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

УДК 519.211; ББК 22.171

Теория вероятностей 4: *подобие идеала*!
ч.IV. Преобразование распределений[1] случайных величин

Бондарчук Игорь Иванович, кандидат физико-математических наук

10 декабря 2019 г.

В существующей теории вероятностей преобразование применяется при решении множества разнообразных задач, но известно в основном под названием «функция от случайных величин». В общем-то, не совсем правильно, и все бы было прекрасно, если бы оно не трактовалось в соответствии с названием, а такая трактовка приводит к тому, что «теряется» его связь с преобразованием действительной функции действительных переменных, а также к появлению утверждений в теории случайных величин, которые не соответствуют реальному положению вещей.

При создании *новой исходной системы* теории вероятностей [19,4-40] 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 прежней [19,1; 29,1] – показать, *какая есть и какой должна быть* теория вероятностей случайных явлений, которые называют *массовыми*[2].**

На основе анализа показано: 1. *Преобразование законов* распределения выполняется *по законам преобразования действительной* функции *действительного* аргумента. 2. *Преобразование* определяет: *функциональную зависимость* между случайными величинами; *влияние распределения* одной величины *на распределение* другой величины; *зависимость распределения* случайной величины от *изменения внутренних или внешних условий проведения эксперимента*. Т.е. качественные рассуждения, приводенные в [19,10-16], «облечены» в математическую форму. 3. *Теоремы о суммах математических* ожиданий и *дисперсий* связаны с 1-м *способом преобразования двумерного* распределении к *одномерному* распределению: они *не имеют отношения к сумме* случайных величин.

---

[1] **Говоря о распределениях мы, конечно же, подразумеваем закон распределения и вероятностную функцию**

[2] **Т.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.

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 принятой теории вероятностей, нумеруются заглавными буквами латинского алфавита. Уточненные предположениия – 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.**

Предисловие

О *преобразовании действительной* функции *действительной* переменной упоминалось неоднократно по разным поводам[3] при изложении исследований, в том числе, при *анализе принятой и разработке новой исходной системы* случайных величин. В существующей теории вероятностей *преобразование* применяется при решении множества задач, но даже если о преобразовании и упоминают, то как-то вскользь, мимоходом. В чем оно заключается и как связано с *преобразованием действительных* функций *действительных* переменных, обычно не поясняется: во всяком случае, как в цитируемых, так и в просмотренных работах (**их много больше**) мы этого не увидели (**возможно – «прозевали»**).

***Комментарий*. В существующей теории вероятностей, чаще всего оно вводтся под названием функция от случайных величин, иногда говорят о суперпозиции функций, весьма редко – о функции случайных аргументов. Реально же, независимо от названия, задача ставится (хотя и в очень общем виде) о преобразовании закона распределения: «...по функции распределения $F(x_1, x_2, ..., x_n)$ совокупности случайных величин $\xi_1, \xi_2, ..., \xi_n$ определить функцию распределения $\Phi(y_1, y_2, ..., y_k)$ величин $\eta_1 = f_1(\xi_1, \xi_2, ..., \xi_n)$, $\eta_2 = f_2(\xi_1, \xi_2, ..., \xi_n)$,..., $\eta_k = f_k(\xi_1, \xi_2, ..., \xi_n)$» [1,135]. Что означает «совокупность величин» не объясняется, но по определениям 'D, D.1' [29,16-17] – многомерная величина уже является функцией от случайных величин $\xi_1, \xi_2, ..., \xi_n$. О преобразовании «молчок», но постановка задачи в цитате говорит сама за себя. Кроме того, среди прочих приводится пример [1,146], непосредственно относящийся к преобразованию – поворот осей координат: об этом при анализе числовых характеристик случайных величин (часть V). Отметим, что в работах приводится много примеров, связанных, косвенно или непосредственно с преобразованием, но за редким исключением – в них одна алгебра и «никакой геометрии».**

Преобразование распределений и применение преобразований в теории вероятностей и надежности рассмотрены в работе [10]. Здесь остановимся *на основных* положениях и рассмотрим более подробно некоторые моменты, которые, во-первых, недостаточно освещены в работе, но важны для понимания преобразования. Во-вторых, *исправление* ошибок, допущенных в [10], которые мы «свалили на въевшиеся стереотипы» (**замечания в приложении II [19,48]**), и *одной серьезной* ошибки допущенной при рассмотрении *преобразование одномерных* распределений с применением *немонотонной формулы преобразования*, которая была замечена только при подготовке настоящей статьи.

[3]**Их перечислением заниматься не будем, ибо при необходимости мы отсылали к этой части работы**

Введение

Мы определили случайную величину как *недействительную* функцию *элементарных* событий, а ее *закон* распределения (**вероятностную функцию**) как *действительную* функцию *действительной* переменной: *именно* поэтому говорим о *преобразовании законов* распределения, а не *самих случайных* величин. При разработке *новой исходной* системы для *одной* величины (**т.е. заданной на множестве *элементарных* событий *одного одномерного или двумерного опыта***) *определены*, фактически, *разные законы* распределения[4] (**вероятностные функции**).

Возможность применения *операций* с *действительными* функциями *действительных* аргументов к *законам* распределения (**вероятностным функциям**) случайных величин следует из определений 2 и 4 [29,48,55], *при условии*, что они *не противоречат определениям* 1 и 3 [29,48,55], а также определению 5 [29,67] возможных операций с ними (**как недействительных функций недействительного аргумента**).

***Замечание* 1. *Операция преобразования законов* распределения *не противоречит* отмеченным *условиям.* Ниже будет показано, что *на основе преобразования* функций, для *одной случайной* величины можно получить огромное число *других законов* распределения. Чтобы не было «путаницы», будем полагать, что *исходный* и *преобразованный законы* распределения (вероятностные функции) *принадлежат разным случайным* величинам и обозначать их *разными* знаками.**

И снова напоминание о некоторых простых и известных «вещах»:

Замечания* 2. 1. *Действительные* числа, как и соответствующие им точки на любой из координатных осей X или Y, *непрерывны* и равномерно заполняют каждую ось. Точка – неопределяемый объект геометрии, не имеющий размера, а между двумя соседними точками, как и *действительными* числами, нет промежутков. Любой конечный отрезок, определенный на оси X или Y, содержит бесконечное (несчетное) множество *действительных* чисел и соответствующих им точек. Отрезок «не сжимается и не растягивается»[5]: его можно *увеличить* (уменьшить) только *дополняя его другим* отрезком (отделяя от него некоторый отрезок) с *бесконечным* множеством точек[6]. 2. Между точками 2-х отрезков *произвольной* длины, *произвольно* расположенных на координатных осях X и Y, можно установить *однозначное соответствие* $y \to x$ с применением *однозначной действительной* функции $y = f(x)$ *действительного* аргумента x. Если функция $y = f(x)$ определяет *обратное однозначное соответствие* $x \to y$, то в этом случае соответствие называют *отображением.

3. *Декартово произведение* $\mathbf{X} \times \mathbf{Y}$ тоже *отображает множество то-*

[4] **При этом их отличие мы связывали с преобразованием законов распределения**

[5] **«Растяжение (сжатие)» отрезка означало бы, что «растягиваются (сжимаются)» точки и/или промежутки между соседними точками, но это невозможно**

[6] **Мы не можем сказать, что одно бесконечное множество меньше, равно или больше другого бесконечного множества. Однако используя понятие длины отрезка $l_j = |b_j - a_j|$ (где a_j , b_j $(a_j < b_j)$ – границы отрезка с номером j=1,2,...) можно «ранжировать» (сравнивать между собой) эти множества. Т.е. подобно сравнению бесконечно малых и бесконечно больших величин в теории действительных функций действительного аргумента [21,136]**

чек отрезка, определенного на одной оси, на *множество точек* отрезка, заданного на другой оси. Однако, учитывая, что *понятие множества всех отображений из одного множества в другое* слишком *широкое* и требует «расшифровки[7]» при *рассмотрении того или иного вида преобразования*, применять понятие *декартова произведения не будем.*

[7] В некоторых случаях она может оказаться существенно «длиннее» пояснения сути преобразования

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

1. *Преобразование одномерных* распределений.

***Функциональная зависимость* распределений. О *равновероятном* и *равномерном* распределениях**

О *преобразовании одномерных* распределений в теории вероятностей говорится мало: фактически оно подробно рассматривается (**под названием «функция Y = U(X) случайного аргумента X»**) в работе [13,263] при *различных* видах *формулы преобразования.* В некоторых из (**как цитируемых, так и просмотренных**) работ рассматривается [16,100; 20,44] только **«линейная функция $\mathbf{Y} = A\mathbf{X} + B$ (где A и B – действительные числа) от величины X».** Почему это так, сказать сложно: можно только «гадать» – толи все понятно и нет вопросов, толи возникает какая-то «закавыка», которая «не поддается объяснению». Вообще-то, нам пришлось «долго и мучительно разбираться» с ошибкой (**раздел 1.2, стр.14**), допущенной в работе [10]. Она рассмотрена весьма подробно, но основной «упор» делается на сути выполняемых преобразований.

Сначала пример из механики[8].

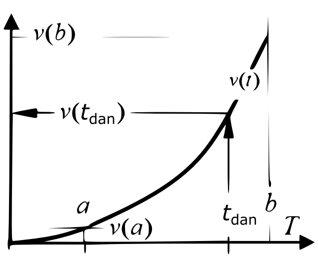

Рис. 1.

***Пример 1.* При действии силы $F = ct$ (где c – постоянная, определяющая переход от размерности времени t к размерности силы F) на точку массой m ее скорость v равна $v = t^2\cdot\{c/(2\cdot m)\}$ {1}. Таким образом, *действительная* переменная v *связана однозначной* зависимостью $v = v(t) = v\{F(t)\}$ (рис.1) с действительной переменной t. Т.е. каждому значению t_j соответствует значение $v_j = v(t_j)$. По другому, функция $v_j = v(t_j)$ преобразует переменную t (в масштабе и размерности времени) в переменную v (в масштабе и размерности скорости). Если переменная t определена на отрезке $[a,\ b]$ оси T (рис.1), то переменная v – на отрезке $[v(a),\ v(b)]$ оси V. Как значения t, так и значения v *<u>равномерно</u>* заполняют отрезки. Очевидно, что каждому значению $v'_{\mathbf{dan}}$ будет соответствовать значение $t'_{\mathbf{dan}} = w(v'_{\mathbf{dan}})$, где $w(v)$ – функция, обратная функции $v(t)$.**

По-другому и в более общем виде:

На множестве *действительных* чисел X *определена монотонно*

[8]**Пример из механики (расширенный вариант – в [10,49]), ибо автор – механик. Во-вторых, аналогию между математическим ожиданием и центром тяжести, дисперсией и осевым моментом инерции (замечание [29,4]) следует продолжить: «задействовать полярный и центробежный моменты инерции; связь моментов с преобразованием координат. Об этом при рассмотрении числовых характеристик случайных величин (часть V исследований)**

возрастающая (в строгом смысле [21,270]) ***непрерывная***[9] ***действительная*** **функция** $y=y(x)$, ***значения*** **которой** ***принадлежат*** **множеству** ***действительных*** **чисел** Y (**т.е. функция отображает (преобразует) множество** X **на множество** Y **или наоборот – множество** Y **на множество** X).

Если на множестве X **определена** ***дискретная действительная*** **функция** $V(x_j)$ (**j=1,2,...,N**), **то функция** $y=f(x)$ ***преобразует*** (**отображает на множество** Y) **ее в функцию** $W(y_j) = V\{x_j = \phi(y_j)\}$ (**где функция** $x=\phi(y)$ **обратная к функции** $y=f(x)$). **Обратно. Если на множестве** Y **определена функция** $W(y_j)$, **то функция** $y=f(x)$ ***преобразует*** (**отображает на множество** X) **ее в функцию** $V(x_j) = W\{y_j=f(y_j)\}$. **При** ***преобразовании дискретной*** **функции** ***изменяются*** **только** ***расстояния между*** **значениями функции,** ***но не ее значения.***

Напомним: этот тип ***отображения*** **в теории** ***действительных*** **функций** ***действительных*** **переменных в общем случае называется [21,483]** ***точечным преобразованием*** **плоскости, а функция** $y=f(x)$ – ***формулой преобразования. Цели преобразования*** **в теории** ***действительных*** **функций и теории вероятностей несколько различаются, но это не имеет значения: – сущность и формулы одинаковы.**

> **{A.1.1}** ***Формула преобразования*** $y = f(x)$ **определяет** ***связь*** **множества** ***действительных*** **чисел** X **с множеством** ***действительных*** **чисел** Y. **Это означает, что появление** ***конкретного*** **значения** x' ($x_{\mathbf{H}} \leq x' \leq x_{\mathbf{K}}$) **из множества** X **однозначно определяет одновременное появление только конкретного значения** $y' = f(x')$ ($y_{\mathbf{H}} \leq y' \leq y_{\mathbf{K}}$) (**где** $y_{\mathbf{H}} = f(x_{\mathbf{H}})$, $y_{\mathbf{K}} = f(x_{\mathbf{K}})$) **из множества** Y.

Формула преобразования $y=f(x)$ не *зависит ни от* функции $V(x)$, *ни от* функции $W(y)$. Все с точностью до «шиворот-навыворот»: **функции** $V(x)$ **и** $W(y)$ ***зависят*** **от функции** $y=f(x)$. Таким образом:

> **{A.1.2}** ***Монотонно возрастающая формула преобразования*** $y=f(x)$ **определяет** ***однозначную функциональную зависимость*** **функции** $W(y)$ **от функции** $V(x)$ **и наоборот: функции** $V(x)$ **от функции** $W(y)$. **Это означает, что** ***по значению*** **функции** $V(x)$ ***точно*** **определяется,** ***какое значение примет*** **функция** $W(y)$ **и наоборот.**

Перейдем к теории вероятностей.

Пример **1.а.** ***Опыт 1.*** **Бросание кубика с разными изображениями на гранях. При проведении** ***опыта,*** **появляется некоторое** ***элементарное событие*** $\mathbf{a_j}$ (**j=1,2,...,6**). ***Поставим*** **ему** ***в соответствие*** **значения времени** $\mathbf{a_j} \to t_j$ **в секундах** (t_j**=1;1.5;2;...;3.5**). **В результате на отрезке** $a=1 \leq t \leq b=3.5$ (**рис.1.а**) **имеем** ***дискретную случайную***[10] **величину T с** ***законом*** **распределения** $P(t_k) = \mathbf{P}(a \leq t_j < t_k) = \sum_{j=0}^{k} p(t_j)$ **{2}** **и** ***вероятностной*** **функцией** $p(t_j) = \mathbf{P}(\mathbf{a_j}) = 1\ /\ 6\ [t_0 = 0;\ \ p(t_0) = 0]$ **{2*}**.

