

*«То все ошибки, все накладки
И заблуждения веков»*
Автор неизвестен.

УДК 519.211; ББК 22.171

Критика теории вероятностей, часть V.

Числовые характеристики случайных величин

Бондарчук Игорь Иванович, кандидат физико-математических наук

12 декабря 2019 г.

Числовые характеристики, применяемые в теории вероятностей и математической статистике разнообразны, имеют различное значение и различные области применения. Из них мы рассмотрим в основном математическое ожидание, дисперсию и ковариацию, а также теоремы о математическом ожидании или дисперсии (суммы и произведения) случайных величин. Без упоминания о них не обходится ни одна работа по теории вероятностей, а в учебниках обязательно приводится доказательство теорем.

При создании *новой исходной системы* теории вероятностей [19,10-45] мы приводили *исходные понятия* теории событий *в согласии с экспериментами*. Принцип анализа в работе [29] *чисто математический*: построение *математических моделей* теории случайных величин *должно основываться исключительно на математических моделях* теории событий. Этот принцип анализа сохраняется и в настоящей статье.

{А.0} *Главная цель исследований* остается неизменной [19,1]
– *показать, какая есть и какой должна быть теория вероятностей случайных явлений, которые называют массовыми*¹.

Анализ показал: 1. Вывода биномиального распределения в имеющейся теории вероятностей не существует. 2. При неограниченном увеличении числа n испытаний к нормальному распределению сходится функция, полученная преобразованиями биномиального распределения, а не само биномиальное распределение. Утверждение справедливо только при равных вероятностях $p = q = 0,5$ событий. При конечных, но достаточно больших n это позволяет приближенно определять значения вероятностей биномиального распределения по нормальному закону. 3. Дисперсия двумерного распределения определяется суммой дисперсий одномерных величин X и Y : она не зависит от поворота системы координат. 4. Ковариация двумерного распределения, как в начальной (смешенный момент 2-го порядка), так и в центральной системе координат характеризует поворот главных осей относительно начальных или центральных осей координат. Она является характеристикой двумерного распределения и не определяет никаких характеристик одномерных случайных величин. 5. Ковариация и дисперсии одномерных распределений, вычисленных по двумерному распределению, взаимосвязаны между собой, но это не означает, что дисперсии определяются ковариацией и наоборот: правила вычисления ковариации и дисперсии разные. 6. Ковариация в начальных координатах (смешенный момент 2-го порядка) не является математическим ожиданием произведения случайных величин: она не имеет никакого отношения к их произведению. 7. Ковариация в центральных (или начальных) координатах не определяет функциональной зависимости случайных величин, а ее равенство нулю означает: направление главных осей совпадает с направлением центральных (или начальных), но никоим образом это не означает отсутствие функциональной зависимости между величинами (раздел 1.3 [30,28]).

¹Т.е. теории, занимающейся изучением только тех случайных событий, которые в принципе могут быть осуществлены неограниченное число раз при неизменных условиях проведения опыта, обладают статистической устойчивостью и в отношении которых имеет смысл не только утверждение об их случайности, но и возможна объективная оценка доли случаев их появления. Подробнее об этом можно почитать в [1,16]

Содержание.

1. Одномерное распределение.....	3
1.1. «Механическая интерпретация» числовых характеристик	3
1.2. Трактовка записи моментов в существующей теории	4
1.3. Применение моментов выше 2-го порядка	6
1.3.1. Какое же биномиальное распределение реально приближается к нормальному распределению? (6)	
1.3.2. Числовые характеристики и закон распределения (13)	
1.3.3. О трактовке математического ожидания и дисперсии (14)	
2. Двумерные (n -мерные) распределения	15
2.1. Математическое ожидание и дисперсия двумерного (n -мерного) распределения	16
2.2. Ковариация как характеристика поворота системы координат двумерного распределения: связь с дисперсиями	17
2.2.1. Отличие понятий дисперсии и ковариации (18).	
2.2.2. Поворот центральных координат на плоскости (20).	
2.2.3. Поворот начальных координат на плоскости (24).	
3. Линейное преобразование двумерного распределения в трехмерном пространстве	28
3.1. О применении нового вида преобразования	30
Список литературы	34

Примечания

1. Ссылки на работы даются в виде [№ в списке, № стр.]. Если работ несколько, то они разделяются знаком «;». При ссылках внутри работы указывается положение, на которое дается ссылка, и № страницы на которой оно находится.

2. Выводы, следующие из анализа, обозначаются номерами W.1, W.2, Чтобы обратить внимание Читателя на некоторые важные моменты, следующие из анализа, они выделены как положения и обозначены номерами {A.1}, {A.2},

3. Понятия, используемые в настоящее время, нумеруются заглавными буквами латинского алфавита. Уточненные предположения – римскими цифрами, а уточненные и вновь введенные определения понятий – словом «определение» с арабскими цифрами.

4. Нумерация формул: в круглых скобках – сквозная; в фигурных скобках – в пределах рассматриваемого примера.

«— Но ведь в природе так не бывает!
— Природа тут не причем.
Мое уравнение. Что хочу, то и пишу!»
Диалог (Физики шутят)

1. Одномерные распределения

Для характеристики *одномерной* случайной величины применяются [6; 8; 9; 11-18; 20]: *начальные* $\nu_k = \nu_k(\mathbf{X}') = \int_a^b (x')^k \cdot p(x) dx$ **{*}** и *центральные* $\mu_k = \mu_k(\mathbf{X}) = \int_a^b (x' - m_X)^k \cdot p(x) dx$ **{**}** моменты² *порядка* $k=1, 2, \dots$ (где $m_X = M(\mathbf{X}')$ – *математическое ожидание* случайной величины \mathbf{X}' ; \mathbf{X} – *центрированная* величина для которой математическое ожидание совмещено с началом координат). *Начальный момент 1-го порядка* называют *математическим ожиданием* $M(\mathbf{X}')$, а *центральный момент 2-го порядка* – *дисперсией* $D(\mathbf{X})$.

Замечания 1. **1.** Записи $\nu_k(\mathbf{X}')$, $\nu_k(\mathbf{X}'_n)$ (или $\mu_k(\mathbf{X})$, $\mu_k(\mathbf{X}_n)$) *условные*: они просто обозначают, *к какой величине* \mathbf{X} или \mathbf{X}_n ($n=1, 2, \dots$ – номер величины) относится характеристика: **Все!** Числовые характеристики определяются *выражениями в формулах справа от знака равенства*, в которых входят значения *вероятностной функции и их координаты*. **2.** Начальный момент 2-го порядка связан с дисперсией равенством $\nu_2(\mathbf{X}') - D(\mathbf{X}) = m_X^2$ **{1}**. В принципе, оно определяется самым простым преобразованием – *параллельным переносом системы координат вдоль оси X*: говорить здесь о «зависимости одинаковых случайных величин (положения {A.18}, {A.19} и замечания к ним [30, 42-43]) как-то язык не ворочается». В механике момент $\nu_2(\mathbf{X}')$ называется *осевым моментом инерции* относительно произвольной точки на оси X , в том числе начала координат: по аналогии будем называть его *начальной дисперсией* и обозначать $D_0(\mathbf{X}')$, подчеркивая этим общность их вычисления. В общем, их смысл тоже одинаков:

{A.1} В обоих случаях вычисляется *средневзвешенное отклонение*, только *относительно разных точек* – $D(\mathbf{X})$ *относительно математического ожидания*, а $D_0(\mathbf{X}')$ – *относительно произвольной точки* $x'_0 \neq m_X$ на оси X' .

Очевидно, что чем больше отклонение x'_0 от математического ожидания m_X , тем больше отличие между значениями $D_0(\mathbf{X}')$ и $D(\mathbf{X})$.

1.1. Механическая интерпретация числовых характеристик

При рассмотрении *математического ожидания* и *дисперсии* в теории вероятностей иногда проводится «механическая интерпретация» этих характеристик (например [6; 12; 13]): если по существу, то это *аналогия* между *распределением вероятностей* и *действительной функцией действительного аргумента*, которой является и *распределение масс*.

²Центральные потому, что *начало* координат находится в точке $x_H = M(\mathbf{X})$, а *начальные* – *начало* координат находится в *любой другой* точке $x'_H \neq M(\mathbf{X})$

Замечания 2. 1. (Механическая интерпретация). Если $\Phi_0(x)$ – *однозначная действительная* функция *действительной* переменной x , характеризующая распределение масс на отрезке $a \leq x \leq b$, а $\phi_0(x) > 0$ – ее производная (например, распределение масс), то для нее определяются те же моменты $\nu_k = \int_a^b (x_0)^k \cdot \phi(x_0) dx_0$, $\mu_k = \int_a^b (x)^k \cdot \phi(x) dx$, что и для случайной величины. Произведя *нормировку* $\Phi(x) = \Phi_0(x) / N$, где $N = \int_a^b \phi(x) dx$, получим функцию, которую можно трактовать как *закон* распределения случайной величины X .

2. *Чем же отличаются распределение масс и вероятностей?* Пусть функция $\phi_0(x) = m_0(x)$ определяет распределение масс на отрезке $a \leq x \leq b$. Тогда интеграл $M(x_1) = \int_a^{x_1} m_0(x) dx$ определяет точное значение массы $M(x_1)$ отрезка $a \leq x \leq x_1$. Теперь положим, что распределение масс является случайной величиной, а функция $\phi_0(x) = p_M(x)$ определяет *вероятностную* функцию величины M на этом отрезке. Тогда интеграл $P_M(x_1) = \int_a^{x_1} p_M(x) dx$ означает, что может появиться одно значение массы $M(x')$ из всех возможных значений масс $M(a \leq x' < x_1)$, принадлежащих интервалу $a < x' < x_1$, с вероятностью $P_M(a < x' \leq x_1)$. Более наглядное представление дает сравнение распределения «*точечных*» масс $m(x_k)$ ($k = 1, 2, \dots, K$) на оси X с *вероятностной* функцией $p(x_k)$ *дискретной* случайной величины.

1.2. Трактовка записи моментов в существующей теории

Обратим внимание на еще один момент: в существующей теории запись моментов порядка $k=2,3,\dots$ в виде $\nu_k = \nu_k(X)$ и $\mu_k = \mu_k(X')$ применяется редко. Обычно их записывают в виде $\nu_k = M(X^k)$, $\mu_k = M\{(X')^k\}$, т.е. как *математическое* ожидание *произведения*, которое определяется *умножением самой на себя k раз* случайной величины X . Хотя в некоторых работах и говорится, что это обозначение, но запись рассматриваются именно как (и в этих работах тоже) *произведение* (и суммы при обозначении $M(X + Y)$) случайных величин и воспринимается «как руководство к действию». Это следует из доказательства множества теорем о числовых характеристиках, в том числе для моментов высших порядков, которые основаны именно на такой записи. Мы не приводим этих «теорем», ибо придется выяснять, есть ли среди них правильные, а делать это нам «жутко как не хочется» (очень «муторная работа»).

Замечания 3. 1. Аналогично записывают и моменты многомерных распределений, например: «Начальным моментом порядка k, v системы $\{(X')^k(Y')^v\}$ называется математическое ожидание произведения $(X')^k$ на $(Y')^v$ » [13,175], т.е. имеем $\nu_{k,s}(X, Y) = \int_D (x')^k (y')^s \cdot p(x, y) dx dy$ и $\mu_{k,s}(X, Y) = \int_D x^k y^s \cdot p(x, y) dx dy$ {***}. С формулами {*-}{***}, определяющие моменты, все в полном порядке: исходя из механической аналогии, в правых частях формул подынтегральное выражение содержит произведение степеней координат на значение вероятностной функции, определяемые в одной точке. Т.е. вычисляется «средневзвешенное» значение степеней $(x')^k$, x^k или $(x')^k (y')^s$, $x^k y^s$ координат. Но то, что формальная запись $M\{(X')^k\}$ или $M\{(X')^k(Y')^s\}$ полагает-

ся «руководством к действию» (п.п.1-3 примера 16 [30,49]), вызывает недоумение.

2. В соответствии с принятым в существующей теории *отождествлением* случайных величин с *действительными* значениями координат рассмотрим две *дискретные* величины X' и Y' , принимающие на осях X и Y значения x'_v ($v = 1, 2, \dots, V$) и y'_w ($w = 1, 2, \dots, W$).

А. Согласно правилу II (приложение I [29,91]) умножения действительных функций, заданных на одной числовой оси, *произведение* $X'_1 X'_1$ величины X'_1 на саму себя определяет V *произведений* вида $x'_v \cdot x'_v$. Учитывая формулы, определяющие *дисперсии* в *начальных* и *центральных* координатах, получим выражения для дисперсий $D_0(X')$ и $D(X)$ выражения $M\{\sum_{v=1}^V (x'_v)^2\} \neq \{\sum_{w=1}^W (x'_v)^2 p(x'_v)\}$ и $M\{\sum_{v=1}^V (x'_v - m_X)^2\} \neq \{\sum_{v=1}^V (x'_v - m_X)^2 p(x_v)\}$ (с учетом равенства $D(X) = M(X^2)$, принятого в существующей теории).

Очевидно, что левые части выражений не соответствуют правым частям: в *левой части* имеем *сумму произведений* координат, а в *правой части* – *средневзвешенное значение произведений* соответствующих координат.

В. Пусть на осях X и Y случайные величины X и Y принимают значения x'_v ($v = 1, 2, \dots, V$) и y'_w ($w = 1, 2, \dots, W$). Согласно правилу IV (приложение I [29,91]) умножения действительных функций, заданных на числовых осях X и Y , *их произведение* определяет VW *произведений* вида $x'_v \cdot y'_w$, которые *представимы в виде двумерной* таблицы.

Центральные моменты 2-го порядка, т.е. *дисперсия* и *ковариация* в *центральных* координатах, вычисляются по *правым* частям формул $\{*\}$ - $\{***\}$. Учитывая принятые в существующей в теории равенства $D(X) = M(X^2)$ и $K(XY) = M(XY)$, получим для дисперсий и ковариации выражения $M\{\sum_{v=1}^V (x'_v - m_X)^2\} \neq \{\sum_{v=1}^V (x'_v - m_X)^2 p(x_v)\}$, $M\{\sum_{w=1}^W (y'_w - m_Y)^2\} \neq \{\sum_{w=1}^W (y'_w - m_Y)^2 p(y_w)\}$ и $K\{\sum_{v=1}^V \sum_{w=1}^W (x'_v - m_X)(y'_w - m_Y)\} \neq \{\sum_{v=1}^V \sum_{w=1}^W (x'_v - m_X)(y'_w - m_Y) p(x_k, y_w)\}$.

Подробный анализ равенства $D(X) = K(XY)$ будет дан при анализе того, что реально определяет ковариация двумерного распределения. Здесь оно приведено для того, чтобы показать: в существующей теории вероятностей не учитывается, что произведение действительных функций, заданных на одной оси, и произведение функций, заданных на 2-х (и более) осях, вычисляются по разным правилам (приложение I, [29,91]).

