $(\alpha=\overline{\alpha}+o(1))$

$$P\left\{\sqrt{\frac{n}{2}}\,\omega_n(\alpha)<z\right\}\underset{n\to\infty}{\to}\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\frac{z}{\sqrt{\overline{\alpha}(1-\overline{\alpha})}}}e^{-\frac{x^2}{2}}dx.$$

в) Для вероятностного интервала $\overline{(\alpha, \beta)}$, дополнительного к интервалу (α, β)

$$P\left\{\sqrt{\frac{n}{2}}\,D_n^+\overline{(\alpha, \beta)}<z\right\}\underset{n\to\infty}{\to}K^+(\alpha, \beta; z),$$

где

$$K^+(\alpha, \beta; z)=\frac{1}{2\pi\sqrt{1-R^2}}\int\limits_{-\infty}^{a}\int\limits_{-\infty}^{b}e^{-\frac{1}{2}\overline{Q}(z_1, z_2)}\,dz_1\,dz_2-$$

$$-\frac{e^{-2z^2}}{2\pi\sqrt{1-R^2}}\left\{\int\limits_{-\infty}^{-a'}\int\limits_{-\infty}^{-b'}e^{-\frac{1}{2}Q(z_1, z_2)}\,dz_1\,dz_2+\int\limits_{-\infty}^{a'}\int\limits_{-\infty}^{b'}e^{-\frac{1}{2}\overline{Q}(z_1, z_2)}\,dz_1\,dz_2\right\},$$

R, a, a', b, b', $\overline{Q}(z_1, z_2)$ — определены как в (9).

Это утверждение содержится в общем случае $n_1\neq n_2$ в работе [5].

§ 3. Совместное распределение двусторонних уклонений на различных вероятностных интервалах

Теорема 5. Вероятность совместного распределения величин $D_n^+(0, \alpha_1)$, $D_n^-(0, \alpha_1)$, $D_n^+(\alpha_1, \alpha_2), \ldots, D_n^+(\alpha_k, 1)$, $D_n^-(\alpha_k, 1)$ при условии $\omega_n(\alpha_1)=\frac{r_1}{n}$, $\omega_n(\alpha_2)=\frac{r_2}{n}, \ldots, \omega_n(\alpha_k)=\frac{r_k}{n}$, равна

$$P\left\{\sqrt{\frac{n}{2}}\,D_n^+(0, \alpha_1)<z_0;\ \sqrt{\frac{n}{2}}\,D_n^-(0, \alpha_1)>t_0; \ldots;\right.$$

$$\left.\sqrt{\frac{n}{2}}\,D_n^-(\alpha_k, 1)>t_k\Big/\omega_n(\alpha_1)=\frac{r_1}{n}; \ldots; \omega_n(\alpha_k)=\frac{r_k}{n}\right\}=$$

$$=\frac{\Phi_{a_0}(c_0, d_0; 0, r_1)}{C_{2na_0}^{na_0-\frac{r_1}{2}}}\,\frac{\Phi_{a_1}(c_1, d_1; r_1, r_2)}{C_{2na_1}^{na_1-\frac{r_2-r_1}{2}}}\cdots\frac{\Phi_{a_k}(c_k, d_k; r_k, 0)}{C_{2na_k}^{na_k-\frac{r_k}{2}}}.$$

Доказательство этой теоремы легко следует из леммы 2.

Укажем на некоторые соотношения, вытекающие из теоремы 5.

а) Для урезанного вероятностного интервала (α, β) $(\beta-\alpha=\gamma)$

$$P\left\{\sqrt{\frac{n}{2}}D_n^+(\alpha,\beta)<z;\ \sqrt{\frac{n}{2}}D_n^-(\alpha,\beta)>t\Big/\omega_n(\alpha)=\frac{r}{n};\ \omega_n(\beta)=\frac{s}{n}\right\}=$$

$$=\frac{\Phi_\gamma(c, d; r, s)}{C_{2n\gamma}^{n\gamma-\frac{r-s}{2}}}.$$

б) Из а) при $z=-t$ получаем условное распределение величины $D_n(\alpha, \beta)$ при условии $\omega_n(\alpha)=\frac{r}{n},\ \ \omega_n(\beta)=\frac{s}{n}$:

$$P\left\{-z<\sqrt{\frac{n}{2}}D_n(\alpha,\beta)<z\Big/\omega_n(\alpha)=\frac{r}{n};\ \omega_n(\beta)=\frac{s}{n}\right\}=\frac{\Phi_\gamma(c, -c; r, s)}{C_{2n\gamma}^{n\gamma-\frac{r-s}{2}}}=$$

$$=\begin{cases}0, & \text{если}\ \ c\leqslant\max(|r|, |s|);\\ \dfrac{1}{C_{2n\gamma}^{n\gamma-\frac{r-s}{2}}}\displaystyle\sum_{(i)}(-1)^i C_{2n\gamma}^{n\gamma+ic-\frac{r-s}{2}}, & c>\max(|r|, |s|).\end{cases}$$

в) При $d<\frac{r+s}{2}-n\gamma$ из а) получаем условное распределение величины $D_n^+(\alpha, \beta)$ при тех же условиях.

г) Для вероятностного интервала $\overline{(\alpha, \beta)}$

$$P\left\{\sqrt{\frac{n}{2}}D_n^+\overline{(\alpha,\beta)}<z;\ \sqrt{\frac{n}{2}}D_n^-\overline{(\alpha,\beta)}>t\Big/\omega_n(\alpha)=\frac{r}{n};\ \omega_n(\beta)=\frac{s}{n}\right\}=$$

$$=\frac{\Phi_\alpha(c, d; 0, r)\,\Phi_{1-\beta}(c, d; s, 0)}{C_{n\alpha}^{n\alpha-\frac{r}{2}}\quad C_{2n(1-\beta)}^{n(1-\beta)-\frac{s}{n}}}.$$

д) Из г) при $z=-t$ получаем условное распределение величины $D_n\overline{(\alpha, \beta)}$

$$P\left\{-z<\sqrt{\frac{n}{2}}D_n(\alpha,\beta)<z\Big/\omega_n(\alpha)=\frac{r}{n};\ \omega_n(\beta)=\frac{s}{n}\right\}=$$

$$=\Phi_\alpha(c, -c; 0, r)\,\Phi_{1-\beta}(c, -c; s, 0)\Big/C_{2n\alpha}^{n\alpha-\frac{r}{2}}\,C_{2n(1-\beta)}^{n(1-\beta)-\frac{s}{n}}.$$

Теорема 6. Пусть $\alpha_i=\overline{\alpha}_i+o(1)$, $\overline{\alpha}_{i+1}-\overline{\alpha}_i=a_i$. В принятых нами обозначениях имеет место следующее предельное соотношение

$$\lim_{n\to\infty}P\left\{\sqrt{\frac{n}{2}}D_n^+(0,\alpha_1)<z_0;\ \sqrt{\frac{n}{2}}D_n^-(0,\alpha_1)>t_0;\ldots;\right.$$

$$\left.\sqrt{\frac{n}{2}}D_n^-(\alpha_k, 1)>t_k\Big/\omega_n(\alpha_1)=\frac{r_1}{n};\ldots\right\}=$$

$$=F_{a_0}(z_0, t_0; 0, x_1)\,F_{a_1}(z_1, t_1; x_1, x_2)\ldots F_{a_k}(z_k, t_k; x_k, 0).$$

Доказательство этого предложения легко следует из леммы 3.

Из теоремы 6 получаем, в частности.