Следовательно, ***случайная*** **величина T имеет** ***равновероятное*** **распределение. Значения функций {2}-{2*} расположены на** ***одинаковых рассто-***

[9] **Полагаем, производная от функции также непрерывна: отличие отмечается в тексте**
[10] **Замечания к выводам Q.II и Q.III [29,44-45])**

яниях, **следовательно, значения распределения величины T *равномерно* расположены на отрезке $a \le t \le b$.**

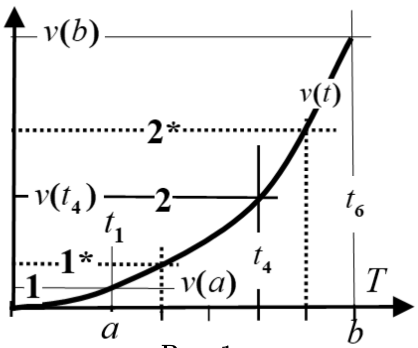

Рис. 1.a

Очевидно, что переменная v (с учетом зависимости $v = v(t)$ {1}) также будет принимать *случайные* значения $v_j = v(t_j)$. Т.е., функция $v = v(t)$ преобразует *закон* распределения (и вероятностную функцию) *случайной* величины T в *закон* распределения $P(v_k) = \sum_{j=0}^{k} p(v_j) = P(t_k = v_k^{1/2}\sqrt{2m/c})$ {3} (где $t = v^{1/2}\sqrt{2m/c}$ {1*} – функция обратная к функции {1}) и вероятностную функцию $p(v_j) = p(t_j = v_j^{1/2}\sqrt{2m/c})$ {3*} случайной величины V. Т.е., ее распределение тоже *равновероятное*. Однако, в отличие от функций {2}-{2*}, значения функций {3}-{3*} расположены на отрезке $v(a) \le v \le v(b)$ *на разных расстояниях* (на рис.1, расстояния между сплошными и штриховыми горизонтальными линиями 1-1* и 2-2* при одинаковых отрезках на оси T). Т.е. значения распределение величины V расположены *неравномерно*.

***Опыт 2.* Можно рассмотреть вынимание наугад шара из урны, в которой находятся шары с номерами j=1,2,...,6 и числами шаров 2,5,10, 25, 28 и 30 соответственно. Закон распределения и вероятностная функция величины T определяются формулами {2}-{2*}, а величины V – формулами {3}-{3*}. Отличаются только значения вероятностной функции. В этом случае распределения величин T и V *не равновероятны*, однако значения распределения (как и в 1-м случае) величины T на отрезке $a \le t \le b$ расположены *равномерно*, а величины V на отрезке $v(a) \le v \le v(b)$– *неравномерно*.**

1.a. *Элементарным* событиям a_j можно поставить в соответствие значения $a_j \to v_j = v(t_j)$ скорости v, вычисленные в заданные моменты времени t_j. Положим: исходным (для обоих опытов) будет распределение величины V, определяемые формулами { 3}-{3*}, т.е. с *неравномерным* расположением значений. Используя функцию {1*} как *формулу преобразования*, получим *равномерные* расположение значений распределения величины T, определяемые формулами {2}-{2*}, т.е. *исходные* распределения, следующие из опытов 1,2.

1.b. Однако нет никакого запрета на то, чтобы положить *расстояния* между значениями v_j одинаковыми. Тогда (для обоих опытов) исходными будут распределения величины V, определяемые формулами типа {2}-{2*}, с *равномерным* расположением значений. Используя функцию {1*} как *формулу преобразования*, получим *законы* распределения (и вероятностные функции) величины T с *неравномерным* расположением значений, определяемые формулами $P(t_k) = \sum_{j=0}^{k} p(t_j) = P\{v_k = t_k^2 c/(2m)\}$ и $p(t_j) = p\{v_j = t_j^2 c/(2m)\}$ соответственно. *Расстояния* между вероятностями соответствуют *значениям*, определяемым *заданной* функцией $v = v(t)$.

Таким образом, с применением формулы *преобразования*[11] $v(t)$, *закон* распределения $\mathbf{P_T}(a \le t_j < t_{\mathbf{dan}})$ *дискретной* случайной величины $\mathbf{T}$ *преобразуется* в *закон* распределения $\mathbf{P_V}[v(a) \le v_j < v(t_{\mathbf{dan}})]$ *дискретной* величины $\mathbf{V}$. Преобразование определяет *функциональную зависи-*

[11] **К функции случайной величины $\mathrm{V} = \mathrm{U(T)} = \mathrm{T}^2$ (в теории множеств – это декартов квадрат, замечание 7.1 [29,20]) от величины T формула преобразования не имеет никакого отношения (и к декартову квадрату тоже)**

мость $v = v\,(t)$ *законов* распределения (**вероятностных функций**) величин $\mathbf{T}$ и $\mathbf{V}$. Она означает (**вывод w.17 [29,52]**), что *элементарные* события величин $\mathbf{T}$ и $\mathbf{V}$ с *метками* «t_j» и «v_j» (**на интервалах** $a \le t_j < t_{\mathbf{dan}}$ **и** $v(a) \le v_j < v(t_{\mathbf{dan}})$) появляются с *одинаковой* вероятностью.

Мы повторили этот простой пример (**в работе [10,49] он дан в расширенном варианте**) для того, чтобы показать: **вероятности как *равновероятного*, так и *не* *равновероятного* распределения *дискретной случайной* величины могут быть *расположены* на отрезке *равномерно* или *неравномерно*.** Во-вторых, лишний раз напомнить, что *закон* распределения (**вероятностная функция**) – это *действительная* функция *действительной* переменной.

Далее будем говорить только о *случайных* величинах и их *законах* распределения (**вероятностных функциях**). Пусть на отрезке $a \le x \le b$ определена *непрерывная случайная* величина $\mathbf{\Phi}$ с *законом* $P_{\mathbf{\Phi}}(a \le x < x_{\mathbf{dan}})$ и *плотностью* $p_{\mathbf{\Phi}}(x)$ распределения соответственно. В общем случае достаточно положить:

{A.2} ***Формула преобразования*** $y = f\,(x)$ **является *непрерывной однозначной* функцией, *ограниченной* на отрезке** $a \le x \le b$**.**

С применением *формулы преобразования* $y = f\,(x)$, *закон* и *плотность* распределения *непрерывной* величины $\mathbf{\Phi}$ преобразуются в *закон* и *плотность* распределения величины $\mathbf{\Psi}$ [10,52] по одному из двух вариантов[12]:

$$\mathbf{P_\Psi}(y_a \le y < y_{\mathbf{dan}}) = \mathbf{P_\Phi}\{\phi(y_a) \le \phi(y) < \phi(y_{\mathbf{dan}})\},\ \textbf{(a)}\quad p_\Psi(y) = p_\Phi\{\phi(y)\}\cdot|\phi'(y)|;\ \textbf{(b)} \qquad \textbf{(1)}$$

$$\mathbf{P_\Psi}\{f(a) \le f(x) < f(x_{\mathbf{dan}})\} = \mathbf{P_\Phi}(a \le x < x_{\mathbf{dan}}),\ \textbf{(a)}\quad p_\Psi\{z = f(x)\} = p_\Phi(x)/|f'(x)|;\ \textbf{(b)} \qquad \textbf{(2)}$$

где $\phi(y)$ **– функция обратная к функции** $f(x)$, $f(a) = y_a \le y_{\mathbf{dan}} < y_b = f(b)$, **значения** z **принадлежат оси** X.

Замечания* 3.** **1.** **В существующей теории (1-й вариант) дифференцируются равенства [13,264; 16,100]:** $P_\Psi(y) = \int_a^{\phi(y)} p_\Phi(x)\,dx$, **если** $f(x)$ ***монотонно возрастает **и** $P_\Psi(y) = \int_{\phi(y)}^b p_\Phi(x)\,dx$, **если** $f(x)$ ***монотонно убывает.*** **2.** **Мы определяем плотность распределения величины** Ψ **определяется дифференцированием равенств: при 1-м варианте – (1.a) по переменной** y**; при 2-м – (2.a) по переменной** x**: это соответствует ее определению.** **3.** **2-й вариант *преобразования* введен нами в [10,53] «на автомате»: *непрерывность закона* распределения и *формулы преобразования* определяет существование обеих производных. При написании работы [10], мы, хотя и рассматривали различия свойств вариантов, но «прошли мимо» одного, пожалуй, наиболее важного. Если *в 1-ом варианте формула преобразования*** $y = f(x)$ **устанавливает *однозначное соответствие только* функции плотности** $p_\Psi(y) \to p_\Phi\{\phi(x)\}$ **с функцией** $p_\Phi(x)$**, то 2-й устанавливает как *прямое*, так и *обратное соответствие*** $p_\Phi(x) \to p_\Psi\{f(x')\}$ **между функциями плотности**

[12] **В существующей теории применяется только 1-й вариант (например [8; 13; 16]). В формулах (3.1) для плотности в работе [10,49] допущена описка: плотность должна умножаться на абсолютное значение производной, что и делается в работе при преобразованиях с немонотонной функцией [10,53]**

и законами распределения. Т.е. *2-й вариант* является *отображением* (замечание 2.2, стр.4). <u>4.</u> *Преобразование непрерывного закона* распределения приводит *не только к изменению расстояния* между значениями вероятностей $P_{\Phi}(x_1)$ и $P_{\Phi}(x_2)$, но и *к изменению* значений *функции плотности*. <u>5.</u> Так как *функции* $y = f(x)$ и $x = \phi(y)$ определяют в системе координат $X0Y$ *одну и ту же кривую* [21,108-109], то *оба варианта* определяют *один и тот же закон* распределения случайной величины Ψ, но *на разных* осях координат. Это будет показано при применении *немонотонной формулы преобразования* (раздел 1.2, стр.14). <u>6.</u> Первый вариант удобнее, если функция $y = f(x)$ разрешается относительно переменной x: закон и плотность распределения величины Ψ непосредственно выражаются через переменную y. Второй – если функция $y = f(x)$ не разрешается относительно переменной x. Утверждение остается в силе, если формула преобразования задана в виде неявной функции $V(y, x) = 0$ <u>7.</u> Из формул (1.a-2.a) следует: преобразование можно рассматривать как суперпозицию функций. Однако напомним: *суперпозиция* функций определяет функцию на *одной числовой оси*, т.е. *одной переменной*. Следовательно; в 1-м *варианте* имеем *суперпозицию* функций в *правой части* равенства (1.a), а во 2-м – в *левой части* равенства (2.a). Учитывая это, *закон* распределения, полученный 1-м *вариантом*, необходимо *представлять на оси* Y, а 2-м – *на оси* X.

Подчеркнем: **из *преобразования* следует *функциональная зависимость*** $y = f(x)$ ***законов* распределений случайных величин Φ и Ψ**.

1.1. *Монотонно возрастающая формула преобразования. Равновероятность дискретных и непрерывных величин*

Пока не рассмотрено *преобразование* с *немонотонной* функцией, полагаем *формулу преобразования* $y = f(x)$ *непрерывной монотонно возрастающей* (**в строгом смысле**), *ограниченной* на отрезке $a \le x \le b$, и рассмотрим *только* 1-й *вариант преобразования.*

***<u>Пример</u>* <u>2</u>. Пусть на отрезке $t_{\mathbf{H}} \le t \le t_{\mathbf{K}}$ задано *<u>равновероятное</u>* («равномерное» в существующей теории) *непрерывное* распределение $P(t) = (t - t_{\mathbf{H}})/(t_{\mathbf{K}} - t_{\mathbf{H}})$ {4}: его *плотность* – $p(t) = 1/(t_{\mathbf{K}} - t_{\mathbf{H}})$ {4*}.**

<u>п.1.</u> Применяя функцию {1} (рис.1, стр.6), получим (обратная функция и ее производная – $t = \sqrt{v \cdot \{(2m)/c\}}$, $t' = v^{-1/2} \cdot \sqrt{m/(2c)}$):

$$P(v) = v^{1/2} \cdot \sqrt{(2m)/[c(t_{\mathbf{K}} - t_{\mathbf{H}})^2]} - t_{\mathbf{H}}/(t_{\mathbf{K}} - t_H); \qquad \{5\}$$

$$p(v) = v^{-1/2} \sqrt{m/\{2c(t_{\mathbf{K}} - t_{\mathbf{H}})^2\}}. \qquad \{5^*\}$$

<u>п.2.a.</u> Если на оси V задано распределение, определяемое формулами {5}-{5*}, то используя *формулу преобразования* {1*}, оно *преобразуется* в исходный (заданный) *закон <u>равновероятного</u> непрерывного* распределения. Т.е. обратное *преобразование* также *однозначно.*

<u>п.2.b.</u> Пусть на отрезке $v_{\mathbf{H}} \le v \le v_{\mathbf{K}}$ задано *<u>равновероятное</u> непрерывное* распределение: $p(v) = 1/(v_{\mathbf{K}} - v_{\mathbf{H}})$, $P(v) = (v - v_{\mathbf{H}})/(v_{\mathbf{K}} - v_{\mathbf{H}})$. Применяя *формулу преобразования* {1*} (обратная функция и ее производная равны $v = t^2 \cdot \{c/(2 \cdot m)\}$, $v' = t \cdot (c/m)$), получим: $p(t) = 2t\big/(t_{\mathbf{K}}^2 - t_{\mathbf{H}}^2)$, $P(t) = (t^2 - t_{\mathbf{H}}^2)\big/(t_{\mathbf{K}}^2 - t_{\mathbf{H}}^2)$. Т.е. *плотность* и *закон* распределения, *соответствующие формуле преобра-*

зования $v = v(t) = t^2 \cdot \{c/(2 \cdot m)\}$ **(формула {1} пример 1),** ***заданной*** **на множестве** ***действительных*** **чисел** T**.**

Из анализа, данного в примере, следует:

{A.3} ***Элементарные*** **события** ***любой непрерывной*** **случайной величины** ***равномерно заполняют*** **область ее определения (закона и плотности распределения соответственно).**
В принципе, это утверждение непосредственно следует из замечаний 2 (стр.4).