И в этом случае *левые* части выражений *не соответствуют правым* частям. В *левой части* имеем *сумму произведений* координат, а в *правой части* – *сумму средневзвешенных значений произведений* координат.

Таким образом, из анализа следует:

{А.2} Запись моментов в виде *степеней* $M(X^k)$ случайной величины X (или *произведений* $M(X^k Y^n)$ *степеней* случайных величин X и Y) *не соответствует* формулам, определяющим моменты даже в рамках существующей теории. Ее *следует считать* просто как *обозначение моментов* соответствующего порядка, но не более.

Восприятие записи как действий над случайными величинами (п.1.2 примера 14 [30,41]) определяет в последующем неверные выводы. Если бы случайная величина и неразрывно связанная с ней вероятностная функ-

ция были отделены друг от друга, то они никогда не появились бы. Однако это рассуждения рода «если бы, да кабы ...».

1.3. Применение моментов выше 2-го порядка

В существующей теории, хотя и определяются моменты *выше* 2-го порядка (для одномерных распределений), но говорится в основном о применении только 2-х *центральных* моментов [12,93; 13,98]. Момент 3-го *порядка*, который представляется в виде $\Sigma_k = \mu_3/\sigma^3$ (где $\sigma = \sqrt{D(\mathbf{X})}$ – *среднеквадратическое отклонение*) и называется *коэффициентом асимметрии*; 4-го *порядка*, который представляется в виде $\varepsilon_k = \mu_4/\sigma^4 - 3$ и называется *эксцессом* (для нормального распределения отношение $\mu_4/\sigma^4 = 3$, т.е. $\varepsilon_k = 0$). По существу, *произвольные законы* распределения сравниваются с «эталонном» – *нормальным распределением*.

Сравнивать некоторое *произвольное* распределение с *нормальным* распределением, конечно же, можно: ничего «противозаконного» в этом нет.

Однако, мы затрудняемся сказать, что дает такое сравнение по двум причинам:

Во-первых, в некоторых случаях это приводит к противоречивым выводам. Во-вторых, *сравнение не позволяет определить влияние центральных моментов на вид распределения: даже математического ожидания и дисперсии* (вообще-то, в первую очередь), не говоря о моментах *более высокого порядка*.

Рассмотрим каждую из причин отдельно.

1.3.1. Какое же биномиальное распределение приближается к нормальному распределению?

Начнем с двух цитат:

В [12,100] определены выражения для *асимметрии* $\Sigma_n = (p - q)/\sqrt{npq}$ и *эксцесса* $\varepsilon_n = (1 - 6p + 6p^2)/(npq)$ *биномиального* распределения, а далее утверждается: «С ростом n ε_n и Σ_n , очевидно, стремятся к нулю. Это, как мы увидим далее, связано с приближением биномиального распределения к симметричному нормальному закону» [12,100]. Однако, далее следует [12,186]: «Мы видим, что теорема (5.3.8) Я. Бернулли утверждает, что среднее арифметическое большого числа независимых величин (частного вида – «двузначных») почти наверное будет как угодно близко к своему математическому ожиданию – постоянной величине p ». Чтобы было, о чем поразмыслить, для значений $p = q = 0,5$, $p = 0,1$; $q = 0,9$ и $p = 0,9$; $q = 0,1$ приведем значения: математических ожиданий $m_1 = 0,5n$, $m_2 = 0,1n$ и $m_3 = 0,9n$; дисперсий $D_1 = 0,25n$ и $D_2 = D_3 = 0,09n$.

Их отношения постоянны, равны значениям $m_1/m_2 = 5$, $m_1/m_3 = 5/9$ и $D_1/D_2 = 25/9$ и не зависят от числа n испытаний.

Как видно, утверждение во 2-й цитате *противоречит* утверждению в 1-й цитате: оно состоит в том, что распределение *одной вероятностной функции не может иметь разные математические ожидания* (например: $p = 0,5$ и $p = 0,1$ или $p = 0,9$). Зато из противоречия косвенно следует:

{А.3} *Биномиальное распределение $P_n(\mu_m = m) = C_n^m \cdot p^m \cdot q^{n-m}$ (где $C_n^m = n!/[m!(n-m)!]$) приближается к нормальному распределению при значении $p = 1/2$.*

Это показано в [10,82-86]: утверждение подтверждается существованием распределения Пуассона, которое также является аппроксимацией биномиального распределения. *А каким будет распределение между этими двумя – $p = 1/2$ и p близких к 0 (или к 1) – «крайними» случаями?*

Мы вряд ли сможем дать ответ на вопрос, поставленный выше, – для этого необходим выдающийся талант А. Муавра, С. Пуассона (или еще кого-то), – но с имеющимися противоречиями попробуем разобраться.

Пример 1. п.1. При значениях $p = q = 0,5$ и $p = 0,9, q = 0,1$ значения *асимметрии и эксцесса* равны: $\Sigma_k = 0, \varepsilon_k = -1/n$ и $\Sigma_k = -2,67/\sqrt{n}, \varepsilon_k = 5,11/n$. Т.е. в обоих случаях с *ростом n* значения Σ_k и ε_k *стремятся к нулю*, следовательно, по утверждению в [12], *биномиальные распределения приближаются к симметричному нормальному закону*. Математические ожидания $M_1 = 0,5n, M_2 = 0,9n$ и *расстояние между ними $M_2 - M_1 = 0,4n$ с ростом n стремятся к бесконечности*. Однако *отношения $M_1/M_2 = 5/9$ и $(M_2 - M_1)/M_2 = 4/9$ постоянны, не зависят от числа n и равны значениям исходных вероятностей*, а это, в принципе, означает: *сумма значений вероятностных функций в точках $m = 0, 1, 2, \dots, M_1$ и $m' = 0, 1, 2, \dots, M_2$ равны значениям $P(0 \leq x_m \leq M_1) = \sum_{m=0}^{M_1} p_n(x_m = m) = 0,5$ и $P(0 \leq x_m \leq M_2) = \sum_{m=0}^{M_2} p_n(x_m = m) = 0,9$ в 1-ом и во 2-м случаях соответственно*.

п.2. «На рис.10.а ординаты изображают значения вероятностей $P_n(m)$ при различных значениях абсциссы m . По рисунку видно, что с увеличением n величины $P_n(m)$ равномерно убывают. Для того чтобы на рисунке точки $\{m, P_n(m)\}$ уже для рассматриваемых значений n не слились с осью абсцисс, мы выберем резко различные масштабы по осям координат» [1,84]. Далее идет пояснение, что означает замена переменных m и $P_n(m)$ переменными $x_n = (m - np)/\sqrt{npq}$ и $y_n(m) = P_n(m) \cdot \sqrt{npq}$. Наиболее подробное пояснение дано в [15]: оно компактно изложено в [10,85]. Отметим следующее:

I. *Вероятностная функция $p_n(m)$ преобразованием $m' = m - np$ приводится к центральной системе координат: в конечном счете, это определяет симметрию конечного распределения.* II. *Расстояния между значениями $p_n(m)$ и $p_n(m+1)$ вероятностной функции преобразованием $m'' = m'/k_n$ уменьшаются в $k_n = \sqrt{npq}$ раз: в пределе (при значении $n \rightarrow \infty$) – это определяет непрерывность распределения вероятностей.* III. Действительно, «...с увеличением n значения $p_n(m)$ равномерно убывают», а в пределе – *обращаются в ноль, образуя на конечном отрезке равновероятное распределение* (выводы Q.1, Q.2 в приложении II [19,50]). Поэтому, значения вероятностной функции $p_n(m)$ увеличиваются в $k_n = \sqrt{npq}$ раз: в пределе это обеспечивает сходимость значений $y_n(m)$ к необходимому пределу, отличному от нуля.

Для любого конечного $n > (pq)^{-1}$ *сумма $\sum_{m=1}^n y_n(m) = \sqrt{npq} \sum_{m=1}^n p_n(m)$ всех значений функции $y_n(m)$ больше 1 и только в пределе равна 1. Т.е. выражение $y_n(m) =$*

$P_n(m) \cdot \sqrt{npq}$ не является вероятностной функцией какого-либо распределения.

Если об *обеспечении непрерывности* говорится в работах (из цитируемых – подробно только в [15,48-49]), то на 1-й и 3-й моменты – *обеспечение симметрии и сходимости к пределу, отличному от нуля* – даже намеков нет.

{А.4} Таким образом, из пунктов I-III, следует, что в цитате *не совсем точно* определена *цель замены переменных: она состоит* в том, чтобы *обеспечить непрерывность и сходимость* распределения при предельном переходе.

А совсем не в том, «... чтобы на рисунке точки $\{m, P_n(m)\}$ уже для рассматриваемых значений n не слились с осью абсцисс ...». Рисунки 10 [1,83] наглядно иллюстрируют сказанное.

Можно было бы на это не обращать внимания, если бы в работах не *утверждалось*, что *биномиальное* распределение при значениях $n \rightarrow \infty$ *сходится к нормальному* распределению для любых p ($0 \leq p \leq 1$). Мы не знаем, когда появилось это утверждение, но, по крайней мере, у Муавра, судя по названию работы, его нет. Приведем еще несколько цитат:

1. «В дальнейшем мы увидим, что при больших значениях n все вероятности $P_n(m)$ становятся близкими к нулю, но только для m , близких к вероятнейшему значению m_0 , вероятности $P_n(m)$ сколько-нибудь заметно отличаются от нуля. Этот факт впоследствии будет доказан нами ...» [1,77]. 2. «Мы вернемся теперь к биномиальному распределению и покажем, что при больших значениях n оно с хорошим приближением (тем точнее, чем больше n) может быть описано с помощью нормального распределения с тем же центром и той же дисперсией, что и у биномиального распределения» [12,145]. 3. «При качественной оценке применимости приближенной формулы $P_n(\mu_n = m) \approx e^{-x_m^2} / \sqrt{2\pi \cdot npq}$ (3.7) нужно оценить величину остаточных членов. При $n \rightarrow \infty$ сумма остаточных членов стремится к 0 при любых фиксированных p ($0 < p < 1$). Однако, при конечных значениях n сумма остаточных членов может оказаться очень большой, если p или q малы. Хорошие приближения формула дает при $p = q = 0,5$ Формулу часто используют при $n > 100$ и $npq > 20$. Указания о границах применения формул³ (3.2) и (3.7) являются очень приближенными и носят скорее качественный характер; к ним следует относиться с большой осторожностью» [16,69]. 4. «Приближенная формула $P_n\{a \leq (\mu_n - np) / \sqrt{npq} \leq b\} \approx 1 / \sqrt{2\pi} \cdot \int_a^b e^{-x_m^2/2} dx$ (3.12) используется в тех случаях, когда возможно использование (3.7)» [16,71]. 5. В работе [12,150] приводится формула, которая «... дает несколько более точное приближение, чем (4.3.11) (аналог формулы (3.12) в работе [16,71], *примечание автора*), особенно при не слишком больших n ».

То, что с утверждением в 1-й цитате что-то не так, следует непосредственно из примеров, приведенных в этой же работе [1,81-83]:

В табл.8-11 и на рис.10 приведены значения m , $x_n = (m - np) / \sqrt{npq}$, $P_n(m)$, $y_n(m) = P_n(m) \sqrt{npq}$ и $\phi(x) = 1 / \sqrt{2\pi} \cdot e^{-x^2/2}$ для значений $p = 0,2$ и $n = 4, 25, 100, 400$. Из анализа данных, например, для максимальных значений $P_n(m)$, $y_n(m)$ и $\phi(x)$: $n = 4 \rightarrow 0.4096, 0.3277$ и 0.3867 ; $n = 25 \rightarrow 0.1960, 0.3920$ и 0.3989 ; $n = 100 \rightarrow 0.0993, 0.3972$ и 0.3989 ; $n = 400 \rightarrow 0.0498, 0.3985$ и 0.3989 следует:

³Что касается формулы (3.2), то речь идет о применимости распределения Пуассона $P_n(\mu_n = m) \approx \lambda_n^m / m! \cdot e^{-\lambda_n}$, $\lambda_n = np$ [16,67]

{A.5} Значения $P_n(m')$ для фиксированного значения x'_n (или $m' = np + x'_n\sqrt{npq}$) уменьшаются при увеличении значения n , приближаются к нулю и удаляются от значения функции $\phi(x'_n)$.

{A.6} Значения функций $y_n(m')$ и $\phi(x'_n)$ для фиксированного значения x'_n (или $m' = np + x'_n\sqrt{npq}$) при увеличении значения n увеличиваются и сближаются между собой.

{A.7} Функции $P_n(m)$ и $y_n(m)$ *несимметричны* относительно *математического ожидания* *np* *биномиального* распределения (в отличие от функции $\phi(x)$ плотности нормального распределения): при этом для *любого конечного n* отношения $M(X)/n = |np/n| = 0,2$ и $\{x_K - M(X)\}/n = |(n - np)/n| = 0,8$ *постоянны и не зависят* от значения n .

Положение **{A.7}** ставит под сомнение и утверждение, данное во второй цитате: сходимости биномиального распределения к нормальному при любых значениях p ($0 \leq p < 1$). Положения **{A.5}**-**{A.7}** получены на основе анализа ограниченных данных при ограниченном числе испытаний: чтобы хоть как-то «узаконить» их, требуются более весомые аргументы. Чтобы показать правильность положений, обратимся непосредственно к доказательству локальной теоремы Муавра-Лапласа.

Замечания 4. 1. После преобразований, отмеченных в п.п. I-III (стр.7), факториалы в биномиальных коэффициентах C_n^m вычисляются по приближенной формуле Стирлинга $n! = \sqrt{2\pi n} e^{-n} (1 + \alpha_n)$: полученную функцию записывают в виде произведения [15,48] $\Pi_n(x_n) = S(x_n)R(x_n)T(x_n)/\sqrt{2\pi}$, где $S(x_n) = \sqrt{npq} \cdot n / \{m(n-m)\}$, $R(x_n) = (np/m)^m \{nq/(n-m)\}^{n-m}$ и $T(x_n) = (1 + \alpha_n) / \{(1 + \alpha_m)(1 + \alpha_{n-m})\}$.

Значения m и $n - m$ заменяются значениями $m = np + x_n\sqrt{npq}$ и $n - m = nq - x_n\sqrt{npq}$.

Мы ограничимся рассмотрением функций $S(x_n)$ и $R(x_n)$: функция $T(x_n)$ при значении $n \rightarrow \infty$ стремится к единице и не влияет на результаты дальнейшего анализа.

2. Функция $S(x_n)$ имеет вид $S(x_n) = 1 / \sqrt{1 + x_n(q-p)/\sqrt{npq} - x_n^2/n}$. При значении $n \rightarrow \infty$ она стремится к единице. Однако если преобразование III не выполнить, получим $S(x_n) = 1 / \{\sqrt{npq} \cdot \sqrt{1 + x_n(q-p)/\sqrt{npq} - x_n^2/n}\}$, т.е. в знаменателе появляется множитель \sqrt{npq} : это приводит к тому, что при значении $n \rightarrow \infty$ функция $S(x_n)$ стремится к нулю. Следовательно, функция $\Pi_n(x_n)$ также стремится к нулю.