а) Для урезанного вероятностного интервала (α, β), $(\alpha = \bar{\alpha} + o(1)$, $\beta = \bar{\beta} + o(1)$, $\gamma = \bar{\beta} - \bar{\alpha})$

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}}\, D_n^+(\alpha, \beta) < z;\right.$$

$$\left.\sqrt{\frac{n}{2}}\, D_n^-(\alpha, \beta) > t \Big/ \sqrt{n}\,\omega_n(\alpha) = \frac{[x\sqrt{2n}]}{\sqrt{n}};\ \sqrt{n}\,\omega_n(\beta) = \frac{[y\sqrt{2n}]}{\sqrt{n}}\right\} = F_\gamma(z, t;\, x, y) =$$

$$= \begin{cases} 0, \text{ если } t \geqslant \min(x, y) \text{ или } z \leqslant \max(x, y); \\ e^{\frac{(x-y)^2}{2\gamma}} \sum\limits_{-\infty}^{\infty} \left(e^{-\frac{(x-y-2i(z-t))^2}{2\gamma}} - e^{-\frac{(x+y-2z-2i(z-t))^2}{2\gamma}} \right) \text{ при } \begin{cases} t < \min(x, y); \\ z > \max(x, y). \end{cases} \end{cases}$$

б) Из а) при $z = -t$ получаем предельное соотношение для условного распределения величины $D_n(\alpha, \beta)$

$$\lim_{n\to\infty} P\left\{ -z < D_n(\alpha, \beta) < z \Big/ \sqrt{n}\,\omega_n(\alpha) = \frac{r}{n};\ \sqrt{n}\,\omega_n(\beta) = \frac{s}{n} \right\} =$$

$$= \begin{cases} 0, \text{ если } z \leqslant \max(|x|, |y|); \\ e^{\frac{(x-y)^2}{2\gamma}} \sum\limits_{j=-\infty}^{\infty} (-1)^j e^{-\frac{(x-y-2iz)^2}{2\gamma}}, \quad z > \max(|x|, |y|). \end{cases}$$

в) Для интервала $\overline{(\alpha, \beta)}$ $(\alpha = \bar{\alpha} + o(1),\ \beta = \bar{\beta} + o(1))$

$$\lim_{n\to\infty} P\left\{ \sqrt{\frac{n}{2}}\, D_n^+\overline{(\alpha, \beta)} < z;\right.$$

$$\left.\sqrt{\frac{n}{2}}\, D_n^-\overline{(\alpha, \beta)} > t \Big/ \sqrt{n}\,\omega_n(\alpha) = \frac{r}{n};\ \sqrt{n}\,\omega_n(\beta) = \frac{s}{n} \right\} =$$

$$= F_\alpha(z, t;\, 0, x)\, F_{1-\beta}(z, t;\, y, 0),$$

$$r = [x\sqrt{2n}], \quad s = [y\sqrt{2n}].$$

г) При $z = -t$ из в) получаем предельное соотношение для условного распределения величины $D_n\overline{(\alpha, \beta)}$.

Теорема 7. В принятых предположениях

$$P\left\{ \sqrt{\frac{n}{2}}\, D_n^+(0, \alpha_1) < z_0;\right.$$

$$\left.\sqrt{\frac{n}{2}}\, D_n^-(0, \alpha_1) > t_0; \ldots;\ \sqrt{\frac{n}{2}}\, D_n^+(\alpha_k, 1) < z_k;\ \sqrt{\frac{n}{2}}\, D_n^-(\alpha_k, 1) > t_k \right\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r_1, r_2, \ldots, r_k)} \Phi_{\alpha_0}(c_0, d_0;\, 0, r_1)\, \Phi_{\alpha_1}(c_1, d_1;\, r_1, r_2) \ldots \Phi_{\alpha_k}(c_k, d_k;\, r_k, 0),$$

где суммирование проводится по всем тем значениям $r_1, r_2, \ldots, r_k$, при которых соответствующие биномиальные коэфициенты не теряют смысла.

Доказательство этой теоремы легко следует из теоремы 5 и формулы полной вероятности.