Применяя *разные формулы преобразования*, для *одного закона* распределения случайной величины Φ получим *множество преобразованных законов* распределения, поэтому и было принято относить их к другим случайным величинам. Из анализа следует:

W.1. ***Преобразование не зависит от того, что преобразуется, действительная функция или закон распределения случайной величины и выполняется абсолютно одинаково*** **(примеры 1, 1.а, стр.6,7).**

W.2. ***Если задана монотонно возрастающая действительная функция*** $y = f(x)$ ***действительной переменной*** x **(формула преобразования), то** ***закон распределения*** $\mathbf{P}_{\mathbf{\Phi}}(a \le x < x_{dan})$ **(в размерности и масштабе переменной** x**)** ***случайной величины Φ может быть преобразован в закон распределения*** $\mathbf{P}_{\mathbf{\Psi}}(y_a \le y < y_{dan})$ ***случайной*** **величины Ψ (в размерности и масштабе переменной** x**).**

W.3. ***Варианты преобразования равноправны – они устанавливают однозначную связь функций*** $P_{\mathbf{\Psi}}(y)$ **и** $p_{\Psi}(y)$ **с** ***функциями*** $P_{\mathbf{\Phi}}(x)$ **и** $p_{\Phi}(x)$**.** ***Их применение определяется решаемой задачей.***

W.4. ***Формула преобразования*** $y = f(x)$ ***не связана ни с законами*** $P_{\Phi}(a \le x < x_{dan})$ **и** $P_{\Psi}(y_a \le y < y_{dan})$ **распределения, ни, тем более, с величинами Φ и Ψ.** ***Следовательно, трактовать формулу преобразования как функцию от случайной величины недопустимо.***

W.5. ***Преобразование определяет функциональную зависимость законов распределения*** **(вероятностных функций) величин Φ и Ψ. Это означает:** ***появление одного из*** **множества** ***элементарных*** **событий $\mathbf{a_j}$ величины Φ,** ***принадлежащих интервалу*** $x_1 \le x < x_2$ **(вывод W.17 [29,52]) и** ***одного из*** **множества** ***элементарных*** **событий $\mathbf{b_k}$ величины Ψ,** ***принадлежащих интервалу*** $y(x_1) \le y < y(x_2)$ **имеют** ***одинаковую*** **вероятность.**

<u>Замечание.</u> **Вычисление вероятностей сложных событий, определяемых величиной Ψ, не требует знания ее закона распределения. Они легко вычисляются по закону величины Φ и функции** $y = v(x)$**. Т.е. построение распределения величины Ψ, вообще говоря, не обязательно.**

Из выводов **W.1, W.5** и положения **{A.1.2}** (стр.7) следует:

W.6. ***Отличие функциональной зависимости*** **между 2-мя** ***действительными*** **функциями** ***действительного*** **аргумента и 2-мя** ***законами*** **распределения** ***единственное*****: в 1-м случае** ***по значению одной*** **функции** ***точно определяется значение другой*****; во 2-м случае –** ***по значению вероятности*** **события, вычисленной** ***по закону*** **распределения** ***одной*** **величины,** ***определяется точное значение вероятности*** **события, соответствующего** ***другому закону.***

В существующей теории о *преобразовании* не говорится (**стр.3, 6**), несмотря на то, что оно применяется: соответственно, *функциональная зависимость* трактуется *только как зависимость* между 2-мя *действительными* функциями *действительных* аргументов: подробный разговор об этом будет при рассмотрении видов зависимости случайных величин (**часть VI исследований**).

{A.4} ***Подчеркнем важное*****: мы говорим о том, что** ***неверна трактовка преобразования*****, а не о том, что** ***оно выполнено неправильно.***
Предупреждение. **Однако в последующих исследованиях** ***неверная трактовка*** **может привести (и приводит) к** ***неверным*** **заключениям.**

В 1-ой части работы было отмечено (**замечание 2 в приложении II [19,50]**), что *правильно* говорить о *равновероятности* *элементарных* событий *опыта*, а не об их *равномерности*. Как следует из анализа, данного в примере 1 (**опыты 1, 2**), для *дискретных* случайных величин *равновероятность* *элементарных* событий также «стоит на первом месте»:

{A.5} ***Закон*** **распределения** ***дискретной*** **величины** ***равновероятен*** **при условии:** ***вероятности 2-х любых*** **событий,** ***состоящих из одинакового числа*** **(в том числе – одного)** ***элементарных*****,** ***одинаковы.*** **Отсюда следует: вероятностная функция дискретной величины** ***равновероятна*****, если** ***равны*** **вероятности** ***всех элем*****ентарных событий.**

При *неравномерном* расположении *равновероятных* *элементарных* событий отличаются *вероятностные* функции и *числовые* характеристики *дискретной случайной* величины, тем не менее, *закон* распределения остается *равновероятным*. Применять называние «равномерный» – «как-то негоже».

Осталось «разобраться» с *непрерывными случайными* величинами.

Пример **3.** **п.1.** ***Бесконечное*** **(несчетное) множество** ***элементарных*** **событий на конечном отрезке является** ***непрерывным*** **(вывод Q.1 [19,50]), следовательно,** ***элементарные*** **события и их вероятности тоже** ***равномерно*** **заполняют отрезок.** ***Преобразование*** **определяет** ***взаимно однозначное*** **соответствие** ***бесконечного*** **множества точек (элементарных событий),** ***равномерно*** **заполняющих отрезок** ***одной длины*****, с** ***бесконечным*** **множеством точек (элементарных событий),** ***равномерно*** **заполняющих отрезок** ***другой длины.***

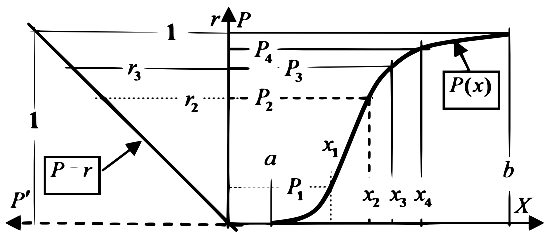

Рис. 2.

При этом *элементарные* события *равновероятны*: *вероятность каждого равна нулю.* Однако *ни равномерность* расположения точек (*элементарных* событий) на отрезке, *ни равновероятность элементарных* событий «нечего не говорят нам» о *законе* распределения и *вероятностной* функции. В [19,50] показано: если *элементарные* события *непрерывно заполняют* отрезок, то *вычисляются вероятности только сложных* событий, определенных на отрезках (сколь угодно малых, но конечных) с *бесконечным* (несчетным) множеством *элементарных* событий.

п.2. Учитывая условие {A.1}, обратимся *непосредственно*[13] к *закону* распределения $P(x) = \mathbf{P}(a \le x < x_{dan})$ (рис.2) случайной *непрерывной* величины Φ, определенной на *бесконечном* (несчетном) множестве *элементарных* событий, принадлежащих точкам отрезка $0 < a \le x \le b$. Рассмотрим *равные* отрезки $[a \le x_1,\ x_2 = x_1 + c \le b]$ и $[a \le x_3,\ x_4 = x_3 + c \le b]$ $(c > 0)$, *произвольно расположенные* в области определения. Имеем одинаковые события, вероятности которых равны $\mathbf{P}(x_1 \le x < x_2) = P_2 - P_1$ и $\mathbf{P}(x_3 \le x < x_4) = P_4 - P_3$, где $P_1 = \mathbf{P}(a \le x < x_1)$, $P_2 = \mathbf{P}(a \le x < x_2)$, $P_3 = \mathbf{P}(a \le x < x_3)$ и $P_4 = \mathbf{P}(a \le x < x_4)$. Если *для любых значений* $x_1 \ge a$, $x_2 \ge a$ *и* $c > 0$ *выполняется равенство* $P_2 - P_1 = P_4 - P_3$, *то закон распределения будет равновероятным*, а плотность распределения – *постоянна* $p(x) = const$. Видимо, можно сказать: *значения вероятностей или производной «равномерны»*, но *будет ли это лучше и понятнее того, что сказано в предыдущем предложении*? Конечно же, нет.

> **{A.6}** *Закон* распределения *непрерывной* случайной величины будет *равновероятным* при условии, что *вероятности 2-х любых* событий, которые определяются *произвольно расположенными равными отрезками, одинаковы.*

Таким образом, правильно говорить о *равновероятном законе* распределения *дискретной* или *непрерывной* случайной величины. Поэтому больше эти термины подчеркивать не будем.

п.3. На оси P *расстояние до произвольного* значения вероятности *пропорционально* значению вероятности $P_j = \mathbf{P}(a \le x < x_j)$, а *расстояние* между *двумя любыми* значениями вероятностей *пропорциональны разности* этих значений $P_k - P_j = \mathbf{P}(a \le x < x_k) - \mathbf{P}(a \le x < x_j)$ $(x_k > x_j)$ **{6}**. Следовательно, вероятности *соответствуют 2-му способу представления опыта* в виде случайной величины [29,46-47] и лежат на прямой $P = r$ (рис. 2, система координат $r0P'$ повернута против часовой стрелки). В [19,50] (вывод Q.1) показано, что в этом случае вероятность события зависит только от длины отрезка, а вероятности двух событий, *определенных на равных* отрезках *имеют равные* вероятности. Следовательно, на отрезке $0 \le r \le 1$ имеем *равновероятное непрерывное* распределение. Отношение (свойства 1-2 [29,46-47]) $\Delta P(r)/\Delta r = 1$ (где $\Delta P(r) = P(r_{j+1}) - P(r_j)$, $\Delta r = r_{j+1} - r_j$) является дискретным аналогом производной. Оно справедливо при *любых сколь угодно близких* последовательных значениях r_j и r_{j+1}: отношение *сохраняется и при предельном* переходе $\lim\limits_{\Delta r \to 0} \Delta P(r)/\Delta r = dP(r)/dx = 1$, что отвечает *равновероятному* распределению на оси r.

---

[13] Плотность распределения не определяет вероятностей

п.4. Функция $P=P(x)$ *одновременно* является *монотонно возрастающей формулой преобразования*: она *отображает непрерывное* множество *элементарных* событий на отрезке $a \le x \le b$ *в непрерывное* множество вероятностей $P(r)$ на отрезке $0 \le r \le 1$.

Пусть $x=\phi(r)$ – функция, обратная функции $P(x)$. Тогда применяя *1-й вариант преобразования* (формулы (1), стр.9), получим *закон* распределения $P(r)$ на *ограниченном* множестве R ($0 \le r \le 1$) *действительных* чисел с равномерным распределением вероятностей.

С другой стороны, в п.3 показано, что на основе равенств {1} определяется *равновероятное* распределение. Это определяет гораздо больший интерес «к обратному ходу» (п.2.b примера 2, стр.10). Пусть на плоскости $X0Y$ задана *монотонно возрастающая* функция $y=f(x)$. Тогда *равновероятное непрерывное* распределение, заданное на отрезке $0 \le r \le 1$ оси Y, с применением функции $x=\phi(y)$, обратной к функции $f(x)$, преобразуется в *закон* распределения $P(x) = \mathbf{P}(a \le x < x_{dan})$ на отрезке $a \le x \le b$ оси X, который соответствует *формуле преобразования* $y=f(x)$.

Таким образом, показано:

> **W.7. *Законы* распределения и *вероятностные* функции, полученные *1-м и 2-м способами представления опыта* в виде случайной величины, связаны между собой *нелинейным преобразованием.***

Вывод следует из замечания 3.2 (стр.9) и выводов W.5- W.6 (стр.11).

> **W.8. На основе *равновероятного непрерывного* распределения, *определенного на отрезке* $0 \le r \le 1$ числовой оси Y, применяя преобразование, можно получить *любой закон* распределения на произвольном отрезке $a \le x \le b$ оси X. Именно это и утверждалось (только голословно) в работе [19,50].**

При рассмотрении *представления одномерных опытов* в виде *одномерных* случайных величин мы положили (**положение {A.15} [29,48]**), что независимо от способа построения, *основными* являются *вероятностные* функции и *законы* распределения *случайных* величин, *определенные* на множествах *действительных* чисел, *ограниченных отрезками* $0 \le r \le 1$. Из проведенного анализа, выводов **W.6** и **W.7** следует:

> **W.9. <u>Основными</u> являются *вероятностная* функции и *закон* распределения (как дискретной, так и непрерывной) случайной величины, которые получены *1-м способом представления видов испытания* (элементарным событиям соответствуют точки равномерно расположенные на отрезке [29,46]) в виде случайных величин и *определенные* на множествах *действительных* чисел R или X, *ограниченных* отрезками $0 \le r \le 1$ или $x_H \le x \le x_K$. На этих отрезках *законы* распределения *связаны* между собой *линейным преобразованием.***

1.2. *Немонотонная формула преобразования*

Сначала о «ляпсусе», допущенном в работе [10,49] при рассмотрении *преобразования законов* распределения. Дело в том, что хотя мы говорили о *преобразовании законов* распределения, но использовали *только преобразование плотностей.* При этом даже не попытались посмотреть, а какой «закон» получается при *немонотонной формуле преобразования.* Восполним этот пробел, исходя из *первого варианта преобразования.*

***Примеры* 4. п.1.** **На рис.3 даны условные изображения: 1. *Закона* распределения $P_{\Phi}(x) = \mathbf{P}_{\Phi}(a \le x < x_{\mathbf{dan}}) = \int_a^{x_{\mathbf{dan}}} p_{\Phi}(x)\,dx$ случайной величины Φ (утолщенная сплошная кривая) с *симметричной* функций *плотности* $p_\Phi(x)$ $(a \le x < b)$ (максимум $p_{\max} = p_\Phi(x_1)$ в точке x_1). 2. *Симметричной* (немонотонной) *формулы преобразования* $y = f(x)$ (максимум $y_{\max} = f(x_2)$ в точке x_2, тонкая сплошная кривая).На рис.3.А максимумы функций совпадают, а на рис.3.В функция $p_{\Phi}(x)$ смещена влево, так что $a = 0$.**

Если два значения функции $p_\Phi(x)$ имеют равные проекции на ось Y, то при преобразовании эти значения суммируются[14].

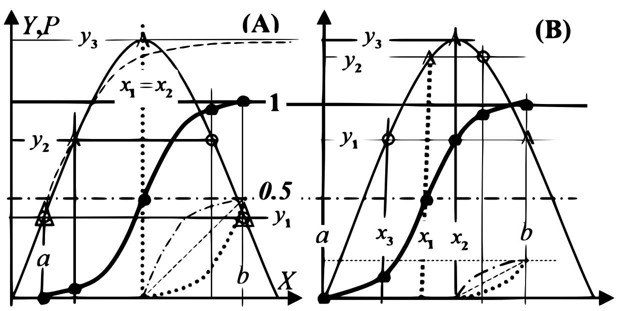

Рис. 3.