Тем самым показана правильность положений **{A.5}**-**{A.6}**.

3. Функция $R(x_n)$ принимает вид $R(x_n) = \{1 + x_n\sqrt{q/(np)}\}^{-m} \{1 - x_n\sqrt{p/(nq)}\}^{-(n-m)}$. Вычисляя логарифм функции $R(x_n)$ по приближенной формуле $\ln(1+z) = z - z^2[1 + \omega(z)]$, получим $\ln R(x_n) = -m\{x_n\sqrt{q/(np)} - x_n^2(1 + \omega')q/(np)\} - (n-m)\{x_n\sqrt{p/(nq)} - x_n^2(1 + \omega'')p/(nq)\}$, где $\omega(z) \rightarrow 0$ при значении $n \rightarrow \infty$, $\omega' = \omega[\sqrt{q/(np)}]$ и $\omega'' = \omega[\sqrt{p/(nq)}]$. Заменяя значения m и $n-m$, получим $\ln R(x_n) = -x_n^2/2 + x_n^2(q\omega' + p\omega'')/2 + x_n^3\{(1 + \omega')q\sqrt{q/p} + (1 + \omega'')p\sqrt{p/n}\}/(2\sqrt{n})$ (с учетом равенства $p+q=1$). Далее доказывается, что при значении $n \rightarrow \infty$ функция $R(x_n)$ стремится к функции $e^{-x^2/2}$ а функция $\Pi_n(x_n)$ к функции плотности $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$, нормального распределения.

Обратим внимание на то, что при значении $n \rightarrow \infty$ слагаемые $x_n^2(q\omega' + p\omega'')/2$ и $x_n^3\{(1 + \omega')q\sqrt{q/p} + (1 + \omega'')p\sqrt{p/n}\}/(2\sqrt{n})$, содержащие исходные вероятности p и q стремятся к нулю. А это означает:

{A.8} Отличие значений вероятностей p и q не влияет на значение предела функции $R(x_n)$. Это является следствием преобразования I и определяет симметрию функции $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$.

Полагая $p = q = 0,5$, получим $m = (n + x_n\sqrt{n})/2$, $n - m = (n - x_n\sqrt{n})/2$, $S_0(x_n) = 1/\sqrt{1 - x_n^2/n}$ и $R_0(x_n) = \{1 + x_n/\sqrt{n}\}^{-m}\{1 - x_n/\sqrt{n}\}^{-(n-m)}$. Функция $T(x_n)$ не изменяется, функции ω' и ω'' становятся одинаковыми $\omega' = \omega'' = \omega(1/\sqrt{n})$, а логарифм функции $R_0(x_n)$ равен $\ln R(x_n) = -x_n^2/2 + x_n^2\omega/2 + x_n^3(1+\omega)/(2\sqrt{n})$. Доказательство сходимости функции $\Pi_n^0(x_n)$ к нормальному распределению проводится точно также как и для функции $\Pi_n(x_n)$. В данном случае в ноль обращаются слагаемые $x_n^2\omega/2$ и $x_n^3(1+\omega)/(2\sqrt{n})$.

Результат тот же, что и полученный выше: при значении $n \rightarrow \infty$ функция $S_0(x_n)$ стремится к единице, а функция $R_0(x_n)$ — к функции $e^{-x^2/2}$. Следовательно, функция стремится к функции $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$.

Таким образом, в замечаниях 4 показано:

{A.9} Именно функция $y_n(m') = p_n(m')\sqrt{npq}$ (где $m' = np + x'_n\sqrt{npq}$), полученная преобразованиями I-III биномиального распределения, сходится к нормальному при неограниченном увеличении числа n испытаний, а не само биномиальное распределение.

{A.10} Анализ (замечания 4.3-4.4) подтверждает положение A.3: преобразованное биномиальное распределение сходится к функции $\phi(x) = e^{-x^2/2}/\sqrt{2\pi}$ при значении вероятности $p = 0,5$.

{A.11} Функции $y_n(m')$ и $\phi(x)$ позволяют вычислить приближенные значения $p_n(m = np + x_n\sqrt{npq}) \approx e^{-x_n^2/2}/\sqrt{2\pi \cdot npq}$ (цитата 3, стр.8) вероятностной функции биномиального распределения при данном значении n .

Выражение $e^{-x_n^2/2}/\sqrt{2\pi \cdot npq}$ в приближенном равенстве при значении $n \rightarrow \infty$ стремится к нулю, а, следовательно, и значения $p_n(m)$ вероятностной функции биномиального распределения также стремятся к нулю. Тем самым лишний раз подтверждаются выводы S6.1-S6.5 [10,106-107] и Q.1-Q.2 приложения II [19,48].

{A.12} На ограниченном отрезке получаем бесконечное (несчетное) множество элементарных событий с нулевыми вероятностями. На равных частях отрезка они определяют равные вероятности, т.е. равновероятное непрерывное распределение.

{A.13} Возможность вычисления значений $p_n(m)$ по функции $e^{-x_n^2/2}/\sqrt{2\pi \cdot npq}$ не определяет сходимости биномиального распределения к нормальному: просто обе функции обращаются в ноль, а при представлении на ограниченном отрезке — сходятся к равновероятному распределению.

Рассмотрим, чем отличается вывод биномиального распределения в существующей теории от его вывода в теории, развитой в настоящей работе. В работе [29,11-14] показано:

{A.14} В *существующей теории* дается только вывод формулы Бернулли $P_n(m) = C_n^m \cdot p^m \cdot q^{n-m}$, который приводится в теории событий, а вот вывода *биномиального* распределения в ней не существует.

По нашему мнению, наиболее понятный и подробный вывод формулы Бернулли, основанный на классической теории, дан в работе [13,59-66].

«Переход» к *биномиальному* распределению весьма прост:

«Обозначим через μ *число появлений* события А в последовательности независимых испытаний, в каждом из которых вероятность его появления постоянна и равна $P(A) = p$ Согласно результатам главы 2 (в этой главе и дается вывод формулы Бернулли) $P_N(m) = P_N(\mu = m) = C_N^m \cdot p^m \cdot q^{N-m}$ » [1,117]. Иногда формулу сразу записывают в виде $p_N(x) = C_N^x \cdot p^x \cdot q^{N-x}$ и называют биномиальным распределением [12,65].

Приведем для интереса комментарии [29,13] к *выводу биномиального* в существующей теории, появившиеся после анализа вывода.

«Комментарии II. 1. Используя понимание *элементарного* события как «неразложимого исхода опыта» [6,26] и *классическую* формулу, вычислены вероятности *сложных* событий А и В. 2. Далее *комбинации* символов (тройки в цитате), считаются *элементарными* событиями. Только о том, что они являются *сложными* событиями, определяемые *произведениями* опять же *сложных* событий А и В, почему-то «ни слова»: – именно по их вероятностям (замените «1» буквой «А», а «0» буквой «В») вычисляются вероятности *произведений*. *Произведения* «объявили» *элементарными* событиями и для них «введена вероятностная мера», следовательно, в соответствии с определениями 'В, В.1' [29,9-10], на них и следует определить случайную величину. 3. Однако это как раз и не делается, а определяются еще более *сложные* события, образованные *суммами произведений*. При переходе к распределению вероятностей их относят к *точкам* числовой оси с координатами $x_m = 0, 1, \dots, N$, которые, вообще говоря, определяют *пространство элементарных* событий (замечание к определению 'В.1'). Т.е. эти *суммы*, по существу, считаются *элементарными* событиями: – только почему-то «по умолчанию». Об этом и говорилось в примере 14 [19,25]».

В соответствии с теорией, развитой в настоящей работе, формула Бернулли $P(A_{n,m}) = C_n^m \cdot p^m \cdot q^{n-m}$ определяет [10,85] вероятности *сложных* событий $A_{n,m} = C_n^m \cdot \prod_{v=1}^m a_1^v \cdot \prod_{k=1}^{n-m} a_2^k$ и не является распределением случайной величины.

Вероятности, соответствующее формуле Бернулли, можно получить исходя из того, что сложные события $A_{n,m}$ *несовместимы*, а *сумма* всех событий $A_{n,m}$ (для данного n) является достоверным событием, т.е. *сумма* их вероятностей равна 1. Поставив в соответствие событиям $A_{n,m}$ действительные числа x_m , такие что $x_{m+1} - x_m = \Delta = \text{const}$. Получим случайную величину с *вероятностной* функцией $p_{n,m}(x_m) = C_n^m \cdot p^m(x_m) \cdot q^{n-m}(x_m)$ и характеристиками: $M(X) = n p \Delta$, $D(X) = n p q \Delta$. Числа x_m не влияют на вероятности событий. Их разность Δ определяет расстояния между 2-мя последовательными значениями вероятностей. Если $\Delta = 1$, то получим значения вероятностей в точках $x = m = 0, 1, \dots, n$. Отметим, что запись $p_n(m)$, используемая в работах, неточна.

Теперь наглядно покажем, что *биномиальное* распределение при выполнении преобразований I-III (стр.7) *сходится к нормальному* только при вероятности $p = 0,5$ (положение {A.3}, стр.7).

Пример 2. Рассмотрим систему, образованную совмещением N одинаковых дискретных случайных величин X^v (индекс $v = 1, 2, \dots, N$ – номер величины), определенных на отрезках $0 \leq r^v \leq 1$, с вероятностными функциями $p(x_1^v = 0) = P(a_1^Y) = q$, $p(x_1^v = 1) = P(a_2^Y) = p$ ($p + q = 1$) и числовыми характеристиками $M(X^v) = p$, $D(X^v) = pq$. Т.е. опыт с одной и той же случайной величиной повторяется N раз. В результате получаем N -мерное распределение произведения N одинаковых вероятностных функций (раздел 4.1.2 [29,74]). Применяя 1-й вариант преобразования N -мерного (с линейной формулой преобразования) распределения к одномерному распределению (пример 12 [30,38]), получим биномиальное распределение, принадлежащее отрезку $0 \leq x \leq N$.

Используем то обстоятельство, что функция $P = P(x)$ закона распределения одновременно является монотонно возрастающей формулой преобразования: она отображает множество элементарных событий, заданных на отрезке $a \leq x \leq b$ оси X , в множество вероятностей $P(a_m \rightarrow r_m)$ на отрезке $0 \leq r \leq 1$ оси Y (п.4 примера 3 [30,13]).

Учитывая это, будем рассматривать распределение в координатах $r0p$ (рис.1). Для того, чтобы распределение принадлежало отрезку $0 \leq r \leq 1$, следует применить линейное преобразование $r = x/N$: получим биномиальное распределение с вероятностной функцией $p_{n,m}(r_m) = C_n^m \cdot p^m(r_m) \cdot q^{n-m}(r_m)$ и числовыми характеристиками $M(X) = p$, $D(X) = pq/N$ [10,85].

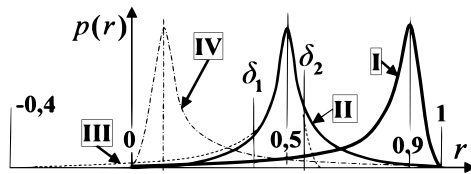


Рис. 1.

на отрезок $0 \leq r \leq 1$ оси Y , учитывая, что математическое ожидание $M(X) = p$ в координатах $r0p$ не изменяется, будет иметь вид, изображенный на рис.1 (сплошная кривая I). Для сравнения в тех же координатах изображена функция плотности нормального распределения (сплошная кривая II). Штрихпунктирная кривая IV изображает биномиальное распределение при значении $p = 0,1$, которая является зеркальным отражением кривой I относительно плоскости, перпендикулярной плоскости рисунка и содержащую линию $p = 0,5$.

Сместим кривую I вдоль оси r так, чтобы ее максимум совпал с максимумом (штриховая кривая III) кривой II: этот прием обычно применяется в существующей теории для иллюстрации характера приближений, которые дает локальная теорема Муавра-Лапласа. В этом случае математические ожидания распределений также совпадают⁴. Функция $y_N(m) = p_N(m)\sqrt{Npq}$ определена на отрезке $-0,4 \leq r \leq 0,6$.

Из рисунка видно: 1. При значении $r = 0,5$ вероятность, определяемая нормальным распределением равна 0,5, а определяемая биномиальным распределением – 0,9. 2. При значении $r = 0,6$, вероятность, определяемая биномиальным распределением равна 1, а при нормальном распределении это значение будет достигнуто только при значении $r = 1$. 3. Относительное отклонение Δ точных значений биномиального распределения $p_n(m) = p_n(m = np + x_n\sqrt{npq})$ от нормального приближения $e^{-x_n^2/2} / \sqrt{2\pi \cdot npq}$ при любом конечном n определяется формулой $\Delta = 1 - e^{-x_n^2/2} / \{\sqrt{2\pi \cdot npq} \cdot p_n(m)\}$ {1}. Требуемые значения $|\Delta| < \varepsilon > 0$ возможных отклонений определяются (рис.1) несимметричным относительно математического ожидания интервалом $\delta_1 < r < \delta_2$ (рис.1). Это также следует из положения {A.6} (стр.9).

⁴Отметим, что максимум функции смещен относительно математического ожидания, однако смещение при значении $p = 0,9$ невелико и при увеличении числа опытов несколько уменьшается. Мы считаем их одинаковыми, что не влияет на рассуждения

Эти отношения *сохраняются* при *любом сколь угодно большом* значении n , в том числе при значении $n \rightarrow \infty$. Отсюда следует:

{А.15} *Биномиальный закон распределения с математическим ожиданием $p = 0,9$ при значениях $n \rightarrow \infty$ стремится к некоторой неизвестной функции, которая никоим образом не может превратиться в нормальный закон.*

1. Очевидно, что аппроксимация биномиального распределения при математическом ожидании $M(X) = 0,9$ распределением Пуассона даст *лучшие* результаты, чем аппроксимация нормальным распределением. 2. Отметим, что при выводе распределения Пуассона, используется единственное, с точки зрения математики вполне естественное, предположение: значение вероятности $q = 1 - p$ близко к нулю.

Кстати, при анализе построения случайных процессов (часть VII исследований) показано: на основе некоторого случайного процесса определяется еще одно биномиальное распределение, в котором вместо произведений вероятностей стоят их суммы, т.е. оно существенно отличается биномиального распределения в существующей трактовке: именно оно сходится к нормальному распределению без использования каких-либо упрощающих предположений.

Значение математического ожидания $p = 0,9$ было принято нами только для примера: приведенные выше рассуждения можно повторить для любого значения $p \neq 0,5$ ($0 < p < 1$). А это означает, что все свойства кривой I будут, в той или иной мере, присущи всем биномиальным распределениям с математическими ожиданиями $M(X) \neq 0,5$.

При значении $M(X) = 0,5$ биномиальное распределение симметрично относительно математического ожидания. Требуемые значения относительных отклонений $|\Delta| = |1 - e^{-x_n^2/2} / \{\sqrt{2\pi \cdot npq} \cdot p_n(m)\}| < \varepsilon$ точного значения биномиального распределения $p_n(m) = p_n(m = np + x_n \sqrt{npq})$ от нормального приближения $e^{-x_n^2/2} / \sqrt{2\pi \cdot npq}$ определяется симметричным интервалом $\delta_1 < r < \delta_2$.