Из теоремы 7, в частности, имеем:

а) для урезанного вероятностного интервала (α, β)

$$P\left\{\sqrt{\frac{n}{2}}\, D_n^+(\alpha, \beta) < z;\ \sqrt{\frac{n}{2}}\, D_n^-(\alpha, \beta) > t\right\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r, s)} C_{2n\alpha}^{n\alpha - \frac{r}{2}}\, C_{2n(1-\beta)}^{n(1-\beta) - \frac{s}{2}}\, \Phi_\gamma(c, d; r, s);$$

б) для вероятностного интервала $\overline{(\alpha, \beta)}$

$$P\left\{\sqrt{\frac{n}{2}}\, D_n^+\overline{(\alpha, \beta)} < z;\ \sqrt{\frac{n}{2}}\, D_n^-\overline{(\alpha, \beta)} > t\right\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r, s)} \Phi_\alpha(c, d; 0, r)\, C_{2n\gamma}^{n\gamma - \frac{r-s}{2}}\, \Phi_{1-\beta}(c, d; s, 0).$$

в) При $c = -d$ из а) и б) получаем распределение величин $D_n(\alpha, \beta)$ и $D_n\overline{(\alpha, \beta)}$.

д) Полагая $\gamma = 1$ $(\alpha = 0,\ \beta = 1)$ из а) получаем совместное распределение величин $D_n^+(0, 1)$ и $D_n^-(0, 1)$ — результат, содержащийся в работе [3]; полагая, кроме того, $c = -d$, получаем распределение величины $D_n(0, 1)$, а при $d < n$ — величины $D_n^+(0, 1)$ — результаты, содержащиеся в [2].

Теорема 8. Для рассмотренной в теореме 7 вероятности совмещения событий $D_n^+(0, \alpha_1) < z_0$, $D_n^-(0, \alpha_1) > t_1$, $D_n^+(\alpha_1, \alpha_2) < z_1, \ldots, D_n^-(\alpha_k, 1) > t_k$ имеет место следующее предельное соотношение

$$\lim_{n\to\infty} P\left\{\sqrt{\frac{n}{2}}\, D_n^+(0, \alpha_1) < z_0;\right.$$

$$\left.\sqrt{\frac{n}{2}}\, D_n^-(0, \alpha_1) > t_0; \ldots;\ \sqrt{\frac{n}{2}}\, D_n^-(\alpha_k, 1) > t_k\right\} =$$

$$= \frac{1}{\sqrt{(2\pi)^k \prod\limits_{i=0}^{k} a_i}} \int\limits_{c_1}^{b_1} \int\limits_{c_2}^{b_2} \cdots \int\limits_{c_k}^{b_k} e^{-\frac{1}{2} s(x_1, \ldots, x_k)} \times$$

$$\times F_{a_0}(z_0, t_0; 0, x_1)\, F_{a_1}(z_1, t_1; x_1, x_2) \ldots F_{a_k}(z_k, t_k; x_k, 0)\, dx_1 \ldots dx_k,$$

где

$$s(x_1, \ldots, x_k) = \frac{x_1^2}{a_0} + \frac{(x_2 - x_1)^2}{a_1} + \ldots + \frac{(x_k - x_{k-1})^2}{a_{k-1}} + \frac{x_k^2}{a_k};$$

$$b_i = \min(z_{i-1}, z_i); \quad c_i = \min(t_{i-1}, t_i).$$

Доказательство этого предложения легко следует из теоремы 7 и леммы 4.

Отсюда, в частности, получаем для случая $n_1 = n_2$ результат, содержащийся в работе [6].

Теорема 9. Пусть p — некоторое фиксированное целое число, $0 < p \leqslant 2n$, $\alpha_i = \frac{i}{2n}$, $i = 0, 1, \ldots, 2n$. В принятых предположениях

$$P\left\{\omega_n(\alpha_p) \geqslant \frac{c}{n} \Big/ \omega_n(\alpha_i) \leqslant \omega_n(\alpha_p), \quad i = 0, 1, \ldots, 2n\right\} =$$

$$= \frac{n+1}{C_{2n}^n} \sum_{\substack{r \geqslant c \\ [r \equiv p \,(\mathrm{mod}\ 2)]}} \varphi_{\alpha_{p-1}}(r+1; 0, r-1)\, \varphi_{1-\alpha_{p+1}}(r+1; r-1, 0) =$$

$$= \begin{cases} 0, \text{ если } c > \varrho; \\ \dfrac{4(n+1)}{C_{2n}^n} \displaystyle\sum_{\substack{x \geqslant c \\ [x \equiv p \,(\mathrm{mod}\ 2)]}} \dfrac{(x+1)^2}{(x+p+2)(2n-p+x+2)} C_p^{\frac{p-x}{2}} C_{2n-p}^{n-\frac{p+x}{2}}, \quad c \leqslant \varrho, \end{cases}$$

где ϱ определено как в (5).