***Преобразованная* функция *плотности* $p_\Psi(y)$ при значении x_2 (так как производная $x' = \phi'(y)$ при приближении к максимуму $y_{\max} = f(x_2)$ стремится к бесконечности) имеет ветвь, уходящую на бесконечность.** **Более подробно об этом – в работе [13]. На рис.3.А дано условное изображение *преобразованной* функции *плотности* $p_\Psi(y)$ (штриховая кривая), заимствованное из работы [13,269]. Кстати, это единственная работа, в которой применяется геометрическое представление при преобразованиях[15], несмотря на другое название. В работе [10,53] было сделано некоторое исключение, но это не дало даже малейшего повода для сомнений в правиле. Одно маленькое замечание к изображению *плотности* $p_\Psi(y)$ на рис.3.А: оно, чтобы «не занимало место», соответствует левосторонней системе координат (ось P в направлении оси X). Для приведения к правосторонней системе ее следует направить влево и зеркально отобразить изображение относительно оси Y. О тонких штрихпунктирных и пунктирных кривых на рис.3, поговорим в замечании 4 после рассмотрения примеров.**

Однако наша *основная цель – преобразование закона*, а не *плотности* распределения. Отметим: если *закон* распределения *сместить* так, что $b \le x_2$ или $a \ge x_2$, то будем «иметь дело» с *монотонно возрастающей* или *убывающей* (левая и правая ветвь кривой $y = f(x)$ относительно максимума) *формулой преобразования* соответственно.

На кривых $P_\Phi(x)$ отмечены точки с вероятностями 0, 0.5, и 1. На рис.3.А они дополнены точкой между координатами a и x_1, а на рис.3.В – точкой пересечения кривой $P_\Phi(x)$ и прямой $y = x_2$. Этого вполне достаточно для нашего анализа. Построение одинаково: через точку на кривой $P_{\Phi}(x)$ и ее координатой на оси X проводится вертикальная прямая линия до пересечения с кривой $y = f(x)$ (отмечена треугольником). Через точки пересечения проводятся горизонтальные прямые $x = y_1$, $x = y_2$ и $x = y_3$ линии до пересечения с противоположной ветвью кривой $y = f(x)$ (обозначены окружностями). Затем через эти точки проводятся вертикальные линии и определяются точки пересечения с кривой $P_{\Phi}(x)$ и их координаты. На рис.3.А точки с координатами $x = a$ и $x = b$

---

[14]**В соответствии с правилом суммы действительных функций (приложение I [29,95]). Как это осуществляется в принятой теории – в замечаниях 4 стр.16**

[15] **Рассуждения и рисунки, приведенные в этой работе, показались нам убедительными, но это впечатление оказалось ошибочным**

одновременно являются «начальными и конечными» в этом построении, поэтому они обозначены окружностью в треугольнике.

При *равенстве* проекций на ось Y двух значений функции $P_\Phi(x)$ эти значения *суммируются* (приложение I [29,95]). *И что же «мы имеем» в результате суммирования*? На рис.3.А получаем *монотонно <u>убывающий</u> закон* распределения случайной величины Ψ, который изменяется от единицы (на линии $x = y_1$) до значения вероятности 0.5 (на линии $x = y_3$). Это *противоречит плотности* распределения (производной), которая на том же отрезке является *монотонно <u>возрастающей</u>* функцией. Да и значение вероятности 0.5 не «вызывает доверия». Но это «только цветочки, ягодки» – на рис.3.В.

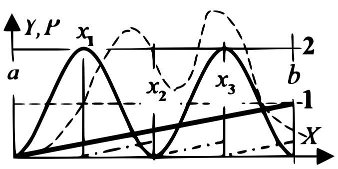

Рис. 4.

До линии $x = y_1$ он «ведет себя, как положено»: – *возрастает* от нуля до некоторого значения $P_\Phi(x_3)$ в соответствии с *возрастанием* производной (при значении x_2 плотность, как и в 1-ом случае, имеет ветвь, уходящую на бесконечность). Однако «благородное поведение» *закона* распределения на этом заканчивается: при переходе через линию $x = y_1$ он изменяется скачком (имеет точку разрыва) – от значения $P_\Phi(x_3)$ до значения $P_\Phi(b) + P_\Phi(x_3) > 1$ большего единицы (этого принципиально быть не может). Затем *монотонно убывает* до некоторого значения $P_\Phi(x_2)$ в точке x_2 максимума функции $y = f(x)$. Таким образом, поведение закона распределения абсолютно не соответствует его производной.

<u>п.2.</u> *А что будет, если экстремум не один, а несколько*? Как на рис.4, на котором утолщенная сплошная кривая определяется функцией $y = 1 - \cos(x)(0 \leq x \leq 4\pi)$ и имеет два максимума и три минимума. При равновероятном законе (утолщенная сплошная прямая) распределения $P_\Phi(x) = x/(4\pi)$ легко определяются значения вероятностей закона распределения $P_\Psi(y)$ в 3-х точках: $P_\Psi(y = 0) = 1.5$, $P_\Psi(y = 1) = 1.875$ и $P_\Psi(y = 2) = 1$. При этом значение в точке $y = 1$ не является максимальным для функции $P_\Psi(y)$: ее максимум будет на интервале $1 < y < 2$. *А если функция $y = f(x)$ будет определяться несимметричной тонкой штриховой кривой, изображенной на рис.4*? Мы на этом не останавливаемся: ответ, по крайней мере, в качественном отношении, легко получить из тех же рассуждений.

«Веселенькие законы» получаются при *немонотонной формуле преобразования* - что тут еще скажешь. Можно, конечно возразить, что наши рассуждения в примерах 4 неверны, ибо в конкретном примере [13,268], формулы дают 1. По этому поводу отметим следующее.

***<u>Замечания</u>* <u>4</u>. <u>1.</u> Можно показать, что в и общем случае формулы, приводимые в некоторых работах (в частности – в [13,267-268]), дают 1. По сути, при немонотонной функции $y = f(x)$ $(a \leq x \leq b)$ закон $P_\Phi(x)$ представляется на двух (рис.3) и более (в примере 4.2 – на 4-х) интервалах: $P_\Phi^{(1)}(x) = \int_a^x p_\Phi(x)dx$ $(a \leq x < x_2)$ {7} и $P_\Phi^{(2)})(x) = \int_{x_2}^x p_\Phi(x)dx$ $(x_2 \leq x < b)$ {7*}. Закон распределения величины Ψ определяется равенствами $P_\Psi^{(k)}(y) = P_\Phi^{(k)}\{x = \phi(y)\}$ (k=1,2) {8}. Очевидно, что при значениях $x = x_2$ и $x = b$ (рис.3.А) интегралы в формулах {7}-{7*} равны $P_\Phi^{(1)}(x_2) = 1 - P_\Phi^{(2)}(b)$ (следовательно, и значения функций {8}), а их сумма – 1. Применяя этот подход, легко показать, что *этот же результат* получится как в примере 4.2, так и в *общем случае немонотонной формулы преобразования и произвольном законе* распределения.**

<u>2.</u> *Что имеем при таком представлении*? Функция {7} возрастает от нуля до значения $P_\Phi^{(1)}(x_2)$, а функция {7*} – от нуля до значения $1 - P_\Phi^{(2)}(x_2)$ (рис.3. тон-

кие штрихпунктирные кривые, рис.4 – штрихпунктирные прямые[16], т.е. в точке $x = x_2$ закон убывает скачком от значения $P_\Phi^{(1)}(x_2)$ до нуля. Соответственно, в примере 4.2 закон убывает скачком от значения $P = 0.25$ до нуля в 3-х точках $x = x_1, x_2, x_3$. Очевидно, что это противоречит пониманию закона как монотонно возрастающей функции. Какие значения функций суммируются, легко определяется по пунктирным кривым[17] (рис.3), которые переносятся (в соответствии с подходом, приведенным в [13,268]) в начало координат. Из рисунка 3.В следует: в общем случае немонотонной формулы преобразования подобные скачки будет иметь и преобразованный закон распределения.

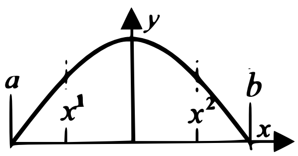

Рис. 5.

<u>3.</u> Теперь рассмотрим пример[18], который приведен в работе [13,268]. С применением функции $y = \cos x$ (рис.5), заданной на отрезке $a = -\pi/\mathbf{2} \le x \le \pi/2 = b$, преобразуется равновероятный закон $P_\Phi(x)$ распределения с плотностью $p_\Phi(x) = 1/\pi$.

В соответствии с подходом, который дан в работе, записаны законы распределения для возрастающей и убывающей ветвей функции $y = \cos x$ соответственно: $P_\Psi(y_1) = \int_a^{x^1} p_\Phi(x)dx$ и $P_\Psi(y_2) = \int_{x^2}^b p_\Phi(x)dx$, где $a \le x^1 = -\arccos y < 0$, $0 \le x^2 = arccosy < b$ и $0 \le y_1, y_2 < 1$. С применением равенства $p_\Psi(y = P'_\Psi(y)$, вычисляется плотность распределения, которая одинакова для обеих ветвей функции $y = \cos x$: $p_\Psi(y_1) = p_\Psi(y_2) = 1/\{\pi(1-y^2)^{1/2}\}$. Затем законы и плотности распределения обеих ветвей суммируются (в [13] сразу записывается их сумма). Обратим внимание на следующее. Функция $P_\Psi(y_1)$ монотонно возрастает от нуля до значения 0.5, а функция $P_\Psi(y_2)$ монотонно убывает от значения 0.5 до нуля (замечание к формулам (1-2), стр.9). Их сумма дает закон распределения, изменяющийся на интервале $0 \le y < 1$ от нуля до единицы.

Все «было бы прекрасно, если бы не одна нестыковочка»: *плотность* распределения $p_\Psi(y_2) = 1/\{\pi(1-y^2)^{1/2}\}$ является *положительной* функцией, следовательно, *функция* $P_\Psi(y_2)$, *производная* от которой и есть *плотность*, *может быть только возрастающей функцией.*

Т.е., должны *суммироваться две возрастающие* функции. Что получится, если в *сумме* $P_\Psi(y_1) + P_\Psi(y_2)$ использовать функцию $P_\Psi(y_2)$, *возрастающую* от значения 0.5 до единицы, очевидно. Таким образом, мы «неизбежно» приходим к результату, полученному в примере 4 (стр.15).

Таким образом, из анализа, данного в замечании, следует:

> **W.10. При существующей трактовке *1-го варианта преобразования с немонотонной* функцией $y = f(x)$ моделируется некоторый *разрывный закон* распределения, *не имеющий отношения к реальному заданному закону и противоречащий принятому* пониманию *закона* распределения.**

Фактически, при принятой трактовке, *преобразование моделирует равенство сумм* вероятностей $P_\Phi^{(1)}(x_2) + P_\Phi^{(2)}(b)$ единице на интервалах $a \le x < x_2$ и $x_2 \le x < b$ (**рис.3, стр.15**), однако *цель преобразования* другая:

---

[16] **В примере 4.2 на каждом отрезке функция $P_\Phi^{(k)}(x)$ (k=1,2,3,4) изменяется от нуля до значения $P = 0.25$**

[17] **Вообще говоря, кривые отличаются от тех, которые изображены на рисунке: об этом в примере 5 (стр.18)**

[18] **Применяются обозначения, принятые в настоящей работе. Рис.5 практически повторяет рис.12.3.2 в [13]: нет только прямой АВ. Сравните его с примером 4.2**

{A.7} При *преобразовании* должен *моделироваться непрерывный закон* распределения, а *не равенство единице суммы* $P_{\Phi}^{(1)}(x_2)+P_{\Phi}^{(2)}(b)$ значений функции $P_{\Phi}(x)$ на участках. Для его *моделирования следует представить* функцию $P_{\Phi}(x)$ на интервале $x_2 \le x < b$ в виде $P_{\Phi}^{(2)}(x) = P_{\Phi}^{(1)}(x_2) + \int_{x_2}^{x} p_{\Phi}(x)\,dx$. Эта функция является *непрерывным продолжением* функции $P_{\Phi}^{(1)}(x)$ и *не суммируется* с ней.

Примеры 4 правильно отражают моделирование непрерывного закона $P_{\Phi}(x)$ распределения при принятой трактовке 1-го варианта преобразования (что очевидно из рисунков) и показывают, что реально получается в результате преобразования.

Напомним: *преобразование одномерных* распределений (**под названием «функция Y=U(X) случайного аргумента X»**) дан только в работе[19] [13,263] (**при различных видах формулы преобразования**). Следуя работе [13], рассмотрим *первый* вариант отдельно при *монотонно возрастающей* и *убывающей формуле преобразования.*

Пример 5. Пусть на отрезке $0 \le x \le b$ определен *закон* $P_{\Phi}(x) = \mathbf{P_{\Phi}}(0 \le x < x_{dan})$ распределения величины Φ с *несимметричной* функцией плотности $p_{\Phi}(x)$ (рис.6, сплошные утолщенная и тонкая кривые соответственно). Рассмотрим простейший случай – линейные *функции* $y = x$ (рис.6, штриховая прямая) и $y = -x + b$ (штрихпунктирная прямая).

п.1. Если $y=x$, то *преобразование* приводит к повороту исходной кривой $P_{\Phi}(x)$ вокруг начала координат на угол $\pi/2$ влево: точка (0,0) остается на месте, точка (b,0) перемещается в точку (0,b), а точка (b,1) – в точку (-1,b).

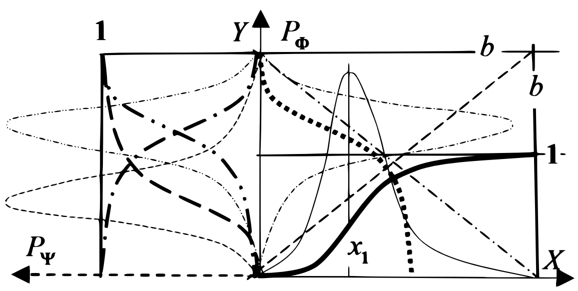

Рис. 6.