Из формулы $\{(1)\}$ следует:

{А.16} *При значении $M(X) = 0,5$ точность вычисления значений $p_n(m)$ наибольшая, ибо на точность влияет только число n испытаний.*

Даже при малом отклонении математического ожидания от значения $p = 0,5$ симметрия распределения нарушается. На точность, кроме числа n испытаний, начинает влиять и несимметричность биномиального распределения. Чем больше отклонение биномиального распределения от симметрии, тем меньше точность. По сути, рассуждения о точности в цитатах 3-5 (стр.8) косвенно подтверждают это.

1.3.2. Числовые характеристики и закон распределения

При разработке новой исходной системы теории случайных величин показано (вывод W.18 [29,52]), что область определения закона распределе-

ния является *ограниченной*, что позволило предложить подход [10,77-82], который позволяет провести исследование *влияние центральных* моментов на *вид* распределения. Он состоит в следующем:

Замечание 5. Вычисление *начальных* моментов по *центральному* моментам⁵ определяется формулой $\nu_k = \mu_k + \sum_{j=1}^k C_k^j m_x^j \mu_{k-j} + m_X^k \{*\}$ (где C_k^j – число сочетаний). Вычислив момент этого вида для *произвольной* точки $x = \xi$ области *определения* $a = x_H \leq \xi \leq x_K = b$ *вероятностной* функции, получим $\beta_k(\xi) = \int_0^b (x - \xi)^k p(x) dx \{**\}$. Раскрывая скобки и учитывая формулу $\{*\}$, получим *для момента* выражение: $\beta_k(\xi) = \mu_k + (m_X - \xi) \sum_{j=1}^{k-2} C_k^j \mu_{k-j} (m_X - \xi)^{j-1} + m_X - \xi)^k$ ($k=2,3,\dots$). (1)

Выражение (1) определяет *общую зависимость моментов* $\beta_k(\xi)$ от переменной ξ в области *определения* $a = x_H \leq \xi \leq x_K = b$ *вероятностной* функции. Удобно применять также функцию $d_k(\xi) = \beta_k(\xi) - \mu_k$. (2)

При значениях $x = a$ и $x = m_X$ получим значения *начального* и *центрального* моментов. *Производная* от функции (1) равна $\beta'_k(\xi) = -k\beta_{k-1}(\xi)$. Легко показать, что *при переходе* через точку $x = \xi = m_X$ *производная от моментов четного* порядка $k=2j$ ($j=1,2,\dots$) *меняют знак*, а от *моментов нечетного* порядка $k=2j+1$ *знака не меняют*.

Отсюда следует:

I. Функции $\beta_k(x_1)$, $d_k(x)$ *четного* порядка $k=2j$ ($j=1,2,\dots$) – *четные относительно точки* $x = \xi = m_X$ и имеют в этой точке *экстремум*.

II. Функции $\beta_k(x_1)$, $d_m(x)$ *нечетного* порядка $k=2j+1$ ($j=1,2,\dots$) – *нечетные относительно точки* $x = \xi = m_X$, т.е. она является *точкой перегиба*.

С применением функций (1-2), в [10,77-82] проведен анализ *влияния центральных* моментов на функцию плотности. Показано:

W.1. *Положение математического ожидания относительно границ области определения определяет форму функции плотности, а дисперсия – масштаб формы. Центральные моменты более высокого порядка уточняют: нечетного порядка – форму, а четного порядка – масштаб функции плотности.*

Отметим то, о чем в [10] не говорилось:

{A.17} Функции (1-2) *непрерывны* как при *непрерывной*, так и *дискретной* случайной величине.

Т.е. функции можно применить для анализа дискретных распределений, в том числе – биномиального, который приводит к интересным результатам, но об этом как-нибудь в другой раз.

⁵В принятой теории вероятностей, исходя из равенства $\mu_k = \sum_{j=1}^V (x'_j - m_X)^k \cdot p(x) dx$ дается только вычисление центральных моментов μ_k по начальным моментам ν_k

1.3.3. О трактовке математического ожидания и дисперсии

В существующей теории математическое ожидание часто называют *средним* или *средневзвешенным* значением случайной величины. Заметим, что оба верны, но название «*среднее* значение» следует относить к *экспериментальному*, а «*средневзвешенное* значение» – к *теоретическому* вычислению математического ожидания: это отмечается в некоторых работах. Однако для дальнейшего анализа более интересны трактовки *математического* ожидания и *дисперсии*, данные в работах⁶ [12; 13]:

«С точки зрения наглядности этот способ подсказывается механической аналогией распределения вероятностей ... с распределением масс Полученная таким образом числовая характеристика ... носит название математического ожидания ... ; ее значение определяет ... абсциссу центра тяжести распределения “масс вероятности”» [12,80]. «Физический смысл математического ожидания как характеристики положения центра группирования распределения можно проиллюстрировать ... » [12,82]. «По аналогии с механикой в качестве меры рассеивания рассматривается центральный момент второго порядка ..., его называют дисперсией ... » [12,87]. «Среди числовых характеристик ... нужно, прежде всего, отметить те, которые характеризуют положение случайной величины на числовой оси, т.е. указывают некоторое среднее, ориентировочное значение, около которого группируются все возможные значения случайной величины» [13,85].

«В механической интерпретации математическое ожидание непрерывной случайной величины сохраняет тот же смысл⁷ – абсциссы центра тяжести в случае, когда масса распределена по оси абсцисс непрерывно ... » [13,89]. «Дисперсия случайной величины есть характеристика рассеивания, разбросанности значений случайной величины около математического ожидания. ... Если обратиться к механической интерпретации распределения, то дисперсия представляет собой не что иное, как момент инерции заданного распределения масс относительно центра тяжести ... » [13,96]. Отметим, что в современных работах часто говорится, математическое ожидание – это число и приводится формула без дальнейших пояснений.

Вроде бы одно и то же, но более точная трактовка следует из [12]:

{А.18} *Математическое* ожидание характеризует *положение центра группирования* распределения вероятностей случайной величины, а *дисперсия* – *рассеяние вероятностей около математического* ожидания.

⁶Иногда эта трактовка «попадает» и в других работах

⁷Как и в случае дискретного распределения [13,87]

2. Двумерные (n -мерные) распределения

Под n -мерным понимается как распределение n -мерной случайной величины, так и распределение *произведения* системы, образованной *совмещением* n одномерных случайных величин.

По плотности распределения $p(x_1, x_2, \dots, x_n)$ n -мерной случайной величины определяются плотности *одномерных* $p(x_j)$, *двумерных* $p(x_j, x_k)$ и т.д. случайных величин. Число случайных величин с *размерностью* $j=1, 2, \dots, n$ равно C_n^j числу сочетаний.

Мы рассмотрим характеристики *двумерного* распределения и обратим внимание только на три из них: *математическое* ожидание, *дисперсию* и *ковариацию*. Рассмотрение характеристик n -мерного распределения ограничим только некоторыми замечаниями.

Замечания 6. В качестве числовых характеристик *двумерного* распределения применяются: **1.** Числовые характеристики *одномерных* величин X и Y с плотностями $p(x) = \int_S p(x, y) dy$ и $p(y) = \int_S p(x, y) dx$ {2} (где S – область определения закона распределения): *математические* ожидания $M(X') = \int_S x' p(x, y) dx dy$ и $M(Y') = \int_S y' p(x, y) dx dy$ {3}; *дисперсии в начальных* $D_0(X') = \int_S (x')^2 p(x, y) dx dy$, $D_0(Y') = \int_S (y')^2 p(x, y) dx dy$ {4a} и *центральных* $D(X) = \int_S (x' - m_X)^2 p(x, y) dx dy$, $D(Y) = \int_S (y' - m_Y)^2 p(x, y) dx dy$ {4b}; другие *начальные* и *центральные* моменты порядка k . **2.** *Смешанные* моменты 2-го порядка $K_0(X', Y') = \int_S x' y' p(x, y) dx dy$ и $K(X, Y) = \int_S (x' - m_X)(y' - m_Y) p(x, y) dx dy$ {5}. Произведя замену переменных $x_0 = x + m_X$ и $y_0 = y + m_Y$ в 1-м интеграле, получим равенство: $\nu_{1,1}(X', Y') - \mu_{1,1}(X, Y) = m_X m_Y$ {6}, которое, в принципе, определяется простым *преобразованием – параллельным переносом системы координат*: говорить здесь о «произведении случайных величин как-то не с руки».

По аналогии с механикой, также как моменты $\nu_2(X')$ и $\mu_2(X)$ (замечание 1.2, стр.3), будем называть их *ковариацией в начальных и центральных координатах*, и обозначать $K_0(X', Y')$, $K(X, Y)$ соответственно, подчеркивая этим *общность* их вычисления.

В существующей теории *центральный* момент $K(X, Y)$ называют *ковариацией*. Считается, что он характеризует *рассеивание* величин и *связь* между ними, а *начальный* момент $K_0(X', Y')$ – есть *математическое* ожидание *произведения* случайных величин X и Y . Далее будет показано: моменты $K_0(X', Y')$ и $K(X, Y)$ определяют *поворот* системы *начальных* или *центральных* координат, т.е. имеют *общий* смысл, являются характеристиками *двумерного* распределения, а начальный момент *не имеет отношения к произведению* случайных величин X и Y .

3. *Смешанные* моменты порядка $k > 2$ $\nu_{k,v}(X', Y') = \int_S (x')^k (y')^v p(x, y) dx dy$ и $\mu_{k,v}(X, Y) = \int_S (x' - m_X)^k \cdot (y' - m_Y)^v p(x, y) dx dy$ ($k > 1$, $v > 1$): легко видно, что эти формулы являются общими для всех моментов.

2.1. Математическое ожидание и дисперсия двумерного (n -мерного) распределения

Сначала обратимся к трактовке *математического ожидания двумерного* (n -мерного) распределения.

«Совокупность математических ожиданий m_X, m_Y представляет собой характеристику положения системы. Геометрически это координата средней точки плоскости, вокруг которой происходит рассеивание точки (X, Y) » [13,176]. «Математическим ожиданием n -мерной случайной величины $(\xi_1, \xi_2, \dots, \xi_n)$ называется совокупность n интегралов ...» [1,163], которые определяют математические ожидания $M(\xi_1), M(\xi_2), \dots, M(\xi_n)$ одномерных величин.

Трактовка *математических ожиданий одномерных* величин X_j ($j=1,2, \dots, n$), как характеристики *положения центра группирования* вероятностей n -мерного распределения, есть и в других работах (например, справочниках [8; 17]).

Однако трактовка существует как бы «сама по себе», хотя из смысла цитаты в [13] «само собой напрашивается» ее продолжение:

Найти *рассеивание пар точек* (x, y) относительно *центра группирования*. В *начальных* координатах x и y *взвешенное отклонение* значения *вероятностной функции* $p(x, y)$ в точке (x, y) от *центра группирования* равно $r_0^2 p(x, y)$ (где $r_0^2 = x'^2 + y'^2$ {a.1} – расстояние от точки до начала координат). *Средневзвешенное отклонение от центра группирования*, т.е. *начальная дисперсия* (замечание 1.2, стр.3) *распределения двумерной величины* $\Phi(x, y)$, определяется интегралом $D_0(\Phi) = \int_S r_0^2 p(x, y) dS$ по области определения S . В *центральных* координатах $x = x' - m_X$ и $y = y' - m_Y$ расстояние от точки до начала координат равно $r^2 = (x - m_X)^2 + (y - m_Y)^2$ {a.2}, соответственно, формула вычисления *центральной дисперсии* имеет вид $D(\Phi) = \int_S r^2 p(x, y) dS$.

Подстановкой равенств {a.1}, {a.2} в интегралы, получим: *дисперсия двумерного* распределения определяются *суммами дисперсий* одномерных случайных величин X и Y :

$$D_0(\Phi') = D_0(X') + D_0(Y'), \text{ (a)} \quad D(\Phi) = D(X) + D(Y). \text{ (b)} \quad (3)$$

В существующей теории величину $D(\Phi)$ называют *дисперсией суммы* случайных величин X и Y . При выводе формул (3) о *сумме* даже не упоминалось, следовательно:

{A.19} *К сумме случайных величин дисперсия* $D(\Phi)$ *никакого отношения не имеет. Это – характеристика двумерного* распределения!

Тем самым подтверждается анализ, проведенный в разделе 2.3 [30,51], положение A.26 [30,54] и косвенно – вывод W.19 [30,38].

Аналогично можно показать, что *дисперсии* распределения *n*-мерной случайной величины $\Psi(x_1, x_2, \dots, x_n)$ определяется *суммами дисперсий одномерных* случайных величин X_1, X_2, \dots, X_n :

$$D_0(\Psi') = \sum_1^n D_0(X'_j), \text{ (a)} \quad D(\Psi) = \sum_1^n D(X_j). \text{ (b)} \quad (3^*)$$

2.2. Ковариация как характеристика поворота осей системы координат двумерного распределения: связь с дисперсиями одномерных распределений

Начнем с некоторых положений существующей теории о ковариации и дисперсиях.

Замечание 7. В существующей теории характеристики $D(X_j)$ и $K(X_j, X_k)$ представляются в виде *двумерной* таблицы (матрицы) размером $n \cdot n$ (учитывая равенства $K(X_j, X_k) = K(X_k, X_j)$), в которой дисперсии $D(X_j)$ являются диагональными элементами. В одних работах (например [1,168]) таблицу называют *дисперсией n-мерной* случайной величины, а в других (например [13,185]) – *корреляционной матрицей* случайных величин (X_1, X_2, \dots, X_n) . Утверждается: «... дисперсия каждой из случайных величин есть, по существу, не что иное, как частный случай корреляционного момента, а именно корреляционный момент величины X_i и той же величины X_i : $D_i = M(X_i^2) = M(X_i X_i) = K_{ii}$ » [13,185]. В работе [1,168] еще проще без пояснений: «Очевидно, что $K_{ii} = D_i$ ».

2.2.1. Отличие понятий дисперсии и ковариации

Запись *дисперсий* $D(X_j)$ и *ковариаций* $K(X_j, X_k)$ в виде таблицы *ничему не противоречит*. Математике не важно, как получена таблица и что она означает: в алгебре она представляется в виде $\sum_{j,k=1}^n a_{j,k} x_j y_k$ ($a_{j,k} = a_{k,j}$) $\{*\}$ квадратичной формы и изучаются свойства таблицы [17,257]. В некоторых случаях квадратичная форма $\{*\}$ определяется непосредственно: например, при определении достаточных условий существования экстремумов функции n аргументов [23,423]. Однако в большей части приложений это не так. Например, в механике при движении твердого *недеформируемого* тела в пространстве можно составить трехмерные таблицы: *осевых и центробежных* моментов *инерции* и *скоростей* или *ускорений*; *кинетических* энергий *поступательного* и *вращательного* движения. Для *деформируемого* тела – таблицы *осевых* и *касательных напряжений и деформаций*. Они определяют *не только отличие* понятий, но и *отличие способов* их вычисления. *Свойства* этих таблиц, следующие из свойств квадратичной формы, характеризуют *определенные положения (теоремы)* механики. В механике обычно не обращаются к представлению $\{*\}$: теоремы доказываются непосредственно, ибо важен *смысл доказываемых* теорем, а он следует из *способа вычисления* коэффициентов таблицы.