Для доказательства этого предложения обратимся к изложенной ранее схеме блуждания частицы. В нашей интерпретации рассматриваемое условие означает, что в данном случае возможными траекториями блуждающей частицы являются траектории, у которых какая-либо из точек (x, p), $(x \equiv p \,(\mathrm{mod}\ 2))$ является одной из „наиболее правых“ (т. е. максимум односторонних уклонений между эмпирическими кривыми достигается на p-шаге); искомая вероятность означает вероятность того, что блуждающая частица движется по одной из таких траекторий, у которой $x \geqslant c$.

Подсчитаем число траекторий, имеющих одной из „наиболее правых“ точку (r, p), $r \equiv p \,(\mathrm{mod}\ 2)$. Очевидно, что в точку (r, p) при этом условии можно попасть только выходя из точки $(r-1, p-1)$, а также из точки (r, p) частица могла двигаться только в точку $(r-1, p+1)$. Таким образом, число путей, удовлетворяющих требуемому условию, равно произведению числа путей из точки $(0, 0)$ в точку $(r-1, p-1)$, не пересекающих прямую $x = r+1$, на число путей из точки $(r-1, p+1)$ в точку $(0, 2n)$, также не пересекающих эту прямую, т. е. равно

$$\varphi_{\alpha_{p-1}}(r+1; 0, r-1)\, \varphi_{1-\alpha_{p+1}}(r+1; r-1, 0).$$

Таким образом, вероятность достижения максимума односторонних уклонений эмпирических кривых на p-шаге равна

$$P\{\omega_n(\alpha_i) \leqslant \omega_n(\alpha_p), \ i = 0, 1, \ldots, 2n\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r)} \varphi_{\alpha_{p-1}}(r+1; 0, r-1)\, \varphi_{1-\alpha_{p+1}}(r+1; r-1, 0).$$

Воспользовавшись известными соотношениями для биномиальных коэфициентов (см., например, Рыжик, Таблицы интегралов, стр. 252).

$$\sum_{(k)} C_m^{n-k} C_l^k = C_{m+l}^n, \tag{10}$$

$$\sum_{(k)} C_m^{n+k} C_l^k = C_{m+l}^{n+l}, \tag{11}$$

находим, что

$$P\{\omega_n(\alpha_i) \leqslant \omega_n(\alpha_p),\ \ i=0,\ldots,2n\} = \frac{1}{n+1},$$

откуда следует наше утверждение.

В частности, получаем известный результат [3]: вероятность нахождения одной эмпирической кривой под другой (т. е. вероятность достижения максимума односторонних уклонений двух эмпирических кривых на $2n$-шаге) равна $\dfrac{1}{n+1}$.

Теорема 10. В предположениях теоремы 9 для $0 < p < 2n$

$$P\left\{\omega_n(\alpha_p) \geqslant \frac{c}{n} \Big/ \omega_n(\alpha_i) < \omega_n(\alpha_p),\ \ i=0,\ldots,p-1,p+1,\ldots,2n\right\} =$$

$$= \frac{2(2n-1)}{C_{2n}^n} \sum_{\substack{r \geq c \\ (x \equiv p \,(\mathrm{mod}\ 2))}} \varphi_{\alpha_{p-1}}(r;\,0,\,r-1)\,\varphi_{1-\alpha_{p+1}}(r;\,r-1,\,0) =$$

$$= \begin{cases} 0, & \text{если } c > \varrho; \\ \dfrac{2(2n-1)}{p(2n-p)} \dfrac{1}{C_{2n}^n} \displaystyle\sum_{\substack{x \geq c \\ (x \equiv p\,(\mathrm{mod}\ 2))}} x^2 C_p^{\frac{p-x}{2}} C_{2n-p}^{n-\frac{p+x}{2}}, & \text{если } c \leqslant \varrho. \end{cases}$$