В результате получим на отрезке $0 \le y \le b$ оси Y *закон* распределения величины Ψ (утолщенная штриховая кривая), *равный исходному закону* $P_{\Psi}(y) = P_{\Phi}(x)$. Чтобы закон $P_{\Psi}(y)$ был определен *в правосторонней системе* координат $Y0P_{\Psi}$, ось P_{Ψ} следует направить противоположно оси X (штриховая стрелка). Ему *соответствует плотность* $p_{\Psi}(y) = p_{\Phi}(x)$ (как производная от функции $P_{\Psi}(y)$: обозначена тонкой штриховой кривой).

п.2. Положим $y = -x + b$. Тогда *преобразование* приводит к смещению исходной кривой $P_{\Phi}(x)$ параллельно оси Y в точку ($x = 0, y = b$) и повороту вокруг этой точки на угол $\pi/2$ вправо: точка (0,0) перемещается в точку (0,b), точка (b,0) перемещается в точку (0,0), а точка (b,1) – в точку (-1,0).

В результате на отрезке $0 \le y \le b$ оси Y получим *закон* $P_{\Psi_1}(y)$ распределения величины Ψ_1, отличающийся от исходного закона. Во-первых. Он находится в отрицательной области значений (пунктирная кривая) оси P_{Ψ}: если по

---

[19] **В этой работе, в отличие от многих других, геометрическое представление применяется часто. Работа в большей степени опирается на классическое определение вероятности, проста, понятна и наглядна. По нашему мнению, это лучший учебник для начинающих изучать теорию вероятностей**

сути, то он расположен в левосторонней системе $Y0P'_{\Psi}$ координат (направление оси P'_{Ψ} и оси X совпадают). Переход к правосторонней системе обеспечивает положительность функции $P_{\Psi_1}(y)$ (утолщенная штрихпунктирная кривая). Во-вторых. *Закон* распределения величины Ψ_1 *монотонно убывает* от значения $P_{\Psi_1}(y = 0) = 1$ до значения $P_{\Psi_1}(y = b) = 0$. Оба отмеченных момента связаны с тем, что закон распределения $P_{\Psi_1}(y) = P(b \le y < y_{dan})$ определяется интегрированием плотности $p_{\Phi}(x)$ в обратном направлении[20] и следуют из свойства интеграла по направленному отрезку [24,108]. Соответственно, *плотность* (как производная от функции $P_{\Psi_1}(y)$) является *отрицательной* функцией (тонкая штрихпунктирная кривая), но *именно она правильно отражает* поведение *монотонно убывающего закона* распределения. Используя равенство $P_{\Psi_2}(y) = 1 - P_{\Psi_1}(y)$, получаем возрастающий закон с положительной плотностью (утолщенная и тонкая штриховые кривые с двойным пунктиром соответственно). Из рисунка видно, что закон и плотность $p_{\Psi_2}(y)$ отличаются от исходного закона (и плотности) распределения.

п.3. Можно провести подобный анализ при произвольной *монотонно возрастающей* $y = f(x)(a \le x \le b)$ или *убывающей* $f_1(x) = -f(x) + f(a) + f(b)$ функции. Конечно же, как-то изменятся преобразованные законы и плотности распределения, но это не оказывает никакого *влияния* на суть *преобразования*: – *результаты преобразования будут подобны тем*, которые отмечены в п.п.1-2.

В работе [13] *преобразования* с видами *монотонных* функций рассматриваются достаточно подробно: мы обратили внимание на некоторые *нюансы преобразования*, которых нет в работе. Вообще говоря, проведенный анализ – это *первая* попытка дать *геометрическую* интерпретацию *преобразования* при *возрастающей* (**п.1**) и *убывающей* (**п.2**) функции и на этой основе построить правильное *преобразование* с *немонотонной* функцией $y = f(x)$, а затем пояснить, почему при *немонотонной* функции, выражение для плотности следует записывать *для каждой из ветвей раздельно* (**примечание 3 [10.53]**). Осталось выяснить: *каков результат, полученный на основе исследования, выполненного в примере 5*?

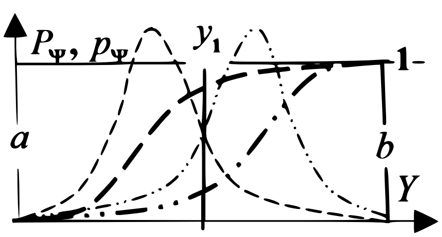

Рис. 7.

Пример 6. **Применив формулу $P_{\Psi_2}(y)\ 1 - P_{\Psi_1}(y)$ (п.2 примера 5,), мы получили *возрастающий* закон распределения (рис.6-7 стр.18-19, утолщенная штриховая кривая с двойным пунктиром). Не геометрия, но эффект подобен тому, который дает *абсолютное значение производной* в формулах (1-2, стр.9). На «первый взгляд» (как показывает опыт – и «последующие взгляды» тоже) – все сделано правильно. Однако. При убывающей формуле преобразования закон и плотность распределения отличались от того, что получено в существующей теории, поэтому желание посмотреть, что дает преобра-**

[20] **В работе [13] этот момент не учитывается, что приводит к противоречию, отмеченному в замечаниях к формулам (1-2) и замечании 4.3 (стр.9, 16). Однако, если закон распределения определять по положительной плотности $p_{\Psi}(y) = p_{\Phi}\{\phi(y)\} \cdot |\phi'(y)|$, то при монотонно убывающей функции $y = f(x)$ получим для него правильное выражение: с точностью до его местоположения, о чем пойдет разговор в примере 8 (стр.23)**

зование с немонотонной формулой преобразования было вполне естественным. Мы не приводим анализа, а только скажем: «чехарда похлеще», чем в примерах 4, но именно эта «чехарда» и привела к проверке существующего подхода (т.е. появлению примеров 4). Отметим только некоторые исходные моменты: возможно, они позволят Читателю лучше понять только что сказанное (по крайней мере, мы думаем, что это так).
1. В качестве немонотонной формулы преобразования рассмотрена ломаная линия, образованная прямыми $y=x$ и $y=-x+b$, точка пересечения которых лежит на линии $y_1=x_1=b/2$. Т.е. применена формула преобразования, производная от которой имеет разрыв 1-го рода[21] (скачок). 2. Для наглядности на рис.7 даны изображения преобразованных функций $P_\Psi(y)$, $p_\Psi(y)$ и $P_{\Psi_1}(y)$, $p_{\Psi_1}(y)$, использованные при анализе и полученные в п.1 и п.2: обозначения кривых соответствуют рис.5. 3. Суммируются значения функций $P_\Psi(y)$, $p_\Psi(y)$, лежащих слева от значения y_1 и значения функций $P_{\Psi_1}(y)$, $p_{\Psi_1}(y)$, лежащих справа от значения y_1.

Мы искренне надеялись получить правильное решение, но получили то, что совсем не ожидалось: в общем-то, «красиво вляпались» или, говоря словами известного политика: **«Хотели как лучше, а получилось как всегда».**

Несмотря на, казалось бы, «полное фиаско» идеи, интуиция[22] «подсказывала»: ***функции закона* и *плотности* распределения, полученные по 1-му варианту, *должны быть аналогичны* (с точностью до линейного преобразования) *функциям*, полученным по 2-му варианту *преобразования.***

Выше шел разговор только о *1-м варианте преобразования*, однако и со *2-м вариантом* тоже не все гладко, что связано со следующим:

> **{А.8} Анализ *преобразования законов* распределения *проводился на основе геометрических построений*, а *они не зависят от конкретного варианта преобразования законов* распределения.**

Учитывая положение {А.8}, анализ, данный в примерах 4 (**стр.15**), 5 (**стр.18**), 6 (**стр.19**) и замечаниях 4 (**стр.16**), справедлив и для *2-го варианта преобразования*. Рассмотрим *отличие плотностей* распределения, полученных по 1-му (**в его существующей трактовке**) и 2-му варианту *преобразования*, которое следует из анализа, данного в работе [10,53].

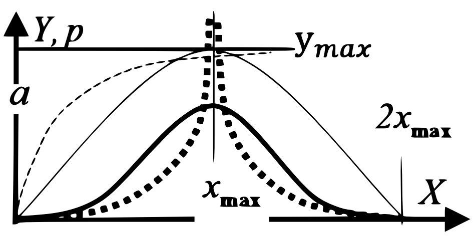

Рис. 8.

***<u>Пример</u>* <u>7</u>. На рис.8 (он почти повторяет рис.8 в [10,53]) даны условные изображения. Симметричной функции плотности (сплошная утолщенная кривая) и формулы преобразования (сплошная тонкая кривая) $y=f(x)$, $y_{\max}=f(x_{max})$. Преобразованных функций плотности, полученных с применением: 1-го варианта в соответствии с принятой теорией (тонкая штриховая кривая); 2-го варианта преобразования (утолщенные пунктирные кривые). Различие плотностей по 2-м вариантам очевидно из рисунка: о *законе* распределения, следующего из преобразования, уже все сказано (примеры 4 стр.15).**

---

[21] **В этом примере также не используется понятие абсолютно непрерывной функции**

[22] **Вообще-то, это была убежденность, основанная на том, что оба варианта определяют один и тот же закон распределения $P_\Psi(y)$ на отрезке $f(a) \le y < f(b)$ (замечание 3.2, стр.9) и результатах, полученных по 2-му варианту**

{A.9} ***Основное*** **отличие 2-го варианта от существующей трактовки 1-го:** ***формула преобразования*** **определяет** ***взаимное однозначное соответствие исходной и преобразованной*** **функций** ***плотности*** **(и законов) распределения.**

Утверждение, в принципе, правильное, но оно *основано* только на *исследовании* функции *плотности* и *представлении преобразованного* распределения на оси X. *Преобразование закона* распределения не рассматривалось[23]. Отсюда естественно следует: **переход** ***от представления*** **распределения на оси** X **к его** ***представлению*** **на оси** Y**, тоже** ***требует пояснений.***

Т.е. мы *должны показать, как осуществляется данный переход*. Если ошибки, связанные с переходом к *непрерывному* распределению и понятием «равномерного» *непрерывного* распределения можно было «свалить на въевшиеся стереотипы» [19.52], то «ляпсус» с *немонотонной формулой преобразования* – «недосмотр», который трудно объяснить: кто его знает почему, но он был «прохлопан».

Замечания 5. **1. В работе [13.267] рисунки и формулы (построенные для** ***немонотонной формулы преобразования*****) указывали на то, что на** ***возрастающей*** **и** ***убывающей*** **ветви (рис.4)** ***формулы преобразования*** **значения вероятностей** ***закона*** **распределения,** ***суммируются.*** **Достаточно было проверить, какой реально получается** ***закон*** **распределения (что выполнено в примере 4, стр.14). Однако, по этому поводу «никакая мысль даже не шевельнулась» при написании работы [10].** **2. Никакой «реакции» не вызвало** ***несоответствие*** **плотностей, полученных по** ***1-му и 2-му*** **вариантам преобразования (рис.8 [10,53]: к этому рисунку и результату, следующему из представления плотности на оси** X**, мы обратимся попозже) с применением** ***немонотонной формулы преобразования,*** **а соответствие должно «было бы быть». Но «прошли мимо, даже не заметив» и не попытавшись пояснить, как получить распределение на оси** Y**.**

3. Далее было отмечено, что (примечание 3 [10.53]): «При немонотонной функции $y = v(x)$**, выражение для плотности распределения лучше записывать раздельно для каждой из ветвей возрастания и убывания функции** $y = v(x)$**», но, к сожалению, только потому, что «. . . невозможно обратное преобразование . . . » и «. . . является важным при решении определенного круга задач».** **4. По сути, эта «оговорка» была сделана на основе результата применения 2-го варианта преобразования и только для того, чтобы показать возможность правильного построения точных распределений выборочной дисперсии и моментов более высокого порядка, определяемых по** N **испытаниям, при известном (произвольном) законе распределения случайной величины X (п.4.6.2 [10.82]). Однако о том, что следует дать пояснение тоже «не подумалось»** **5. Не говоря уж о том, чтобы подумать, а** ***что означает абсолютное значение производной*** $|\phi'(y)|$**?**

По большому счету, все было готово к тому, чтобы преобразование выполнить так, как «ему положено быть». Можно сказать, что «интуиция не подвела»: решение было найдено, и существенно проще, чем казалось после анализа, данного в примере 4 (**стр.15**): только времени на это

[23] **Подразумевалось, что он определяется интегрированием функции плотности**

«ушло немало». Как «засияла мысль» о *зеркальной симметрии* сказать сложно[24].

***Замечание* 6. Зеркальная симметрия применяется в существующей теории. Знак плотности определяется знаком производной $\phi'(y)$. А*бсолютное* значение $|\phi'(y)|$ производной в формулах (1.b-2. b, стр.9) приводит к зеркальному отображению плотности $p_{\Psi_1}(y)$ в область *положительных* значений. При этом, вообще говоря, решается и 2-я задача: *закон* распределения становится *монотонно возрастающей* функцией.**

Обратить внимание на самую простую «вещь» – **в «зеркале отражается» *не плотность*, а *производная* $f'(x)$ от функции** $f(x)$ – оказалось не менее сложным. Но как «только это случилось, тут же последовало»:

***Замечания* 7. 1. При *зеркальной симметрии* необходимо определить «место», в котором находится «зеркало, отражающее» плоскую кривую (расстояния от точек кривой до прямой линии, через которую проходит плоскость, перпендикулярная плоскости кривой) и где находится «отраженная» кривая. 2. В геометрии не существует ни функций, ни обратных функций, ни, тем более, их производных: это «территория» алгебры (теории действительных функций действительного аргумента). Следовательно, производная «геометрических задач не решает». 3. В геометрии отображаются прямые и кривые линии, их отрезки, фигуры и т.п. 4. Связь между геометрией и алгеброй определяется схемой Ферми-Декарта (пример 3 [29,21]).**

Из замечаний следует:

> **{A.10} *Абсолютное значение производной* $|f'(x)|$ определяет *зеркальное отображение* кривой, которая *определяется убывающей* функцией $y=f(x)$. Только в этом случае абсолютное значение изменяет знак производной от функции.**

Была также просмотрена работа[25] [21,102-114], ибо сразу вспомнилось, что в ней *подробно рассмотрены* (**с применением геометрического представления и связи между алгеброй и геометрией**) *элементарные* функции, а также *даны подробные* пояснения к понятию *обратной* функции.

Обратим также внимание на некоторые свойства *однозначной немонотонной* функции, очевидные из геометрии, но часто «упускаемые из виду» при алгебраической записи.