Равенства $D_i = M(X_i^2) = M(X_i X_i) = K_{ii}$ в цитатах определяют трактовку дисперсий и ковариаций: отсюда вопрос: *является ли дисперсия ковариацией или ковариация дисперсией?*

Пример 3. Пусть $p(x_v, y_w)$ ($v=1,2,\dots,V$) ($w=1,2,\dots,W$) *дискретная вероятностная функция двумерной* величины $\Phi(x, y)$, определенная в прямоугольной области S . По

функции вычисляются: функции $p(x_v) = \sum_{w=1}^W p(x_v, y_w)$, $p(y_w) = \sum_{v=1}^V p(x_v, y_w)$ {2*}, математические ожидания $m_X = \sum_{v=1}^V x'_v p(x'_v)$, $m_Y = \sum_{w=1}^W y'_w p(y'_w)$ {3*} и дисперсии $D(X) = \sum_{v=1}^V x_v^2 p(x_v)$, $D(Y) = \sum_{w=1}^W y_w^2 p(y_w)$ {4*} *одномерных* величин X и Y . *Смешанный момент в начальных* мы не приводим. *Ковариация в центральных* координатах равна $K(X, Y) = y_1 \sum_{v=1}^V x_v p(x_v, y_1) + y_2 \sum_{v=1}^V x_v p(x_v, y_2) + \dots + x_V \sum_{w=1}^W y_w p(x_V, y_w)$ {5*a}. Другой вид $K(Y, X) = x_1 \sum_{w=1}^W y_w p(x_1, y_w) + x_2 \sum_{w=1}^W y_w p(x_2, y_w) + \dots + x_V \sum_{w=1}^W y_w p(x_V, y_w)$ {5*b}.

Замечания. 8. 1. В формулах {2*}-{5*} принято: $x_v = \Delta_v \cdot v$ ($v=0,1,\dots,V$) и $y_w = \Delta_w \cdot w$ ($w=0,1,\dots,W$), т.е. значения вероятностной функции расположены *равномерно* на плоскости; положение значений $p(x_1, y_w)$ и $p(x_v, y_1)$ не совпадают с началом координат. Второе условие принято потому, что только при этом условии *обеспечивается равенство нулю математических ожиданий централизованных* случайных величин X и Y . В этом случае, например, для *биномиального* распределения получим значение *математического* ожидания $m_X = 1 + np$, вместо обычного значения $m_X = np$; дисперсия не изменяется и равна $D(X) = npq$. **2.** Формулы {2*} промежуточные и можно обойтись без них, записав выражения {3*}-{4*} в виде двойных сумм $m_X = \sum_{v=1}^V x'_v \sum_{w=1}^W p(x'_v, y'_w)$, $m_Y = \sum_{w=1}^W y'_w \sum_{v=1}^V p(x'_v, y'_w)$ и $D(X) = \sum_{v=1}^V \{x'_v\}^2 \sum_{w=1}^W p(x'_v, y'_w)$, $D(Y) = \sum_{w=1}^W \{y'_w\}^2 \sum_{v=1}^V p(x'_v, y'_w)$. Что лучше не знаем, но нам больше нравится раздельная запись: дает лучшее понимание «что есть что».

Теперь небольшой анализ формул {2*}-{5*}.

Начнем с формул {5*}, которые определяют одно и то же, т.е. $K(Y, X) = K(X, Y)$ — это просто *разная группировка одной суммы* членов $\sum_{v=1}^V \sum_{w=1}^W x_v y_w p(x_v, y_w)$ *двумерного* распределения. Легко видно, что в этих *суммах* нет *ни математических* ожиданий (формулы {3*}) *одномерных* величин X и Y , ни даже их *вероятностных* функций (формулы {2*}). Есть то, что в принятой теории называют «условными» математическими ожиданиями $\sum_{v=1}^V x_v p(x_v, y_1)$ или $\sum_{w=1}^W y_w p(x_1, y_w)$, но это ничего не меняет — *суммы разные для каждого конкретного значения* x_j ($j=1,2,\dots,V$) или y_k ($k=1,2,\dots,W$).

{А.20} Следовательно, *ковариация* $K(X, Y)$ определяется *по двумерному* распределению, является именно его характеристикой и *не имеет непосредственного отношения* к числовым характеристикам *одномерных* случайных величин X и Y .

Замечания. 1. Вообще говоря, более правильно применять обозначение $K\{\Phi(x, y)\}$ или кратко $K_{x,y}$. Далее будем использовать: краткое обозначение или обозначение $K(X, Y)$, принятое в замечании 6.2 (стр.16), которое «не претендует» на произведение величин. **2.** Ниже будет показано, что имеется *взаимосвязь* между *ковариацией* *двумерного* распределения и *дисперсиями одномерных* величин.

Положим, что *вероятностная* функция $p(x_v, y_w)$ *симметрична* относительно *одной* из прямых линий $y' = m_X$ или $x' = m_Y$, т.е. *симметрична* относительно *одной из центральных осей* Y или X . В этом случае все суммы $\sum_{v=1}^V x_v p(x_v, y_1)$ или $\sum_{w=1}^W y_w p(x_1, y_w)$ равны нулю в силу того, что слева $x_v < m_X$ и справа $x_v > m_X$ от оси Y (или снизу $y_w < m_Y$ и сверху $y_w > m_Y$ от оси X) имеется равное число *одинаковых произведений с противоположными* знаками. Следовательно, в обоих случаях *ковариация* обращается в ноль $K_{x,y} = 0$.

Пусть *вероятностная* функция $p(x_v, y_w) = p(x_v) \cdot p(y_w)$ есть *произведение системы одномерных* случайных величин X и Y , образованной их *совмещением*. В этом случае, учитывая, что *суммы* $\sum_{v=1}^V x_v p(x_v, y_1)$ или $\sum_{w=1}^W y_w p(x_1, y_w)$ принимают вид $p(y_w) \sum_{v=1}^V x_v p(x_v)$ или $p(x_v) \sum_{w=1}^W y_w p(y_w)$, а значения *всех сумм* $\sum_{v=1}^V x_v p(x_v)$ или $\sum_{w=1}^W y_w p(y_w)$, входящих в формулы {5*a} или {5*b} одинаковы (положение {A.19} [30,45]), формулы сводятся к виду $K_{x,y} = p \sum_{v=1}^V x_v p(x_v) \cdot \sum_{w=1}^W y_w p(y_w)$.

Ковариация равна $K_{x,y} = 0$, ибо каждый из *смножителей* равен нулю в силу того, что он является *математическим* ожиданием случайной величины X или Y в *центральных* координатах. Но это не значит, что *ковариация* $K_{x,y}$ есть *математическое* ожидание *произведения* случайных величин. Отсюда следует:

{A.21} *Ковариация $K_{x,y}$ распределения в центральных координатах равна 0 при выполнении одного из условий:* а) *распределение симметрично относительно одной из центральных осей;* б) *распределение есть произведение $p(x_v, y_w) = p(x_v) \cdot p(y_w)$ вероятностных функций системы одномерных случайных величин X и Y , образованной их совмещением.*

Вообще говоря, распределения *одномерных* величин X и Y , полученные по *вероятностной* функции *двумерной* величины $\Phi(x, y)$ могут быть *одинаковы* (пример 16, табл.5 [30,52]). При равенстве распределений *одномерных* величин X и Y , можно записать равенство $Y=X$ (или $X=Y$): в этом случае *ковариация* *двумерного* распределения (формулы {5*a}-{5*b}) примет вид $K_{x,x} = x_1 \sum_{v=1}^V x_v p(x_1, x_v) + x_2 \sum_{v=1}^V x_v p(x_2, x_v) + \dots + x_V \sum_{v=1}^V x_v p(x_V, x_v)$. Из сравнения полученного выражения с 1-й из формул {4*}, очевидно, что между ними нет ничего общего. Перемножая *математические* ожидания величин X и Y (формулы {3*}), исходя из того, что *ковариация* $K_{x,y}$ есть «*математическое* ожидание их *произведения*», получим $K_{x,x} = x_1 p(x_1) \sum_{v=1}^V x_v p(x_v) + x_2 p(x_2) \sum_{v=1}^V x_v p(x_v) + \dots + x_V p(x_V) \sum_{v=1}^V x_v p(x_v)$. Из сравнения с 1-й из формул {4*}, очевидно, что это выражение тоже не равно дисперсии. А если положить $p(x_v, y_w) = p(x_v) \cdot p(y_w)$, то дисперсия вообще равна нулю.

{A.22} *Равенства $D(XX) = K_{x,x}$ и $D(YY) = K_{y,y}$ не имеют места, ибо, по существу: дисперсии $D(X)$ и $D(Y)$ – характеристики одномерных распределений, а ковариация $K_{x,y}$ – характеристика двумерного распределения. Правила их вычисления значительно отличаются (приложение I [29,93]).*

Далее проведем анализ того, что в *реально* определяет ковариация⁸ $K(\mathbf{X}, \mathbf{Y})$ двумерного распределения, которая в существующей теории трактуется как «... характеристика системы, описывающая, помимо *рассеивания* величин, еще и связь между ними» [13,177].

2.2.2. Поворот центральных координат на плоскости

В предыдущем разделе, на основе анализа формул, определяющих дисперсии и ковариацию в центральных координатах, во-первых, определены условия, при которых ковариации обращается в ноль. Во-вторых, показано: равенств $D(\mathbf{X}\mathbf{X}) = K_{x,x}$ и $D(\mathbf{Y}\mathbf{Y}) = K_{y,y}$ между дисперсиями и ковариациями не существует.

Выясним, что же на самом деле определяет ковариация.

Пример 4. Пример с *дискретным* распределением вероятностей рассмотрен потому, что он дает наглядное представление о различии рассмотренных характеристик. Чтобы выяснить, что определяет ковариация $K(\mathbf{X}, \mathbf{Y})$ в *центральных* координатах, удобнее использовать *непрерывное* распределение $p(x, y)$ (не надо писать громоздкие выражения для сумм): его числовые характеристики вычисляются по формулам {2}-{5} (замечания 6,2 стр.16).

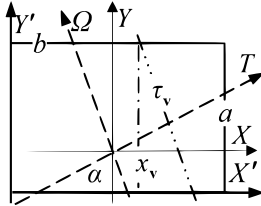


Рис. 2.

п.1. Введем *новые* координаты $T\Omega$ (рис.2, штриховые стрелки), связанные с координатами x_0y выражениями $\tau = x \cos \alpha + y \sin \alpha$, $\omega = -x \sin \alpha + y \cos \alpha$: для определенности⁹ $0 < \alpha < \pi/2$. Это означает: координаты $T\Omega$ повернуты относительно *центральных* координат X_0Y на угол α , что следует из геометрии. В новых координатах *дисперсии* одномерных величин T, Ω и *ковариация* равны: $D(T) = \int_S \tau^2 p(x, y) dx dy$, $D(\Omega) = \int_S \omega^2 p(x, y) dx dy$, $K(T, \Omega) = \int_S \tau \omega p(x, y) dx dy$. После преобразований¹⁰ получим равенства:

$$D(T) = [D(\mathbf{X}) + D(\mathbf{Y})]/2 + \{[D(\mathbf{X}) - D(\mathbf{Y})]/2\} \cdot \cos(2\alpha) + K(\mathbf{X}, \mathbf{Y}) \cdot \sin(2\alpha), \quad (\text{A})$$

$$D(\Omega) = [D(\mathbf{X}) + D(\mathbf{Y})]/2 - \{[D(\mathbf{X}) - D(\mathbf{Y})]/2\} \cdot \cos(2\alpha) - K(\mathbf{X}, \mathbf{Y}) \cdot \sin(2\alpha), \quad (\text{B})$$

$$K(T, \Omega) = -[D(\mathbf{X}) - D(\mathbf{Y})]/2 \cdot \sin(2\alpha) - K(\mathbf{X}, \mathbf{Y}) \cdot \cos(2\alpha). \quad (\text{C})$$

⁸Обычно применяется так называемый коэффициент корреляции $r = K(\mathbf{X}, \mathbf{Y})/\{\sigma_x \sigma_y\}$, который «является мерой силы связи (линейной связи) между величинами» [1,168]

⁹Принято для упрощения изложения, чтобы «не возиться со знаками»

¹⁰Преобразования, приводящие к формулам (A), (B) и (C), можно посмотреть в работе [28]: они легко следуют из геометрических построений

При непрерывном изменении угла α равенства определяют зависимость числовых характеристик *центрированных* величин в координатах $T0\Omega$ от характеристик в координатах $X0Y$ при повороте осей координат.

{А.23.1} Изменение числовых характеристик при повороте *центральных* координат связано с изменением распределения вероятностей – это означает; *исходное* распределение $p(x, y)$, заданное в координатах $X0Y$, преобразуется в распределение $p(\tau, \omega)$, определенное в координатах $T0\Omega$.

Замечания. 1. Отсюда следует необходимость: анализа равенств (А-С), определяющих *взаимосвязь и свойства числовых* характеристик при повороте координат; определение *функции плотности (закона распределения)*, полученной поворотом *центральных* координат. 2. Решение второй задачи является более *важным*: оно позволяет определять вероятности в некоторой части области определения распределения, а не только числовые характеристики.

п.2. Сначала рассмотрим свойства *дисперсий и ковариации* в *центральных* координатах при повороте осей координат.

I. Сложив (А) и (В), получим $D(T) + D(\Omega) = D(X) + D(Y) = D\{\Phi(x, y)\}$ (4). При значении $\cos(2\alpha) = 0$ получим равенства $D(T) = [D(X) + D(Y)]/2 \pm K(X, Y)$, $D(\Omega) = [D(X) + D(Y)]/2 \mp K(X, Y)$ (4*), где знак зависит от значения $\sin(2\alpha)$. Третий независимый подход снова привел нас к *сумме дисперсий*, которая «никак не хочет свестись к *дисперсии* суммы случайных величин X и Y ».

{А.23.2} Из равенства (4) следует, что *дисперсия двумерного* распределения не зависит от поворота системы *центральных* координат. Т.е. равенства (4*) *подтверждают*, что теоремы $D(X + Y) = D(X) + [D(Y) \pm K(X, Y)]$ *не существует* (п.2 примера 16 [30,53]).

Прямые линии $y = x$ и $y = -x$, определяющие преобразование *двумерного* распределения (п.2 примера 16 [30,53]) перпендикулярны: *ортогональность формул преобразования* просто обеспечивает поворот системы координат на угол $\alpha = \pi/2$.