Доказательство этого утверждения отличается от доказательства теоремы 9 только тем, что теперь точки $(x,\ p)$ являются единственными наиболее правыми на соответствующих траекториях. Таким образом, число путей, проходящих через точку $(x,\ p)$ и удовлетворяющих требуемому условию, равно произведению числа путей, приводящих из точки $(0,\ 0)$ в точку $(r-1,\ p-1)$ и не пересекающих прямую $x=r$, на число путей, выходящих из точки $(r-1,\ p+1)$ и приводящих в точку $(0, 2n)$ и не пересекающих прямую $x=r$, т. е. равно

$$\varphi_{\alpha_{p-1}}(r;\,0,\,r-1)\,\varphi_{1-\alpha_{p+1}}(r;\,r-1,\,0).$$

Таким образом, вероятность достижения единственного максимума односторонних уклонений двух эмпирических кривых на p-шаге равна

$$P\{\omega_n(\alpha_i) < \omega_n(\alpha_p),\ i=0,\ldots,p-1,p+1,\ldots,2n\} =$$

$$= \frac{1}{C_{2n}^n} \sum_{(r)} \varphi_{\alpha_{p-1}}(r;\ 0,\ r-1)\,\varphi_{1-\alpha_{p+1}}(r;\ r-1,\,0).$$

Воспользовавшись соотношениями (10), (11), находим

$$P\{\omega_n(\alpha_i) < \omega_n(\alpha_p),\ i=0,\ldots,p-1,p+1,\ldots,2n\} = \frac{1}{2(2n-1)},$$

откуда следует наше утверждение.

Теорема 11. Из теорем 9 и 10 при $n \to \infty$ получаем

$$\lim_{n\to\infty} P\left\{\sqrt{\frac{n}{2}}\,\omega_n(\alpha_p) \geqslant z \Big/ \omega_n(\alpha_i) \leqslant \omega_n(\alpha_p),\ \ i=0,1,\ldots,2n\right\} =$$

$$= \lim_{n\to\infty} P\left\{\sqrt{\frac{n}{2}}\,\omega_n(\alpha_p) \geqslant z \Big/ \omega_n(\alpha_i) < \omega_n(\alpha_p),\ \ i=0,\ldots,p-1,p+1,\ldots,2n\right\} =$$

$$= f\left(\frac{z}{\alpha(1-\alpha)}\right).$$

где обозначено

$$f(z) = \frac{1}{\sqrt{2\pi}} \int\limits_{z}^{\infty} z^2 e^{-\frac{z^2}{2}} dz.$$

Последний результат в общем случае $n_1 \neq n_2$ содержится в работе [6].

Укажем еще на один результат, имеющий место при любом значении n.

Теорема 12. Вероятность того, что

$$D_n^+ = \omega_n(u_i)$$

при единственном значении i, равна $\frac{1}{2}$.

Это утверждение легко следует из теоремы 10. Действительно, как мы показали, вероятность достижения единственного максимума односторонних уклонений двух эмпирических кривых на p-шаге $(0 < p < 2n)$ не зависит от p и равна

$$\frac{1}{2(2n-1)},$$

причем p может принимать одно из $2n-1$ значений: $p = 1, 2, \ldots, 2n-1$, что и доказывает требуемое.

В заключение приношу глубокую благодарность проф. Б. В. Гнеденко за постановку задачи и ценные указания при ее решении.

ЛИТЕРАТУРА

1. Н. В. Смирнов, Бюлл. Моск. ун-та, (А), 2, в. 2, 1939.
2. Б. В. Гнеденко и В. С. Королюк, ДАН, 80, № 4, 1951.
3. Б. В. Гнеденко и Е. Л. Рвачева, ДАН, 82, № 4, 1952.
4. И. Д. Квит, ДАН, 81, № 2, 1950.
5. Х. Л. Берлянд и И. Д. Квит, ДАН УССР, № 1, 1952.
6. И. И. Гихман, ДАН, 82, № 6, 1952.

Получена 29.IV 1952 г.

Киев,
Институт математики АН УССР.