***Замечания* 8. 1. *Немонотонная* функция $y=f(x)$ (рис.9, утолщенная кривая) имеет один или несколько *экстремумов*: *возрастающая ветвь* за-**

---

[24] **Вроде бы симметрия применялась и часто, но «знание о ее простейшем виде сидело где-то слишком глубоко». Упоминание в примере 4 (стр.15) – результат замены «мутноватого» пояснения, данного до появления «догадки»**

[25] **По нашему мнению – это лучший учебник по математическому анализу**

канчивается *в точке максимума* $y_1 = f(x_1)$ функции, в которой начинается *убывающая ветвь*, продолжающаяся *до точки минимума* $y_2 = f(x_2)$, после которой функция снова становится *возрастающей* до *следующего максимума* $y_3 = f(x_3)$ и т.д. Т.е. *возрастающие* и *убывающие* ветви кривой расположены на *прилегающих* отрезках $x_{j-1} \le x \le x_j$ и $x_j \le x \le x_{j+1}$ (где j=1,2,...,N-1; $x_0 = a$, $x_N = b$, N – число ветвей) оси X.

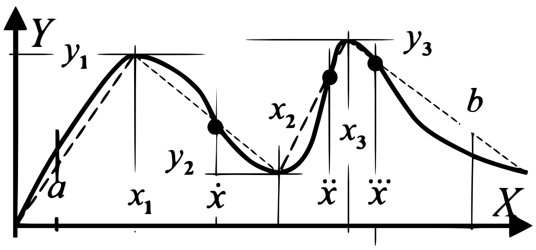

Рис. 9.

2. Отрезки прямых линий, соединяющих пары точек $[x_{j-1},\ f^{(j-1)}(\,x_{j-1})]$ и $[x_j,\ f^{j-1}(\,x_j)]$, образуют некоторую *ломаную* (штриховая линия) линию и определяют положение ветвей на плоскости $X0Y$.

3. Каждая ветвь кривой, находящаяся между *максимумом* и *минимумом* (или минимумом и максимумом) имеет, по крайней мере, *одну точку перегиба* (непрерывность производной в точках экстремумов определяет возможность появления других точек перегиба только парами): начальная или конечная ветвь могут их и не иметь (рис.9, точкам перегиба соответствуют координаты $x = \dot{x}, \ddot{x}, \dddot{x}$).

Теперь, следуя работе [21], т.е. *совмещая геометрическое построение* и *алгебру*, будем исправлять «прохлопанный недосмотр».

Последующий анализ *преобразования законов* распределения будем проводить с учетом замечаний 7-8 (**стр.22, 23**) и положений **{A.8}** (**стр.20**), **{A.10}** (**стр.22**), не обращаясь к *конкретному варианту*. Сначала рассмотрим преобразование с *монотонно возрастающей* и *убывающей формулой преобразования* отдельно.

Пример 8. Рассмотрим *линейную возрастающую* функцию $y = c_0^{(1)}x$ ($c_0^{(1)} > 0$) и определим *закон* распределения $P_\Phi(x)$ на отрезке $0 = a \le x \le b$ (рис.10, ($c_0^{(1)} < 1$) сплошная утолщенная кривая). Положив $c_0^{(2)} = 1$, получим пример 5 (стр.18).

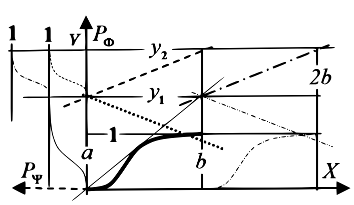

Рис. 10.

п.1. На декартовой плоскости функция определяет прямую линию, которая проходит через точки ($0,\ 0$) и ($b,\ y_1$) (тонкая сплошная прямая). Ее зеркальным отображением относительно оси X (или оси Y, что дает тот же результат) является прямая линия, проходящая через точки $(0, 0)$ и $(-b, y_1)$ (на рисунке она не изображена). Алгебраически – это убывающая функция $y = -\,c_0^{(1)}x$, соответствующая функции $y = c_0^{(1)}x$. Очевидно, что зеркальное отображение функции $y = -\,c_0^{(1)}x$ относительно оси X (или оси Y) есть функция $y = c_0^{(1)}x$.

***Преобразование* с функцией $y = c_0x$ определяет на отрезке ($0 \le y \le y_1 =$**

$c_0^{(1)}b$) ***закон*** **распределения** $P_\Psi(y)$ **(тонкая сплошная кривая) случайной величины** $\boldsymbol{\Psi}$ **в правосторонней системе координат** $Y0P_\Psi$ **(п.1 примера 5, стр.18).** **Абсолютное значение производной от функции** $y = -\ c_0^{(1)}x$ **определяет ее зеркальное отображение относительно оси** X **или** Y**), т.е. функцию** $y = c_0^{(1)}x$. **Следовательно,** ***преобразование*** **с функцией** $y = -\ c_0^{(1)}x$ **определяет точно такой же закон и на том же отрезке оси** Y**, который получен с функцией** $y = c_0^{(1)}x$.

п.2. ***А какой же тогда закон определяет преобразование с функцией*** $y = -\ c_0^{(2)}x + y_1$ $(0 < c_0^{(2)} < c_0^{(1)})$ **(пунктирная прямая),** ***с помощью которой в примере 5*** $(c_0^{(2)} = 1)$ ***мы получили преобразованный закон распределения при убывающей функции на том же отрезке*****?** **Ее зеркальное отображение относительно прямой** $x = y_1 = c_0^{(1)}b$ **(или оси** Y**, что дает тот же результат) есть функция** $y = c_0^{(2)}x + y_1$ **(штриховая прямая). Эта функция** ***преобразует закон*** **распределения** $P_\Phi(x)$ **величины** $\boldsymbol{\Phi}$ ***в закон*** **распределения** $P_{\Psi_1}(y)$ **величины** $\boldsymbol{\Psi}_1$**, который, конечно же, отличается от закона** $P_\Psi(y)$ **и расположен на другом отрезке** $(y_1 \le y \le y_2 = y_1 + c_0^{(2)}b)$ **оси** Y **(штриховая кривая).**

При значении $c_0^{(2)} = c_0^{(1)}$ **закон** $P_{\Psi_1}(y)$ **будет равен с закону** $P_{\Psi_1}(y) = P_\Psi(y)$**, однако и в этом случае он будет определен на другом отрезке. Изменение законов и плотностей распределения при преобразовании с линейной функцией** $y = \pm\ c_0x + c_1$ **показано в работе [13]. Формулы для этого случая (несмотря на неточность) написаны правильные, однако из их вывода следует, что они (п.2 примера 5, стр.18) находятся на одном отрезке оси** Y**, что, как показано выше, тоже ошибочно.**

п.3. Переместим функцию $y = -\ c_0^{(2)}x + y_1$ **параллельно оси** X **так, чтобы прямая линия проходила через точку** $(b,\ y_1)$ **(штриховая прямая с двойным пунктиром), т.е. имеем функцию** $y = -\ c_0^{(2)}(x - b) + y_1$**, и определим** ***закон*** **распределения** $P_{\Phi_1}(x) = P_\Phi(x)$ **на прилегающем отрезке** $b \le x \le 2b$.

Зеркальное отображение функции $y = -c_0^{(2)}(x - b) + y_1$ **относительно прямой** $x = y_1 = c_0^{(2)}b$ **или прямой** $y = b$**) есть функция** $y = c_0^{(2)}(x - b) + y_1$ **(штрихпунктирная прямая). При преобразовании значения** $P_\Phi(b) = 1$ **и** $P_{\Phi_1}(b) = 0$ **законов распределения** $P_\Phi(x)$ **и** $P_{\Phi_1}(x)$ **на оси** P_Ψ **оказываются в одной точке** $P_\Psi = 1$.

В результате ***преобразования*** **получим** ***закон*** **распределения** $P_{\Psi_1}(y)$ **величины** $\boldsymbol{\Psi}_1$ ***одинаковый*** **с** ***законом*** **распределения** $P_{\Psi_1}(y) = P_\Psi(y)$ **величины** $\boldsymbol{\Psi}$**, но** ***расположенный*** **на** ***другом*** **отрезке, смещенном на 1 относительно закона** P_Ψ**, (штрихпунктирная кривая) вдоль оси** P_Ψ.

Можно «продолжить процесс»: задать законы $P_{\Phi_2}(x) = P_\Phi(x)$ **и** $P_{\Phi_3}(x) = P_\Phi(x)$ **на отрезках** $2b \le x \le 3b$, $3b \le x \le 4b$ **и применить для преобразования прямые линии** $y = c_0^{(1)}(x - 2b) + y_2$, $y = -\ c_0^{(2)}(x - 3b) + y_3$ **соответственно. В результате получим законы** $P_{\Psi_2}(y) = P_\Psi(y)$**и** $P_{\Psi_3}(y) = P_{\Psi_1}(y)$**, расположенные на отрезках** $y_2 \le y \le y_3 = y_2 + y_1$ **и** $y_3 \le y \le y_4 = y_3 + c_0^{(2)}b$ **оси** Y **соответственно. Вдоль оси** P_Ψ **закон** $P_{\Psi_2}(y)$ **будет смещен на 1 относительно закона** $P_{\Psi_1}(y)$**, а закон** $P_{\Psi_3}(y)$ **– на 1 относительно закона** $P_{\Psi_2}(y)$**. Построение несложно выполнить на основе пояснений, поэтому мы не приводим их на рисунке.**

Можно задать разные законы распределения на прилегающих отрезках $b \le x \le b_1$, $b_1 \le x \le b_2$, $b_2 \le x \le b_3$ **и применить для преобразования прямые линии** $y^{(1)} = c_0^{(1)}x$, $y^{(2)} = -c_0^{(2)}(x - b_1) + y_1$, $y^{(3)} = c_0^{(3)}(x - b_2) + y_2$ **и** $y^{(4)} = -\ c_0^{(4)}(x - b_3) + y_3$ **(где** $c_0^{(3)}$, $c_0^{(4)}$ **отличаются от значений** $c_0^{1)}$, $c_0^{(2)}$, $y_2 = y_1 + c_0^{(2)}b_1$, $y_3 = y_2 + c_0^{(3)}b_2$, $y_3 = y_2 + c_0^{(3)}b_2$, $y_4 = y_3 + c_0^{(4)}b_3$**). В результате преобразования получим разные законы распределения на прилегающих отрезках** $y_0 \le y \le y_1$, $y_1 \le y \le y_2$, $y_2 \le y \le y_3$ **и** $y_3 \le y \le y_4$**. И вдоль оси** P_Ψ **каждый**

последующий закон $P_{\Psi_1}(y)$ будет смещен на 1 относительно предыдущего закона. Т.е. принципиального отличия не существует.

Из данного анализа очевидно:

> **{A.11.1}** ***Отрезки прямых*** **линий $y^{(1)}$, $y^{(2*)}$, $y^{(3)}$ и $y^{(4*)}$ (звездочка возле номера означает зеркально отображенные линии), определяемые** ***парами точек*** **$(0, 0)$ и (b, y_1); (b, y_1) и (b_1, y_2); (b_1, y_2) и (b_2, y_3); (b_2, y_3) и (b_3, y_4), образуют на плоскости** ***возрастающую ломаную*** **линию, которая и определяет** ***преобразование заданных*** **законов распределения.**

Также понятно, что наше *предположение* о том (**пример 6, стр.19**), что ***формула*** **$P_{\Psi_2}(y = 1-P_{\Psi_1}(y)$ и** ***абсолютное значение производной*** **от функции $y = f(x)$ в формулах (1-2, стр.9) дают** ***одинаковый результат*** ***<u>абсолютно неверно</u>*****.**

<u>Замечание</u> **<u>9</u>. В п.п.2-3 примера 8** ***возрастающие*** **и** ***убывающие*** **линейные функции брались с** ***разными*** **коэффициентами $c_0^{(1)}$, $c_0^{(2)}$, $c_0^{(3)}$ и $c_0^{(4)}$ с** ***конкретной целью*****:** ***моделировать*** **различное положение ветвей** ***немонотонной формулы преобразования*** **$y = f(x)$ (замечание 8.2, стр.24).**

В примере 8 каждую из прямых линий мы задавали и рассматривали (**в том числе, зеркальное отображение**) отдельно, однако:

> **{A.11.2} В отличие от примера 8 отрезки,** ***характеризующие положение*** **ветвей,** ***жестко связаны*** **с функцией и** ***между собой*** **(замечание 8.2, стр.24), поэтому** ***при зеркальном отображении*** **отрезка,** ***полностью отображается немонотонная функция*** **$y = f(x)$.**

Теперь собственно о *преобразовании* с *немонотонной формулой преобразования*, которая отвечает условию **{A.2}** (**стр.9**).

<u>Пример</u> **<u>9</u>. Исходная функция $y = f(x)$ на рис.11 (тонкая сплошная кривая) соответствуют кривой на рис.9 (стр.23). На рисунке также изображена функция равновероятного закона распределения $P_\Phi(x) = x/b$ (утолщенная сплошная прямая).**

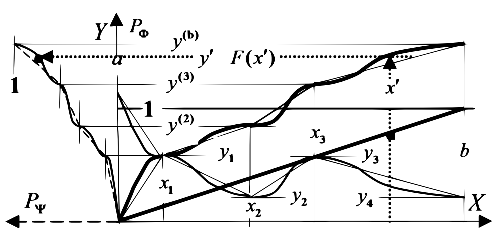

Рис. 11.

п.1. Функция $y = f(x)$ представима в виде 4-х функций, определенных на соответствующих отрезках. ***Возрастающие*** **$f_1^{(0)}(x)$ $(a = x_0 \leq x \leq x_1)$, $f_3^{(0)}(x)$ $(x_2 \leq x \leq x_3)$ {9} и** ***убывающие*** **ветви[26] $f_2^{(0)}(x) = -f(x) + y_1 + y_2$ $(x_1 \leq x \leq x_2)$, $f_4^{(0)}(x) = -f(x) + y_3 + y_4$ $(x_3 \leq x \leq x_4 = b)$ {9*}.**

п.2. На отрезке $(x_0 \leq x \leq x_1)$ функция ***возрастает*****, следовательно (п.1**

[26]**Нижний индекс обозначает номер участка, верхний (0) – исходную функцию, а далее – номер преобразованной функции $y^{(j)} = f(x^{(j)})$; (j=0,1,...,N)**

примера 5 стр.18), *преобразование* проводится с функцией $f_1^{(0)}(x) = f(x)$ (утолщенная кривая, определяющая эту ветвь функции). На оси Y значения функции $f^{(0)}(x)$ ограничены отрезком $(0 = y_0 \le y \le y_1)$.