II. Если $\alpha = 0$, то из равенств (А), (В) и (С) получим $D(T) = D(X)$, $D(\Omega) = D(Y)$ и $K_{T,\Omega} = K_{x,y}$, если $\alpha = \pi/2$ – $D(T) = D(Y)$; $D(\Omega) = D(X)$ и $K_{T,\Omega} = -K_{x,y}$. Т.е. *ковариация* $K_{T,\Omega}$ *изменяет* знак. Следовательно: *при повороте осей* $\tau0\omega$ на некоторый угол $0 < \alpha_0 < \pi/2$ *ковариация* $K_{T,\Omega} = 0$.

{А.23.3} Оси $T0\Omega$, в которых *ковариация* равна $K(T, \Omega) = 0$ нулю, будем называть *главными осями* рассеивания (или просто *главными осями*). Таким образом, *ковариация* $K(X, Y)$ характеризует *поворот главных осей* относительно *центральных осей* координат.

Замечания. 1. В математике это называется *приведением квадратичной* формы к *каноническому (диагональной матрице)* виду или к *главным осям* [17,258]. 2. Т.е. фактически, положение является определением: оно дает название координатным осям, в которых *ковариация двумерного* распределения *равна нулю*, при этом дисперсии $D(T)$ и $D(\Omega)$ принимают наибольшее и наименьшее значения (п.4 данного примера). 3. Название *главные оси* рассеивания позаимствовано из [13,194], которое введено при рассмотрении нормального закона на плоскости. Отметим, что в работе [13,349] при обработке результатов стрельб даны формулы, преобразование которых приводит к формулам (А)-(С). Но опять же – для нормального закона на плоскости и никакого анализа этих связей. Поэтому мы следуем работам [27; 28], результаты которых легко

распространяются на произвольные двумерные распределения.

III. Обращение ковариации $K_{T,\Omega}$ в ноль означает (положение {A.21}, стр.20): либо вероятностная функция $p(\tau, \omega)$ симметрична относительно одной из осей T или Ω , либо вероятностная функция $p(\tau, \omega)$ является произведением $p(\tau, \omega) = p(\tau)p(\omega)$ вероятностных функций совмещенных одномерных величин в координатах $T0\Omega$.

IV. При значении $K(T, \Omega) = 0$ из выражений (A), (B) и (C) определяется формула $\text{tg}(2\alpha_0) = 2K_{x,y}/\{D(X) - D(Y)\}$ (5) (учитывая знаки ковариации и разности дисперсий) вычисления угла α_0 между главными и центральными осями. При значении угла $\alpha_1 = \alpha_0 + \pi/4$ ковариация принимает максимальное или минимальное значение $K(T, \Omega)|_{\min}^{\max} = \pm[D(X) - D(Y)]/2$, при этом дисперсии $D(T)$, $D(\Omega)$ одинаковы и равны значению $D(T) = D(\Omega) = [D(X) + D(Y)]/2$.

п.3. Свойства I-IV наглядно проявляются, когда исходное распределение $p(x, y)$ отвечает одному из условий положения {A.21} (стр.20). В этом случае (рис.2, стр.21), имеем $K_{x,y} = 0$, т.е. при значении $\alpha = \alpha_0 = 0$ главные оси совпадают с центральными осями $x0y$. Из равенств (A), (B) и (C) следует: 1. При повороте осей τ, ω на углы $\alpha = k \cdot \pi/2$ ($k=0,1,2,\dots$) имеем $K_{T,\Omega} = K_{x,y} = 0$. Соответственно, дисперсии $D(T)$ и $D(\Omega)$ при углах $\alpha = 0, \pi$ равны $D(T) = D(X)$ и $D(\Omega) = D(Y)$, а при углах $\alpha = \pi/2, 3\pi/2$ $D(T) = D(Y)$ и $D(\Omega) = D(X)$. Так как ковариация $K_{T,\Omega}(\alpha)$ является непрерывной функцией угла α , то между ее нулями имеются экстремумы. Приравнивая производную от выражения (C) нулю (учитывая, что $K_{x,y} = 0$), получим значения необходимых углов $\alpha = (2k - 1) \cdot \pi/4$ ($k=1,2,\dots$). Пусть, например, при значении $\alpha = 0$ имеем $D(T) = D(X) > D(Y) = D(\Omega)$, тогда: 1. При углах $\alpha = \pi/4, 5\pi/4$ ковариация имеет минимум $K(T, \Omega)_{\min} = -[D(X) - D(Y)]/2$, а при углах $\alpha = 3\pi/4, 7\pi/4$ — максимум $K(T, \Omega)^{\max} = [D(X) - D(Y)]/2$. Дисперсии $D(T)$ и $D(\Omega)$ при этих значениях углов одинаковы и равны значению $D(T) = D(\Omega) = [D(X) + D(Y)]/2$. 2. При изменении угла в пределах $0 < \alpha < \pi/2$ дисперсия $D(T)$ уменьшается, а дисперсия $D(\Omega)$ — увеличивается, при угле $\alpha = \pi/4$ они становятся равными, а при угле $\alpha = \pi/2$ — имеем обратное неравенство $D(T) = D(Y) < D(X) = D(\Omega)$.

При дальнейшем увеличении угла процесс идет в обратном направлении, пока при угле $\alpha = \pi$ не придем к исходному неравенству $D(T) = D(X) > D(Y) = D(\Omega)$. Затем процесс повторяется. Таким образом, при значениях ковариации $K_{\tau,\omega} = 0$, дисперсии $D(T)$ и $D(\Omega)$ принимают наибольшее и наименьшее значения: если одна имеет максимум, то вторая — минимум и наоборот. Ни при каких значениях углов поворота они не могут быть равными нулю!

Из п.3 и свойства III следует еще одно свойство:

V. Если двумерное распределение $p(x, y)$ является произведением $p(x, y) = p(x)p(y)$ системы одномерных случайных величин X и Y , образованной совмещением, то при повороте осей координат на угол α , отличающийся от значений $\alpha_0 = (2k - 1)\pi/4$ ($k=1,2,\dots$) оно превращается в координатах $T0\Omega$ в распределение $p(\tau, \omega)$ двумерной величины $\Phi(\tau, \omega)$.

п.4. При значениях углов $\alpha_0 = k\pi/2$ ($k=1,2,\dots$) распределение каждый раз является произведением совмещенных величин. С принятой точки зре-

ния, при отклонении от этих углов происходит «превращение некоррелированных случайных величин в зависимые и коррелированные случайные величины».

Утверждение легко проверяется, кроме того оно следует из работы [1,146], в которой рассмотрено преобразование при повороте осей координат, только выполнено оно наоборот: рассмотрена плотность двумерного распределения 2-х линейно «коррелированных» нормально распределенных величин X и Y .

В работе утверждается: 1. Поворотом координат «любая нормально распределенная случайная величина ... может быть приведена к системе 2-х нормально распределенных независимых величин» [1,147]. 2. Если «... двумя различными способами можно было на плоскости выбрать оси координат $\xi_1 O \xi_2$ и $\eta_1 O \eta_2$ такие, что координаты ξ_1 и ξ_2 , η_1 и η_2 , рассматриваемые как случайные величины с заданным распределением вероятностей, были бы независимы» [1,147], то распределение будет нормальным.

Зависимости (А), (В) и (С) вообще не рассматриваются, в отличие от работы¹¹ [13]. Что второе утверждение не верно, очевидно из свойства III: произведение нормальных законов – только частный случай распределения произведения системы, определяемой совмещением одномерных величин. А еще существуют системы, образованные совмещением двумерных (n -мерных) случайных величин или разной размерности, которые мы не рассматривали.

Что касается 1-го утверждения, то из п.4 следует: с принятой точки зрения, случайные величины «некоррелированные и независимые» только при значениях углов $\alpha = k \cdot \pi/2$ ($k=1,2,\dots$): при любом отклонении от этих углов происходит «их превращение в зависимые и коррелированные» случайные величины. Зависимости (А), (В) и (С) определяют связь числовых характеристик плотности в координатах $T0\Omega$ с числовыми характеристиками плотности в координатах $X0Y$. Чтобы определить какое распределение определяет характеристики в координатах $T0\Omega$, следует определить функцию плотности $p(\tau, \omega)$, которое связано с заменой координат $x = \tau \cos \alpha - \omega \sin \alpha$ и $y = \tau \sin \alpha + \omega \cos \alpha$ в функции плотности $p(x, y)$.

п.5. Запись в виде $p\{\tau \cos \alpha - \omega \sin \alpha, \tau \sin \alpha + \omega \cos \alpha\}$ – это просто запись, которая означает, что в выражении плотности (закона) распределения $p(x, y)$ следует координаты x и y выразить через координаты τ и ω .

Однако есть ли смысл в таком преобразовании?

Во-первых, функция $p(x, y)$ определена в прямоугольной области ($0 \leq x < l_x, 0 \leq y < l_y$). Определяя ее в координатах $T0\Omega$ мы, в общем случае усложняем ее. Например, если плотность $p(x_v, y_w)$ является произведением плотностей $p(x_v)$ и $p(y_w)$, то в результате преобразования она таковой не будет. Кроме того, для вычисления вероятностей необходимо преобразовать к координатам $T0\Omega$ и границы области определения функции плотности.

Во-вторых, учитывая, что при повороте координат, распределение и границы области его определения не изменяются, проще вычислять вероятности по исходному распределению $p(x, y)$. Если одна из границ области интегрирования в координатах $T0\Omega$ определяется, например, некоторой кривой $\tau = f(\omega)$ (на стр.21 рис.2 использованы прямые¹² линии), то ее следует преобразовать к координатам $X0Y$. Любая из прямых линий $y = c_0 + c_1 \cdot x$ в исходных координатах $X0Y$, преобразуется в прямую линию в новых координатах $T0\Omega$ и наоборот. Так например, прямая линии $y = x_v$, параллельная оси Y (обозначена штрихпунктиром), преобразуется в прямую линию, проходящую под углом α к оси ω . Аналогично, прямая линии $\omega = \tau_w$, параллельная оси Ω (обозначена

¹¹ Появляются впервые при рассмотрении нормального закона плоскости [13,194]. Вывод формул дан в разделе, посвященном обработке результатов трельб [13,348-349]

¹² Наиболее простые варианты выбраны для ясности изложения

штрихами с двойным пунктиром), преобразуется в прямую линию, проходящую под углом α к оси Y .

2.2.2. Поворот начальных координат на плоскости

Применим подход, использованный при анализе ковариации в центральных координатах, для анализа ковариации $K_0(X', Y')$ двумерного распределения в начальных координатах, которая в существующей теории трактуется¹³ как математическое ожидание произведения одномерных случайных величин X и Y .

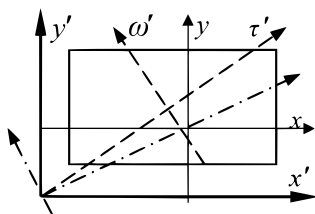


Рис. 4.

Пример 4. п.1. Пусть точка (a, b) – центр группирования распределения $p(x, y)$ в начальных координатах $X'OY'$ (рис.4). Дисперсии и ковариация в начальных координатах x, y связаны равенствами $D_0(X') - D(X) = a^2$, $D_0(Y') - D(Y) = b^2$ (формула {1}, замечание 1.2, стр.3) и $K_0(X', Y') - K(X, Y) = ab$ (формула {6}, замечание 6.2, стр.16). В новых координатах $\tau' = x' \cdot \cos \beta + y' \cdot \sin \beta$ и $\omega' = -x' \cdot \sin \beta + y' \cdot \cos \beta$ (рис.4, штриховые стрелки) получим выражения, аналогичные выражениям (А), (В) и (С), в которые будут входить числовые характеристики с индексом «0», координаты и случайные величины со штрихами, а также угол β .

Оси $\tau'\omega'$ в которых ковариация равна $K_0(\tau', \omega') = 0$, являются главными осями в начальной системе координат.

{А.23.4} Таким образом, ковариация $K_0(X', Y')$ характеризует поворот главных осей относительно начальных осей координат. Соответственно, свойства дисперсий и ковариации в начальных координатах при повороте осей координат, аналогичны свойствам, отмеченным в п.2 примера 4 стр.22.

Учитывая, что смысл характеристик $K(X, Y)$ и $K_0(X', Y')$ одинаков, мы и дали им одинаковые названия – ковариация в центральных и начальных координатах соответственно.

Функция плотности преобразованного распределения и границы его области определения, как и при повороте центральных координат, остаются неизменными и определяется на основе таких же рассуждений (п.5 примера 4 стр.24).

Ковариация в начальных координатах и математические ожидания имеют свойства, которых нет (да и не может быть в принципе) у центральных моментов.

¹³ Характеристики $K(X, Y)$ и $K_0(X', Y')$ отличаются только тем, что они вычисляются относительно разных точек. Стоит только «чуть-чуть сдвинуть» начало координат относительно центра группирования, как ее «толкование» резко изменяется: не кажется ли Вам это несколько странным?

VI. Если плотность распределения $p(x, y)$ отвечает одному из условий положения {A.21} (стр.20), то ковариация $K(X, Y) = 0$: из формулы {2} имеем $K_0(X', Y') = ab$ {4}.

Это означает, что оси (стр.25 рис.4, штрихпунктирная стрелка) *повернуты* относительно *начальных* осей координат на угол β_1 , значение которого определяется из равенства $\operatorname{tg} \beta_1 = b/a$ с учетом *знаков математических ожиданий* a и b . *Произведение величин* X и Y , соответственно, и теорема о *произведении их математических ожиданий* тут «вообще не причем».

VII. При значении $\beta = \beta_1$ ось T'_1 проходит через точку (a, b) , т.е. через центр группирования распределения. В этом случае математические ожидания равны $M(T') = \sqrt{a^2 + b^2}$ и $M(\Omega') = 0$, а из формулы {2} следует равенство $K_0(T', \Omega') = K(T, \Omega)$: *ковариации в начальных и центральных координатах обращаются в нуль*.

Это означает, что в системе координат $T'_1 \Omega'_1$, начальные и центральные координаты повернуты относительно исходных координат $X'OY'$ (или XOY) на одинаковый угол $\beta = \beta_1$.

VIII. Математические ожидания распределений в новой системе координат определяются выражениями $M(T') = M(X')\cos\beta + M(Y')\sin\beta$ и $M(\Omega') = -M(X')\sin\beta + M(Y')\cos\beta$.

При значении $\beta = 0$ получим $M(T') = M(X')$ и $M(\Omega') = M(Y'_1)$, а при значении $\beta = \pi/2$ — $M(T') = M(Y')$ и $M(\Omega') = -M(X')$. При повороте на угол $\beta = \pi/2$ имеем *правостороннюю* систему: переходя к *левосторонней* системе координат $\tau'0\omega'$, получим равенство $M(T') + M(\Omega') = M(Y') + M(X')$. Отметим, что подобные равенства имеют место только при углах $\beta = k\pi/2$ ($k=0,1,\dots$).

Учитывая, что функции $y = x$ и $y = -x$ *ортогональны*, становится понятно, почему *одномерные* распределения, полученные *преобразованием двумерного* распределения с *формулами преобразования* $y = (z - x)$ или $y = (x - z')$ (п.1. примера 16 [30,52]), имеют *одинаковые математические ожидания*.

IX. С другой стороны, возведя в квадрат выражения для математических ожиданий $M(T')$ и $M(\Omega')$ в новой системе координат, и сложив их результаты, получим «неожиданное» равенство $\{M(T')\}^2 + \{M(\Omega')\}^2 = \{M(X')\}^2 + \{M(Y')\}^2$ {5}. Хотя и «неожиданное», но вполне «законное» — расстояние между началами 2-х систем координат равно $\sqrt{a^2 + b^2}$ и *не зависит от поворота системы* координат относительно ее начала.

Почти все, что написано в разделах 2.2.1 и 2.2.2, изложено (в чем-то больше, в чем-то меньше) в учебниках по теоретической механике и сопротивлению материалов (например [27; 28]): там нет, конечно же, «произведения масс» (и естественно вероятностей). Изложение, естественно, дается применительно к *осевым и центробежным моментам инерции* твердых тел [27] или *геометрическим характеристикам площадей* [28] (при выводе полагается: масса распределена равномерно по площади) — если принципиально,

то действительным функциям действительных аргументов.

В частности доказывается [27], что любое твердое тело имеет не менее трех главных осей инерции. Мы только продолжили аналогию с механикой, о чем было сказано ранее.

Замечания 9. 1. Сумма осевых моментов (дисперсий в теории вероятностей) называется полярным моментом и доказывается, что он не зависит от поворота координат. 2. Число главных осей может быть бесконечным. Например, в круге задано распределение с центральной симметрией. В частности, если определить произведение 2-х (или 3-х) одинаковых симметричных, в том числе – нормальных¹⁴, распределений в полярной r, θ (или сферической r, θ, ϕ) системе координат, то «превращения независимых и некоррелированных в зависимые и коррелированные» случайные величины «не состоятся»: в этом случае ковариация в центральных координатах не зависит от поворота координат и всегда равна нулю.

Из анализа, данного в разделах 2.2.1 и 2.2.2, следует:

W.2. Дисперсия двумерного (или n -мерного) распределения определяется суммой дисперсий одномерных величин X и Y (или n величин X_1, X_2, \dots, X_n): она не зависит от поворота системы координат.

W.3. Ковариация двумерного распределения, как в начальной, так и в центральной системе координат характеризует поворот главных осей относительно начальных или центральных осей координат.

W.4. Ковариация в центральных координатах является характеристикой двумерного распределения и не определяет никаких характеристик одномерных случайных величин.

W.5. Ковариация в центральных координатах и дисперсии одномерных распределений взаимосвязаны между собой, но это совсем не означает, что дисперсии являются частным случаем ковариации (замечание 7 стр. 18).

W.6. Ковариация в начальных координатах не является математическим ожиданием произведения случайных величин: она не имеет никакого отношения к их произведению.

W.7. Ковариация в центральных (или начальных) координатах не определяет функциональной зависимости, а ее равенство нулю означает: направление главных осей совпадает с направлением центральных (или начальных), но никоим образом – об отсутствии функциональной зависимости между величинами.

Отличие ковариации от нуля можно воспринимать как признак возможного наличия функциональной зависимости между случайными величинами, но не более: для определения функциональной зависимости необходимы другие методы, о которых поговорим при рассмотрении видов зависимости случайных величин (часть VI исследований).

¹⁴Для произведения 2-х нормальных законов – это распределение Релея [13,200]

W.8. Поворот на угол α центральных (или начальных) координат на плоскости определяет преобразование распределения $p(x, y)$, заданного в координатах XOY (или $X'OY'$), в распределение $p(\tau, \omega)$, определенное в координатах $T0\Omega$ (или $T'O\Omega'$). Оно устанавливает однозначную функциональную связь между распределениями вероятностей.

Осталось выяснить, описывает ли ковариация $K(\mathbf{X}, \mathbf{Y})$ в центральных координатах то, что утверждается в существующей теории: **рассеивание случайных величин и связь между ними**. Вообще говоря, уже из анализа, данного в разделах 2.2.1 и 2.2.2, понятно: ковариация $K(\mathbf{X}, \mathbf{Y})$ не имеет отношения, по крайней мере, к рассеиванию случайных величин \mathbf{X} и \mathbf{Y} . Однако окончательные выводы будут сделаны после более подробного обсуждения этих вопросов при рассмотрении видов зависимости случайных величин (часть VI исследований).

3. Линейное преобразование двумерного распределения в трехмерном пространстве

С точки зрения теории действительных функций действительных аргументов, поворот координат также относится к точечному преобразованию плоскости. Подход привел к идее применения некоторого преобразования в пространстве: она определила определенный вид линейного преобразования двумерного распределения, который рассмотрен в данном разделе. Если говорить по сути, то этому способствовало решение задачи определения вероятности (эффективности) поражения цели, рассмотренной в следующем разделе. Вообще говоря, решение было найдено намного раньше, чем его обоснование, полученное в данном разделе.

Пример 5. Введем систему координат $X'Y'Z'$ и определим на плоскости $X'OY'$ двумерное распределение $p(x, y)$ в прямоугольной области $(0 \leq x' \leq a, 0 \leq y' \leq b)$, где $a = l_{\text{dia}} \cos \beta$, $b = l_{\text{dia}} \sin \beta$ — стороны прямоугольника, l_{dia} — длина диагонали (рис.5, штриховая линия, между точками $(0, 0, 0)$ и $(a, b, 0)$), β — угол ее наклона к оси X' .

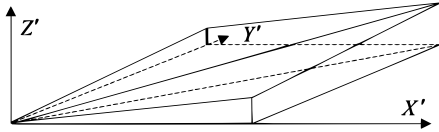


Рис. 5.

п.1. Проведем в плоскости $X'OZ'$ прямую линию $z'_x = x' \operatorname{tg} \alpha_1$ под углом α_1 к оси X' . На плоскости, образованной линией z'_x и осью Y' , получим распределение, ограниченное параллелограммом со сторонами $s_{x,a}^1 = a / \cos \alpha_1$ и $s_{y,b}^2 = b$. Введем на плоскости систему координат $X'_1Y'_1Z'_1$, ось X'_1 которой совпадают с линией z'_x , а ось Z'_1 перпендикулярна этой плоскости. Т.е. ось Z'_1 находится в плоскости $X'OZ'$ под углом α_1 к оси Z' . Проведем в плоскости $Y'OZ'_1$ прямую линию $z'_{1y} = y' \operatorname{tg} \alpha_2$ под углом α_2 к оси Y' . На плоскости, образованной линией

$z'_{1|y}$ и осью X'_1 , получим распределение, ограниченное параллелограммом со сторонами $s'_{x,a} = a/\cos\alpha_1$ и $s'_{y,b} = b/\cos\alpha_2$ **{**}**.

Здесь и далее для упрощения изложения полагается $0 \leq \alpha_1, \alpha_2 \leq \pi/2$.

п.2. Введем систему *косоугольных* координат $T'0\Omega'$, *оси* которой T' и Ω' *совпадают с линиями* z'_x и $z'_{1|y}$. На плоскости $T'0\Omega'$ имеем распределение вероятностей $p(\tau, \omega)$, которое ограничено параллелограммом со сторонами $l_T = l_{\text{dia}} \cos\beta / \cos\alpha_1$, $l_\Omega = l_{\text{dia}} \sin\beta / \cos\alpha_2$ (с учетом **{*}**, **{**}**).

Любая прямая линия $x' = y'_1$ в области определения исходного распределения $p(x, y)$, преобразуется в прямую линию $\tau' = x' \text{tg}\alpha_1$, параллельную оси T' . Пусть на линии $x' = y'_1$ расположены значения функции плотности $p(x, y_1)$. В координатах $T'0\Omega'$ они принадлежат линии $\tau' = \omega'_1$, где значение ω'_1 определяется линейным преобразованием $\omega'_1 = y'_1 \text{tg}\alpha_2$, т.е. $p(\tau, \omega_1) = p(x \text{tg}\alpha_1, y_1 \text{tg}\alpha_2)$. Аналогичными рассуждениями придем к тому, что значения функции $p(\tau, \omega)$ на линии $\omega' = \tau'_1$ определяется линейным преобразованием $p(\tau = \tau_1, \omega = p(x_1 \text{tg}\alpha_1, y \text{tg}\alpha_2))$.

п.3. В общем виде функции *плотности преобразованного и исходного* распределений связаны *равенством* $p(\tau, \omega) = p(x \text{tg}\alpha_1, y \text{tg}\alpha_2)$ (7).

Математические ожидания одномерных случайных величин T и Ω равны $m_T = m_X \text{tg}\alpha_1$ и $m_\Omega = m_Y \text{tg}\alpha_2$, а среднеквадратические отклонения¹⁵ – $\sigma_T = \sigma_X \text{tg}\alpha_1$ и $\sigma_\Omega = \sigma_Y \text{tg}\alpha_2$ соответственно. *Центральные* оси T и Ω *преобразованного* распределения *параллельны начальным* осям T' и Ω' .

Если *исходное двумерное* распределение *симметрично* относительно *центральных осей* X, Y (т.е. без преобразования!), то *преобразованное* распределение *симметрично* относительно *центральных косоугольных осей* T, Ω . Если *исходное* распределение есть *произведение системы*, образованной *совмещением одномерных* величин X, Y , то *преобразованное* распределение будет *произведением системы*, образованной *совмещением* величин T, Ω .

Если повернуть *исходное* $p(x, y)$, а также *преобразованное* $p(\tau, \omega)$ распределения вместе с плоскостью $T'0\Omega'$ на угол ϕ вокруг оси Z (т.е. без преобразования распределения!), проходящей через точку (m_X, m_Y) , то на угол ϕ *изменится* только *направление осей* T и Ω *относительно исходных координат* $X'Y'$. **Всё!!!** Распределение вероятностей $p(\tau, \omega)$ *после этого поворота* будет *точно таким же*, как и *до поворота*.

Из данного анализа следует:

{А.24.1} Если *ковариация исходного* распределения $p(x, y)$ в *начальных* или *центральных* координатах *равна нулю*, то *ковариация преобразованного* распределения в *начальных* или *центральных косоугольных* координатах *тоже равна нулю*.

Т.е. имеем некоторую аналогию с *преобразованием*, которое определяется *поворотом двумерного* распределения на плоскости. На основе проведенного анализа можно утверждать:

¹⁵ Напомним, среднеквадратическое отклонение – корень квадратный из дисперсии

{А.24.2} Таким образом, определено *пространственное линейное* $\tau' = x' \operatorname{tg} \alpha_1$, $\omega = y' \operatorname{tg} \alpha_2$ *преобразование* величины $\Phi(x, y)$, заданной на плоскости $X'OY'$ в величину $\Psi(\tau, \omega)$, определенную в *косоугольной системе* координат $T'O\Omega'$ на плоскости, образованной прямыми линиями $z'_x = x' \operatorname{tg} \alpha_1$ и $z'_{1|y} = y' \operatorname{tg} \alpha_2$ в координатах $X'Y'Z'$.

Законы распределения величин $\Psi(\tau, \omega)$ и $\Phi(x, y)$ связаны *равенством*:
 $P_{\Psi}(\tau < \tau_d, \omega < \omega_d) = P_{\Phi}(x < x_d, y < y_d)$ ($0 \leq \tau_d < z_{x,a}, 0 \leq \omega_d < z_{y,b}$), ($0 \leq x'_d < a, 0 \leq y'_d < b$) (7).

Т.е., получен вид *линейного преобразования*, которому нет аналога в существующей теории случайных величин. По нашему мнению, он более *интересен*, чем *поворот осей* координат. Во-первых, этот вид *преобразования* определяет *существование* распределений в *косоугольных* координатах, которые обладают *такими же свойствами*, как и распределения в *прямоугольных* координатах. Во-вторых, *изменяя значения* углов α_1 и α_2 , получим бесконечное множество распределений в *косоугольных* координатах, которые определяются *одним и тем же* распределением в координатах на плоскости $X'OY'$. В-третьих, в плоскостях $X'OZ'$ и $Y'_0Z'_1$ можно задать (как и при *преобразовании одномерных* случайных величин) *нелинейные однозначные функции* $z'_x = f(x')$ и $z'_{1|y} = g(y')$, тогда в *косоугольных* координатах (например, определяемых точками $(0, 0, 0)$, $\{z'_{x,a} = f(a), 0, 0\}$ и $\{0, z'_{y,b} = g(b), 0\}$) получим распределение на *кривой* поверхности, вероятности которого *однозначно*¹⁶ *определяются* вероятностями *исходного* распределения на плоскости $X'OY'$.

3.1. О применении нового вида преобразования

По-видимому, можно сказать – интересно-то интересно, но (говоря словами кота Матроскина из известного мультфильма) *какая от этого польза?*

В работе [13,199] дан пример определения вероятности поражения цели осколками. Отметим: *определение вероятности поражения* цели – *прекрасный пример полезности* применения понятия *условной вероятности* (что утверждалось в [10.96], вывод S.5.4). Она позволяет рассматривать *раздельно попадание в цель при условии*, что *вероятность поражения* равна 1 (в [13] – *попадание одного осколка приводит к поражению*), и *поражение цели при условии*, что *вероятность попадания* равна 1. Это приводит к *весьма существенному сокращению экспериментальных исследований* и *упрощению расчетов*.

¹⁶Об этом в следующем разделе

Рассмотрим более подробно некоторые моменты оценки вероятности попадания в цель.

Пример 6. п.1. В принятой теории применяется так называемый нормальный закон на плоскости¹⁷ [1; 12; 13], который зависит от пяти параметров. Мы не знаем кто, когда и с какой целью его придумал, но хорошо понимаем, что в рамках принятой теории придумать что-либо лучшее вряд ли возможно.

Хотя и говорится о его применении во многих случаях практики, но в просмотренной литературе показывается его применение в теории стрельбы. Есть примеры [12,176, 323], касающийся измерений, но по-видимому, это исключение (возможно плохо или не там искали). Поэтому выбрана работа [13], ибо, во-первых, в ней рассматривается задача, которую мы решаем далее. Во-вторых, она наиболее простая из просмотренных: некоторые моменты «собранные по крупницам» из других источников.

Сначала приводится запись плотности распределения и доказывается, что она определяет *двумерное* распределение двух *линейно коррелированных* величин X и Y , а его параметры являются математическими ожиданиями m_X, m_Y , дисперсиями $D(X) = \sigma_X^2$, $D(Y) = \sigma_Y^2$ и коэффициентом корреляции¹⁸ $r = K(X, Y) / (\sigma_X \cdot \sigma_Y)$ случайных величин X и Y . Показывается, что сечения поверхности распределения плоскостями, *параллельными* плоскости $X'OY'$, в координатах x, y являются эллипсами, *ориентация* которых относительно координатных осей *зависит от коэффициента корреляции*: их называют *эллипсами равной плотности* или *эллипсами рассеивания*. Общие оси всех эллипсов рассеивания называются *главными осями рассеивания*. Цитата 1: «Рассмотрим специально один частный случай, когда главные среднеквадратические отклонения равны друг к другу $\sigma_X = \sigma_Y = \sigma$. Тогда все *эллипсы рассеивания* обращаются в *круги*, и рассеивание называется *круговым*. ... При некруговом рассеивании случайные величины X, Y независимы тогда и только тогда, когда координатные оси параллельны главным осям рассеивания: при круговом рассеивании независимы при любом выборе прямоугольной системы координат. Эта особенность кругового рассеивания приводит к тому, что оперировать с круговым рассеиванием гораздо удобнее, чем с эллиптическим. Поэтому на практике, где только возможно, стремятся приближенно заменять некруговое рассеивание круговым» [13,196].