п.3. На отрезке $(x_1 \le x \le x_2)$ функция *убывает*. *Абсолютное* значение *производной* от функции $f_2^{(0)}(x)$ *зеркально отображает* функцию $y = f(x)$ *относительно прямой* $x = y_1$: в результате *убывающие* ветви становятся *возрастающими* (утолщенная сплошная кривая), а *возрастающие – убывающими* (тонкая штриховая кривая, изображающая 1-ю ветвь) ветвями. *Преобразование определяется* функцией $f^{(1)}(x) = 2y_1 - f(x)$. Производная от функции $f^{(1)}(x)$ на данном отрезке, с учетом знака производной $f'(x)$, положительная, а на оси Y значения функции $f_2^{(1)}(x)$ ограничены отрезком $(y_1 \le y \le 2y_1 - y_2 = y^{(2)})$.

п.4. На отрезке $(x_2 \le x \le x_3)$ *возрастающая* ветвь исходной функции стала *убывающей*: *зеркально отобразив* функцию $f^{(1)}(x)$ *относительно прямой* $x = y^{(2)}$, получим *возрастающую* функцию $f^{(2)}(x) = y^{(2)} + f(x)$. На оси Y значения функции $f_3^{(2)}(x)$ ограничены отрезком $(y^{(2)} \le y \le 2y_1 - 2y_2 + y_3 = y^{(3)})$. Функция $f^{(2)}(x)$ на отрезке $(x_3 \le x \le x_4)$ снова стала *убывающей*. Ее зеркальное отображение относительно прямой $x = y^{(3)}$ определяет функцию $f^{(3)}(x) = y^{(3)} + y_3 - f(x)$. Производная от функции $f^{(3)}(x)$ на данном отрезке, с учетом знака производной $f'(x)$, положительная, а на оси Y значения функции $f_4^{(3)}(x)$ ограничены отрезком $(y^{(3)} \le y \le 2y_1 - 2y_2 + 2y_3 - y_4 = y^{(\mathbf{b})})$.

Для изображенной функции процесс построения закончен, но если функция $y = f(x)$ имеет N участков возрастания и убывания, т.е. N-1 экстремумов, то данный процесс повторяется N-1 раз.

п.5. В результате *преобразования с немонотонной* функции $y = f(x)$, учитывающего, что *абсолютное значение производной* $|f'(x)|$ (на отрезке $(x_{j-1} \le x \le x_j)$ $(j = 1, 2, ..., N)$) определяет *зеркальное отражение* функции $f(x)$ (относительно прямой $x = f(x_j)$), получена *монотонно возрастающая* функция $F(x)$, определенная (как и исходная функция) на N отрезках, соответствующих *возрастающим и убывающим ветвям* функции $y = f(x)$.

Таким образом, функция $F(x)$ определяется *общими свойствами непрерывного закона* распределения: он должен быть *монотонно возрастающей* функцией, а *производная – больше нуля*. Из *геометрии преобразования* следует: *значение вероятности закона* распределения $P_\Psi\{y' = F(x')\}$ величины Ψ *в точке с координатой* $0 = y_a \le \{y' = F(x')\} \le y^{(b)} = F(b)$ *равно значению вероятности закона* распределения $P_\Phi(x')$ (рис.11, точка на равновероятном законе распределения) величины Φ *в точке с координатой* $a = x_0 \le x' \le x_N = b$. Направление преобразования показано пунктирными стрелками.

> **{A.12.1}** Следовательно, при *немонотонной формуле преобразования* из геометрии естественным образом следует **2-й** *вариант преобразования*: в принципе, его необходимо *представить на оси* X (замечание 3.4, стр.10).

Представление на оси X рассмотрим в п.6 данного примера, в котором будет определен 1-й вариант преобразовании и показано, что 1-й и 2-й варианты определяют один и тот же закон распределения.

Преобразование **с функцией** $y = F(x)$ **определяет на** N **отрезках** $(y_{j-1} \leq y \leq y_j)$ **(j=1,2,. . . ,N) (где** $y_k = F(x_k)$ **(k=0,1,. . . ,N),** $y_0 = F(a)$, $y_N = F(b)$**)** ***непрерывную*** $P^j_\Psi\{F^{(j-1)}_j(x)\} = P^j_\Phi(x)$ **{10} функцию *закона*** $P_\Psi(y)$ **распределения случайной величины** Ψ **на оси** Y**. Точки** $y_k = F(x_k)$**, соответствующие экстремумам функции** $y = f(x)$**, определяют точки перегиба *закона* распределения** $P_\Psi(y)$**, а его производная имеет в этих точках разрывы 2-го рода.**

Для *немонотонной* функции $y = f(x)$**, изображенной на рис. 11, при равновероятном исходном законе** $P_\Phi(x)$ **(утолщенная сплошная прямая) он будет подобен функции** $F(x)$ **(кривая слева от оси Y).**

О функции *плотности* $p_\Psi(y)$ **распределения величины** Ψ**. При приближении к точкам** $y_k = F(x_k)$ **(экстремумов исходной функции** $y = f(x)$**) слева и справа функция плотности имеет ветви, уходящие на бесконечность. В точках перегиба на ветвях исходной функции (рис.9, стр.23) функция плотности имеет минимумы: если точек перегиба на ветви равно** n **(замечание 8.3, стр.23), то на этой ветви она будет иметь** n **экстремумов (минимумов).**

Замечание **к п.5.** ***Закон*** **распределения** $\mathbf{P}_\Psi(y_0 \leq y < y_{\mathbf{d}an})$ **случайной величины** Ψ **определен на интервале** $y_0 \leq y_{\mathbf{d}an} < y^{(\mathbf{b})}$**: значение** $y^{(\mathbf{b})}$ **определяется *суммой* значений *максимумов и минимумов немонотонной формулы преобразования*** $y = f(x)$**.**

п.6. В замечании 3.2 (стр.9) мы голословно утверждали: *оба варианта* определяют *один и тот же закон* распределения случайной величины Ψ**, но на *разных* осях координат. Однако! Наше «катание плоского» (функции** $y^1 = f(x)$**) в направлении оси** Y **привело к *закону* распределения {10} на оси** Y **(а не на оси** X**), *область определения* которого определяется *значениями ординат*** y_n **в точках *экстремумов формулы преобразования*. Во-вторых, в разделе 1.1 использовался *только* 1-й *вариант* преобразования, но пояснения, как он получен, до сих пор не последовало.**

1. Чтобы «узаконить» утверждение, рассмотрим *связь обратной* $x = v(y)$ **и *исходной* функций** $y = f(x)$**. Алгебра дает нам *запись обратной* функции** $x = v(y)$ **и некоторое понимание о ней (однозначность, многозначность и т.п.), но не более.**

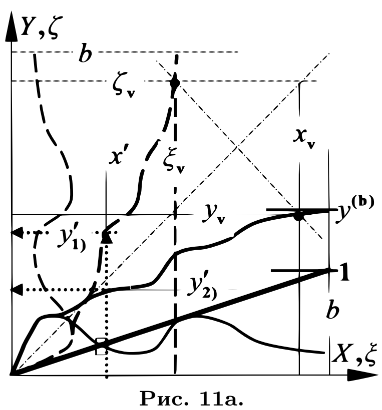

Рис. 11а.

Представление о *связи обратной и исходной* функций следует из геометрии: *зеркальное отображение* графика функции $y = f(x)$ ***относительно биссектрисы*** **(штрихпунктирная прямая на рис.11а) 1-го координатного угла [21,108]. На рис.11а изображены *формула преобразования*** $y = f(x)$**, приведенная на рис.11, и ее «монотонный аналог» (функция** $y = F(x)$**), полученный в пунктах 2-4 (сплошные кривые), а также результат их *зеркального отображения относительно биссектрисы* (штриховые кривые) и *равновероятный* закон распределения** $P_\Phi(x) = x/b$ **(утолщенная прямая).**

Отметим: функцию $x = V(y)$**, обратную к функции** $y = F(x)$**, можно получить как и функцию** $y = F(x)$**. Только надо функцию** $x = v(y)$ **«катать» в направлении оси** X**: в результате получим «монотонный аналог» (функцию** $x = V(y)$**) немонотонной функции** $x = v(y)$**.**

***Зеркальное отображение* допускает другую *интерпретацию*. Введем систему координат** $\xi 0 \zeta$**, оси которой совпадают с осями** X **и** Y **соответственно, и для каждой точки функции** $y = f(x)$**, определяемой парой координат** $\{x, y = f(x)\}$**, положим** $\xi = y$ **и** $\zeta = x$**. Тогда при изменении** ζ **на отрезке** $0 = \zeta_H \leq \zeta \leq \zeta_K = b$ **функция** $\xi = f(\zeta)$ **определяет на оси** Y **функцию** $x = V(y)$ **обратную к функции** $y = F(x)$ **. Любые точки** (x_v, y_v) **и** (ξ_v, y_v)**, принадлежащие кривым** $y = F(x)$ **и** $x = V(y)$ **соответственно, лежат на**

равных расстояниях от биссектрисы. Это легко следует из треугольника, образованного пересечением линий $y = x_v$ и $\xi = \zeta_v$ с прямой, перпендикулярной биссектрисе, проходящей через данные точки. Такая же связь определяется между функциями $y = f(x)$ и $x = v(y)$. Таким образом имеем линейное преобразование функций относительно биссектрисы. Его свойства известны, а переименование координат, примененное выше[27], не надобно. Из анализа, данного в примерах 8-9 и работы [21,108], следует общее *свойство зеркальной симметрии.*:

> **{A.12.2}** ***Зеркальная симметрия относительно*** **прямой $y = c_0 + c_1 x$ в координатах $T0\Omega$ (ось T совпадает с линией $y = c_0 + c_1 x$) определяет на оси Ω *функцию* $\omega = v(\tau)$, *обратную к функции* $\tau = f(\omega)$.**

При значениях $c_0 = 0$, $c_1 = 1$ имеем случай рассмотренный в [21,108]. Очевидно, преобразование определяет взаимно однозначное соответствие и между функциями $y = F(x)$ и $x = V(y)$.

2. Но наша цель – преобразование закона распределения. Теперь все просто. Для каждого значения координаты $x' = V(y_{1)})$. (индекс обозначает вариант преобразования) по исходному закону распределения определяется значение вероятности (обозначено окружностью) $P_{\mathbf{\Phi}}\{(x' = V(y_{1)}))\} = x'/b$, которое присваивается координате $y_{1)}$. Т.е. имеем равенство $P^{1)}_{\mathbf{\Psi}}(y) = P_{\mathbf{\Phi}}\{x' = V(y))\}$, которое соответствует 1-му варианту преобразования. Во-вторых. Закон распределения $\mathbf{P}^{2)}_{\mathbf{\Psi}}(0 \le y < y_d)$ получен на отрезке $0 \le y_d \le b$ оси Y. Линейным преобразованием $x' = (b/y^{b)} \cdot y$ он преобразуется в закон $\mathbf{P}^{2)}_{\mathbf{\Psi}}(0 \le x' < x'_d)$ распределения на отрезке $0 \le x'_d \le b$ оси X. В-третьих. Линейным преобразованием $x' = y$ закон распределения $\mathbf{P}^{1)}_{\mathbf{\Psi}}(0 \le y < y_d)$ преобразуется в закон $\mathbf{P}^{1)}_{\mathbf{\Psi}}(0 \le x' < x'_d)$ распределения на отрезке $0 \le x'_d \le b$ оси X. Из свойств линейного преобразования следует равенство $\mathbf{P}^{1)}_{\mathbf{\Psi}}(0 \le y < y_d) = \mathbf{P}^{2)}_{\mathbf{\Psi}}(0 \le x' < x'_d)$. Таким образом, наше голословное утверждение «узаконено».

Из анализа, данного в примерах 8-9, следуют выводы:

> **W.11. Как *монотонная*, так и *немонотонная формула преобразования* определяют *монотонно возрастающий закон* распределения величины $\mathbf{\Psi}$ и *взаимное однозначное соответствие* между *исходным и преобразованным законами* (и их плотностями) распределения.**

> **W.12. *Принципиального отличия* распределений как при *монотонной*, так *немонотонной формуле преобразования*, полученных *с применением 1-го или 2-го варианта преобразования не существует*: они определяют *один и тот же закон* распределения величины $\mathbf{\Psi}$: 1-й на оси Y, а 2-й на оси X.**

Небольшое отступление (**«философствование»**). Весь текст, предшествующий примеру 9 (**после названия раздела 1.2, стр.14**), конечно же, можно было пропустить, сказать о *линейном преобразовании* (**сократив пример 9**) и сразу дать правильное решение, что обычно и делается.

---

[27] **Переименование, это история того, как мы пришли к общему свойству зеркального отображения: в геометрии это известно и очень давно**

Мысль все изложить так, как было в реальности[28], связана с интересными размышлениями В. Феллера [7,15]: «Роли строгости и интуиции часто понимаются неправильно. Как уже отмечалось в первом томе, природная интуиция и естественный способ мышления дают мало, но они становятся сильнее по мере развития математической теории. Сегодняшние интуиция и применения опираются на сложные теории вчерашнего дня. Кроме того, строгая теория не роскошь, а способ экономии мышления. В самом деле, наблюдения показывают, что в применениях многие авторы полагаются на длинные вычисления, а не на простые рассуждения, которые кажутся им рискованными». Мы практически не ссылались на эту работу, хотя вначале предполагалось другое: просто при написании стало понятно, что это ведет к усложнению изложения и, соответственно, «затуманивает» понимание рассматриваемых положений теории.

Размышления вполне созвучны с нашими мыслями и наблюдениями. Но дело не в этом, а в том, что проведение исследований (**как при анализе существующей системы исходных понятий теории вероятности, так и при разработке новой исходной системы**) основано как раз на этих «простых рассуждениях»: естественно, что это отразилось и на последующем их изложении[29] в работах [10; 19; 25]. А простые рассуждения действительно рискованные: мы в этом убедились не только по наблюдениям, но и на собственном опыте, в том числе, при проводимых исследованиях: показать это также послужило причиной изложения указанного текста.