Затем полагается: главные оси рассеивания параллельны координатным осям и величины X и Y независимы: в этом предположении выводятся формулы [13,197] вероятности попадания в прямоугольную область. Цитата 2: «Если стороны прямоугольника не параллельны координатным осям, то формулы уже неприменимы. Только при круговом рассеивании вероятность попадания в прямоугольник любой ориентации вычисляется по формулам. Формулы широко применяются при вычислении вероятностей попадания в цели: прямоугольные, близкие к прямоугольным, составленные из прямоугольников или приближенно заменяемые таковыми». Цитата 3: «К числу немногих фигур, вероятность попадания в которые может быть вычислена в конечном виде, принадлежит эллипс рассеивания» [13, 198]. Исходя из канонической формы закона, т.е. для случая совпадения направлений главных осей рассеивания с направлением системы координат $X'OY'$ (или XOY), выводится формула вероятности попадания в эллипс

¹⁷ Наиболее понятное пояснение и его применение дано в [13,188-204]. «Попадается» он и в других работах, например [1; 18], как пример применения тех или иных теорем: понять же, что он собой представляет, по этим работам весьма сложно. Кроме того, в [13,347-351] рассмотрена задача определения его числовых характеристик по экспериментальным данным

¹⁸ Рассматриваются также условные распределения величин X, Y и их математические ожидания, а также линии регрессии: об этом – при рассмотрении теории корреляции (часть VI исследований)

рассеивания. Однако. «Формула чаще всего применяется для вычисления вероятности попадания в круг при круговом рассеивании» [13, 199].

П.2. Судя по литературе, круговое поле осколков и упрощение моделей, принятое в цитате 2, используются не только там, «где это возможно», но и там, где это, мягко говоря, не совсем возможно. В нашем понимании это приводит к значительному снижению точности расчетов.

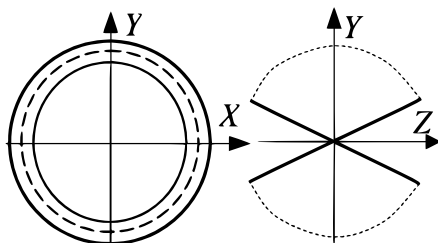


Рис. 6.

есть «картина» кругового поля осколков, описанная в [13]. Она типична для описания поля осколков снарядов. 2. Форма поля осколков может быть иной. В проекции на плоскость $Y0Z$, или любую плоскость, полученную поворотом плоскости $X0Y$ относительно оси Z , их движение может ограничиваться конусом. Т.е. поле осколков расположено между 2-мя концентрическими сферами, ограниченных коническими поверхностями (рис.6 справа, штрихами обозначена только «усредненная» часть сферы). В этом случае скорость осколков в направлении оси Z уменьшается от значения V_{\max} в плоскости $X0Y$ до значения V_{\min} на границах сферической поверхности, что искажает поверхность поля.

Есть и другие причины, влияющие на точность расчета. Например. 1. Движение осколков происходит между 2-мя концентрическими сферами в плоскости $X0Y$, перпендикулярной основной оси Z боеприпаса (рис.6 слева): как следствие случайного различия их скоростей. Направление движения каждого осколка тоже имеет, хотя и небольшой, разброс, что приводит к случайной неравномерности плотности поля осколков. Обычно полагается, что скорости одинаковы и рассматривается некоторая «усредненная» сфера (штриховая окружность). Это и

Но есть «проблемка». Решение задачи возможно только методом Монте-Карло, который работает медленно (даже очень) и для получения приемлемой точности требует большого числа реализаций. А вот объем работ, необходимых для полномасштабной оценки вероятности очень большой. При разработке, учитывая возможности ЭВМ, приходилось выбирать «золотую середину»: чем-то «жертвовать» в модели расчета «во имя» выполнение расчетов в обозримое время. 1. В первую очередь «отмели» причины, связанные с технологией, например, расчеты проводятся по плотности поля осколков на усредненной сфере или ее части, ограниченной коническими поверхностями. 2. Число реализаций свели к некоторому минимуму.

П.3. Конечно же, применяя мощные ЭВМ, отмеченные проблемы можно решить.

Однако основная беда абсолютно не связана с ЭВМ: она определяется *полным отсутствием* в принятой теории *непосредственного преобразования двумерных* (n -мерных) распределений. Впрочем, как и для одномерных распределений. Так называемая «функция от случайных величин» сводится только к преобразованию двумерного распределения к одномерному (раздел 2.1 [30,34]).

Из п.1 очевидно: в принятой теории «вся пляска начинается и продолжается» вокруг записи нормального закона на плоскости. Речь, конечно же, идет об «арсенале средств», имеющихся в *принятой* теории вероятностей: «зависимости и коррелированности» случайных величин, так называемом коэффициенте корреляции и линиях регрессии¹⁹, а *о распределении вероятностей, кроме намека* на то, что оно *меняется*, а вычисления *усложняются – больше ничего*. Даже для простого поворота системы координат на плоскости распределение не определено: есть только запись в новых координатах (п.5 примера 3, стр.22).

Дело в том, что боеприпас имеет *скорость, направление которой* в общем случае не совпадает с основной осью Z боеприпаса: цель также может иметь *определенную ско-*

¹⁹ О теории корреляции – в части VI исследований

рость, направление которой не совпадает ни с осью Z , ни с направлением его скорости. Боеприпас может подходить к цели по самыми различными углами: воздушные цели – с любой точки сферы; морские и наземные цели – ограничиваются поверхностью, на которой они расположены. Отсутствие преобразования двумерных (трехмерных) законов распределения предопределило «неизбежные жертвы»: Во-первых, к упрощению пространственных моделей цели – аппроксимации всех поверхностей цели плоскостями: в принятой теории распределений на кривых поверхностях не существует. Во-вторых, как следствие, поворот моделей цели²⁰ в пространстве: определению проекций моделей целей на плоскость, перпендикулярную основной оси боеприпаса или направлению относительной скорости боеприпаса и цели, в зависимости от вида их взаимодействия. Осуществление этой операции на ЭВМ достаточно сложно и, вообще говоря, занимает тоже немало времени, особенно, если поверхности кривые.

п.4. При применении линейного преобразования, полученного в разделе 4, задача решается очень просто. Вначале простой пример. Рассмотрим распределение $p(x, y)$ двумерной случайной величины на плоскости $X'OY'$, заданное в круге диаметром d . Положим, что центральные координаты параллельны начальным координатам X' и Y' . На плоскости, определяемую прямой линией $z' = x'tg\alpha_1$ и осью Y' , распределение $p(x, y)$ преобразуется в распределение $p(x_1, y)$, ограниченное эллипсом с осями $d_1 = dtg\alpha_1$ и $d_2 = d$ в координатах X'_1OY' . На плоскости, определяемую осью X'_1 и прямой линией $z'_{1|y} = y'tg\alpha_2$, оно преобразуется в распределение $p(\tau, \omega)$, ограниченное эллипсом с осями $d_1 = dtg\alpha_1$, $d_2 = dtg\alpha_2$ [*] в координатах $T'O\Omega'$.

Если повернуть исходное $p(x, y)$, а также преобразованное $p(\tau, \omega)$ распределение вместе с плоскостью $T'O\Omega'$ на угол ϕ вокруг оси Z , проходящей через точку (m_X, m_Y) , то на угол ϕ изменится только направление осей эллипса относительно исходных координат $X'OY'$.

В обоих случаях вероятность события в произвольной части области определения функции $p(\tau, \omega)$ вычисляются (равенство (7)) по ее проекции в области определения функции $p(x, y)$. Однако основная цель рассмотренного примера другая.

п.5. С точки зрения теории, развитой в настоящей работе, двумерный нормальный закон – это произведение одномерных нормальных распределений (при дисперсиях $D(X) = \sigma_X^2 \leq D(YX) = \sigma_Y^2$) системы, образованной совмещением величин X и Y : его дисперсия равна $D\{\Phi(x/y)\} = \sigma_X^2 + \sigma_Y^2$. Т.е. имеем некруговое рассеивание. При одинаковых значениях $\sigma_X = \sigma_Y = \sigma$ [13,196] имеем двумерное распределение с дисперсией $D\{\Phi(x/y)\} = 2\sigma^2$, т.е. с круговым рассеиванием.

Проведя плоскость, например, определяемую осью Y' и прямой линией под углом $\alpha_1 = \arctg(\sigma_X/\sigma)$ оси X' , получим распределение с некруговым рассеиванием, определяемого значениями $\sigma_X = tg\alpha_1$ и $\sigma_Y = \sigma$. Преобразование даже проще, чем при ограниченном распределении вероятностей. Значительно упрощаются и вычисления: определив угол α_1 и применив равенство (7), вычисляем вероятность события, определяемого эллипсами рассеивания по круговому рассеиванию. Т.е. вместо функции $p(\tau, \omega)$ применяем функцию $p(r_n)$, где $r_n = \sqrt{2}(\sigma/n)$ – радиус круга, $n > 0$ – целое число.

Таким образом, достаточно вычислить углы α_1 и α_2 , определяющие направление основной оси Z или относительной скорости боеприпаса по отношению к проекции цели на любую из координатных плоскостей, связанных с координатами $X'Y'Z'$ цели. Далее применяя линейное преобразование, определенное в положении {А.24.2}, (стр.29), по законам распределения, заданных в координатах боеприпаса, вычисляется, напри-

²⁰ Их несколько, что связано с видами взаимодействия боеприпаса и цели

мер, *вероятность* попадания осколков в цель или ее составные части, а также другие *необходимые вероятности*. Никаких поворотов моделей цели не требуется, ибо они создаются в координатах $X'Y'Z'$. Применение преобразования позволяет существенно сократить время расчетов. Во-вторых, учитывать индивидуальность движения каждого осколка. В-третьих, проводить аппроксимацию кривыми поверхностями, по крайней мере, приближенные к реальным поверхностям цели.

Список литературы

1. В.В. Гнеденко. Курс теории вероятностей. Учебник – Изд. 6-е. – М. “Наука”. Главная редакция физико-математической литературы, 1988г. – 448с
2. Л.Е Майстров. Теория вероятностей. Исторический очерк.– М: Наука, 1967. – 321с
3. О.Б. Шейнин. Теория вероятностей до П. Л. Чебышёва // Историко-математические исследования. – М.: Наука, 1978. – № 23. – с. 284-306
4. Я. Бернулли. О законе больших чисел. – М: Наука, 1986. – 176с
5. Википедия: История теории вероятностей.
6. В. Феллер. Введение в теорию вероятностей и ее приложения, В 2-х томах. Т.1: Пер. с англ. – М. «Мир» 1984. 528с.
7. В. Феллер. Введение в теорию вероятностей и ее приложения, В 2-х томах. Т.2: Пер. с англ. – М. «Мир» 1984. 738с.
8. Г. Корн, Т. Корн. Справочник по математике для научных работников и инженеров. – М. Главная редакция физико-математической литературы, 1984г. – 834с
9. А.Н. Колмогоров. Основные понятия теории вероятностей. – Изд.2-е – М. “Наука”. Главная редакция физико-математической литературы, 1980г. – 120с.
10. И.И. Бондарчук. Теория вероятностей и теория надежности: результаты уточнения основных понятий. – Вена. “East-West” Association for Advanced Studies and Education, 2017. – 166с
11. С.Н. Берштейн. Теория вероятностей. – Изд. 2-е – М.-Л. Государственное технико-теоретическое издательство, 1927, 364с
12. Н.В. Смирнов, И.В. Дунин-Барковский. Курс теории вероятностей и математической статистики. – Изд. 3-е - М. Главная редакция физико-математической литературы, 1969г. – 511с
13. Е.С. Вентцель. Теория вероятностей.- Изд. 7-е – М. Издательский центр «Академия», 2003. – 576с
14. А.А. Марков. Исчисление вероятностей. – Изд.4-е – М.-Л.: – Госиздат, – 1924. – 589с.
15. В.И. Гливенко. Курс теории вероятностей. – М.-Л. Государственное объединенное научно-техническое издательство, 1939г. ? 220с.
16. В.П. Чистяков. Курс теории вероятностей. – М. – «Наука». Главная редакция физико-математической литературы, 1978. – 224с
17. И.Н. Бронштейн, К.А. Семендяев. Справочник по математике для инженеров и учащихся ВТУЗОВ. – М. “Наука”, 1980г. – 976с
18. А.А. Боровков. Теория вероятностей: учебное пособие для вузов. – Изд.2-е – М. – «Наука». Главная редакция физико-математической литературы, 1986. – 432с.
19. И.И. Бондарчук. Критика теории вероятностей, часть I. Теория событий: критика исходных понятий и создание новой исходной система.

20. В.И. Лотов. Теория вероятностей. Конспект лекций для студентов механико-математического факультета Новосибирского Государственного Университета
21. Г.М. Фихтенгольц. Курс дифференциального и интегрального исчисления. Т.1 – Изд. 6-е, стереотипное. – М. “Наука”. Главная редакция физико-математической литературы, 1966г. – 607с
22. Г.М. Фихтенгольц. Курс дифференциального и интегрального исчисления. Т.2 – Изд. 7-е, стереотипное. – М. “Наука”. Главная редакция физико-математической литературы, 1969г. – 800с
23. Г.М. Фихтенгольц. Курс дифференциального и интегрального исчисления. Т.3 – Изд. 5-е, стереотипное. – М. “Наука”. Главная редакция физико-математической литературы, 1969г. – 656с
24. Р. Дедекинд. Непрерывность и иррациональные числа. – Изд. 4-е.– Одесса. Mathesis, 1923. – 44с
25. С.А. Ануфриенко. Введение в теорию множеств и в комбинаторику, Учебное пособие. – Екатеринбург. УрГУ, 1998. – 62с
26. Критика теории вероятностей, часть II. О распределениях скоростей Максвелла-Больцмана, Бозе-Эйнштейна, Ферми-Дирака
27. М.М. Филоненко-Бородич, С.М. Изюмов, Б.А. Олисов, Л.И. Мальгинов. Курс сопротивления материалов. Часть первая. – Изд. 5-е – М. Государственное издательство физико-математической литературы, 1961г. – 656с
28. М. Тарг. Краткий курс теоретической механики. – Изд.5-е – М. “Наука”. Главная редакция физико-математической литературы, 1967г. – 480с
29. И.И. Бондарчук. Критика теории вероятностей, часть III. Теория случайных величин: критика существующих исходных понятий и создание новой исходной системы
30. И.И. Бондарчук. Критика теории вероятностей, часть IV. Преобразование законов распределения случайных величин