Мы не знаем, когда впервые была рассмотрена «функция от случайной величины», кто автор разработки подхода к решению задачи и какие рассуждения привели его к конкретному алгоритму. Но одно можно сказать точно – мы (в примере 5, стр.18), как и первый автор, разработавший подход, в своих «простых рассуждениях» *допустили* одну и ту же «*простую ошибку*». *Не учли связь преобразования* этого вида с *зеркальной симметрией, из которой следует, что абсолютное значение производной от монотонно убывающей* функции определяет *зеркальное отображение* функции, вследствие чего она становится *возрастающей*. «Малюсенькая ошибочка» в простых рассуждениях, однако, ее последствия достаточно серьезные. Подобных ошибок было немало, особенно при рассуждениях на предварительном этапе исследований и при разработке новой исходной системы теории событий. Далеко не все из них перечислены, но каждый раз, положения, введенные на основе рассуждений, тщательно проверялись, в том числе, и «длинными вычислениями»: согласие с предыдущими и влияние па последующие положения. Как уже было сказано, мы не знаем, почему эта «привычка не сработала» при решении рассмотренной задачи.

Подчеркнем важные моменты, следующие из данных рассуждений:

1. *Схема Ферми-Декарта* определила *возможность* изучать *свойства кривых* (*геометрических* фигур) на плоскости или в пространстве, используя *алгебраические* уравнения, и *объединила* две *математические* науки –

[28] Первая попытка такого рода была предпринята при подведении итогов (пример 26 [29,88), однако стало очевидно, что восстановить «ход» рассуждений, каким он был не получится: давно (годиков этак 6-7 минуло) это было. Поэтому рассуждения в примере – это в большей степени «взгляд с позиций сегодняшнего дня». А примеры 5-6 «появились практически сегодня»

[29] По крайней мере, в работах [19; 25; 29] (и, конечно же, настоящей) мы пытались пояснить все просто и понятно «не прибегая к высокому математическому штилю»: насколько это удалось – это другой вопрос

***геометрию и алгебру – в единое целое.* Однако *это вовсе не означает,* что *кривые или геометрические* фигуры перестали *быть таковыми*: *какими они были* (пример 3, [29,21]), *такими и остались*! 2. В свою очередь, это определяет *необходимость* (вообще говоря – обязательность) *геометрических построений* при рассмотрении тех или иных положений в механике, физике и, конечно же, в математике тоже. Об этом говорилось в самом начале анализа существующего понятия случайной величины [29,7]).**

«Длинные вычисления[30]» тоже требуют проверки и перепроверки, но чаще всего это делается опять же с применением «длинных вычислений, обходя сторонкой простые рассуждения». На этом закончим «философствование» и вернемся к основному изложению.

1.3. *Условия «независимости»* случайных величин и *функциональная зависимость* распределения от *условий проведения испытания*

Рассмотрим *понятие* и *условие независимости* 2-х случайных величин, которое в существующей теории вероятностей имеет смысл, подобный понятию *независимости* 2-х *сложных* событий *испытания*:

> **E.** «Случайные величины **X** и **Y** называются *независимыми*, если *закон* распределения *каждой из них не зависит* от того, какое *значение приняла другая. В противном случае* величины **X** и **Y** *зависимы*» [13,172].

Смысл определения, вроде бы, понятен и, казалось бы, не вызывает возражений, однако здесь эту сторону понятия рассматривать не будем[31], а обратимся к «*условиям независимости*» величин **X** и **Y**.

***Дискретные* величины X и Y *независимы*, если *для любой* пары значений x_v v=1,2,...,V и y_w w=1,2,...,W *вероятностной* функции $p(x_v, y_w)$ выполнено условие $p(x_v, y_k) = p(x_v) \cdot p(y_k)$ (3). При *непрерывных* случайных величинах X и Y *условие независимости* можно записать для *плотностей*[32] $p(x, y) = p(x) \cdot p(y)$. (4)**

[30] **Под «длинными вычислениями» мы подразумеваем не только непосредственные вычисления, необходимые для решения какой-то конкретной задачи, но и создание более сложных математических моделей, определяющих развитие теории и возможность ее применения к более сложным явлениям**

[31] **Смысл понятия будет будет пояснен при рассмотрении видов зависимости случайных величин (часть VI исследований)**

[32] **В современных работах оно чаще записывается с применением условных плотностей (законов) распределения $p(x|y)$ или $p(y|x)$, например, $p(x|y) = p(x)$ или $p(y|x) = p(y)$. Это, вообще говоря, в принципе неверно**

В существующей теории вероятностей, *условия* (3-4) рассматриваются как *необходимое и достаточное условие независимости* двух случайных величин.

Дискретные случайные величины являются другим представлением *испытания* (**раздел 2 [29,39], более подробно в разделах 3 и 4 [29,42,66]**), следовательно, **анализ, проведенный в приложении III [19,53] и выводы Q.4-Q.6 [19,60] следующие из анализа**, справедливы и для *дискретных* случайных величин.

Поэтому рассмотрим *непрерывные* величины, для этого обратимся к *преобразованию непрерывного закона* распределения *одномерной* случайной величины Φ в *закон* распределения величины Ψ, которое определяется *формулой преобразования* $y = f(x)$.

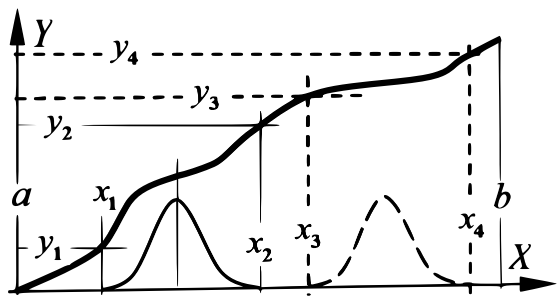

Рис. 12.

***Пример* 10. Пусть на отрезке $a \le x \le b$ задана монотонно возрастающая функция $y = f(x)$ (рис.12, утолщенная сплошная кривая), на части ($x_1 \le x \le x_2$) которого определена функция плотности $p_\Phi(x)$ (тонкая сплошная кривая) случайной величины Φ_X. *Преобразованная* функция *плотности* величины Ψ_Y^1, полученная *с применением* функции $y = f(x)$, определена на отрезке $y_1 \le y' \le y_2$ и *зависит только* от переменной y'. Это очевидно для 2-го варианта преобразования (формулы 2 и замечание 3.3, стр.9), если функция $x = \phi(y')$, обратная функции $y = f(x)$, явно выражается через переменную y'. Утверждение остается справедливым, если обратная функция в явном виде не выражается. Отсюда следует:**

> **W.13. *Совместая плотность* распределения *системы*, образованной *совмещением* случайны величин Φ_X и Ψ_Y^1 определяется *произведением плотностей*, т.е. отвечает *условию* (4): согласно принятой теории они «*независимы*».**

Замечание. **Определив функцию плотности $p_\Phi(x)$ на отрезке $x_3 \le x \le x_4$ (штриховая кривая), получим плотность распределения величины Ψ_Y^2, определенную на отрезке $y_3 \le y'' \le y_4$ и зависящую только от переменной y'', следовательно, законы распределения величин Φ_X и Ψ_Y^2, в соответствии с существующим пониманием, тоже независимы. Распределение случайной величины Ψ_Y^1 зависит только от переменной y', а случайной величины Ψ_y^2 – только переменой y'', т.е. произведение величин Ψ_Y^1 и Ψ_y^2 отвечает *условиям* (4), следовательно, они тоже «*независимы*». *Изменяя положение закона* распределения величины Φ_X (в пределах отрезка $a \le x \le b$) *относительно* кривой $y = f(x)$, получим на отрезке ($a \le x \le b$) множество (в принципе – бесконечное несчетное) *законов* распределения «независмых» (согласно принятым условиям) величин Ψ_j (j=1,2,. . .), которые *связаны функциональной зависимостью* $y = f(x)$ с величиной Φ и *между собой.***

{A.13.1} ***Каждое из произведений*** **величин $\mathbf{\Phi_X}$ и $\mathbf{\Phi^j_Y}$ (j=1,2,...)** ***отвечает условию*** **(3), следовательно, величины $\mathbf{\Phi_X}$ и $\mathbf{\Phi^j_Y}$ «*независимы*», но в то же время законы распределения величин $\mathbf{\Phi^j_Y}$** ***связаны функциональной зависимостью*** $y = f(x)$ ***с законом распределения*** **случайной величины $\mathbf{\Phi_X}$.**

{A.13.2} Если рассматривается ***одновременное*** **появление** ***элементарных*** **событий, принадлежащих случайным величинам $\mathbf{\Phi}$ и Ψ_j, то они являются** ***системой*** **случайных величин, образованной их** ***совмещением.***

Анализ также показывает, что функцию $y = f\,(x)$ можно трактовать как *влияние* случайной величины $\mathbf{\Phi}$ на случайную величину $\mathbf{\Psi}$ (**и обратно**) с одной стороны, а с другой – как *изменение некоторого условия* x проведения *эксперимента* с величиной $\mathbf{\Phi}$ (**или $\mathbf{\Psi}$**). Из данного анализа следуют выводы:

W.14. Имеется ***явное противоречие*****:** ***функционально зависимые*** **случайные величины $\mathbf{\Phi}$ и $\mathbf{\Psi}$ являются,** ***в соответствии с принятыми условиями,*** $y = f\,(x)$ **«*независимыми*» .**

W.15. ***Действительная*** **функция (формула преобразования)** $y = f\,(x)$ **определяет: 1.** ***Преобразование*** **распределения величины $\mathbf{\Phi}$ в** ***распределение*** **величины $\mathbf{\Psi}$. 2.** ***Функциональную зависимость распределений*** **случайных величин $\mathbf{\Phi}$ и $\mathbf{\Psi}$. 3.** ***Влияние распределения*** **случайной величины $\mathbf{\Phi}$** ***на распределение*** **случайной величины $\mathbf{\Psi}$ (и наоборот). 4.** ***Зависимость распределения*** **величины $\mathbf{\Psi}$** ***от изменения условия*** x**, определяющего** ***проведение эксперимента*** **с величиной $\mathbf{\Phi}$.**

W.16. ***Зависимость*** **случайных величин** ***определяется изменением внутренних и/или внешних условий опытов.*** **Существующее понимание** ***понятия*** **и** ***условий независимости не применимы ни*** **к событиям** ***испытания, ни к случайным величинам.***

W.17. ***Если*** **случайные величины** ***связаны функциональной зависимостью,*** **то они являются** ***системой*** **случайных величин. Следовательно,** ***одномерные*** **величины,** ***образованные по распределению многомерной*** **случайной величины** ***не могут быть функционально зависимыми.***

W.18. ***Если действительная функция*** $y = v(x)$ ***определяет функциональную зависимость закона распределения случайной величины*** **$\mathbf{\Phi}$** ***от изменения условия проведения данного опыта, то*** **2-а (и более)** ***положения закона распределения величины определяют законы распределения*** **2-х (и более)** ***случайных величин, которые связаны функциональной зависимостью*** $y = v(x)$ ***между собой.***

В *новой исходной системе* теории случайных величин определена *двумерная* функция плотности $p_{\Pi}|_{\times}(x, y) = p(x) \cdot p(y)$ *произведения системы одномерных* случайных величин $\mathbf{X}$ и $\mathbf{Y}$ (**определение 9 [29.76]**) при *их совмещении*. При ее построении понятия *независимости* (*зависимости*) вообще не упоминались: *оно просто лишнее*. Соответственно, *никаких противоречий не возникает*.

Зависимость распределения величины $\mathbf{\Psi}$ *от условий проведения опыта* с величиной $\mathbf{\Phi}$ приводит и другим последствиям, которые будут рассмотрены при анализе *видов зависимости* случайных величин (**часть VI**). Там же будет показан *реальный смысл* существующей трактовки понятия «*независимости*», с которого был начат данный подраздел, и что в действительности определяют *условия* **(3-4)**.

2. Преобразование двумерных распределений

Преобразование, рассмотренное в разделе, в теории *действительных* функций не применяются (**по крайней мере, об этом прямо не говорится**), но его можно связать с вычислением двойных (**многократных**) интегралов.

В существующей теории вероятностей задача ставится в очень общем виде (**под названием функция от случайных величин, комментарий на стр.3**), однако в примерах в основном приводится преобразование *двумерных непрерывных* распределений. Б.В. Гнеденко рассмотрены три важных частных случая:

1. «Функция распределения суммы» [1,136]. 2. «Функция распределения частного» [1,143]. 3. «Поворот осей координат» [1,146]. В других работах в основном рассматривается 1-й случай, в некоторых – также и 2-й, но иногда под названием (например [13; 17]) произведение случайных величин. Что касается поворота координат, то из цитируемых (и просмотренных тоже) работ, он рассматривается (кроме [1]) также в [13] при обработке результатов стрельб, хотя название «поворот осей координат» не применяется.

«Нашлись» примеры *преобразования* и *дискретных* распределений: как с «*независимыми*» (**[12,188], рассмотрен в [29,32]**), так и «*зависимыми*» величинами (**в работе[33] [6] – мы его не рассматриваем: пояснений требуется много, а наглядности не добавляет**).

Преобразование координат применяется в математическом анализе, но подробно рассматривается в аналитической геометрии (**поворот декартовых осей – как частный случай общего преобразования координат**). *Поворот осей* координат был рассмотрен нами при анализе понятия *ковариации* [10,67] и *вероятностной зависимости* случайных величин [10,99]. О *повороте* координат, подробнее поговорим при рассмотрении *числовых характеристик двумерных* (**многомерных**) распределений (**часть V**): именно одна из них – так называемая *ковариация двумерного* распределения – его и определяет (**а «не момент связи», как считается в принятой теории вероятностей**).

Некоторые моменты, связанные с *преобразованием двумерных* распределений, рассмотрены при анализе *видов зависимости* случайных величин (**часть VI исссследований**).

Здесь немного поговорим о *преобразовании непрерывных* распределений и более подробно – о *преобразовании многомерных дискретных* распределений: в работе [10,82] оно «затронуто вскользь» в связи с рассмот-

---

[33] **В [6] на стр.234 (пример ж) приведены только распределения вероятностей «суммы» и «произведения» величин X_1 и X_2; построение «совместного» распределения величин на стр.228-229, а исходные данные для построения на стр.26-27. Отметим, что в этом томе в основном рассматриваются дискретные случайные величины**

рением *аппроксимации биномиального* распределения кривой[34] $\mathbf{exp}(-t^2/2)$, которое блестяще выполнено А. Муавром.

Исходя из *новой исходной* системы теории случайных величин, *преобразование двумерного* распределения к *одномерному* распределению можно выполнить множеством способов (**примечание 4 [10,56]**), из которых в работе [10] выделены два. Они, как показано далее, *применяются* в существующей теории вероятностей, несмотря на то, что при *их трактовке* понятие *преобразования* пратически не упоминается.

I. *1-й способ преобразования* – определяется *перемещением однозначной* кривой (функции $y(x)$, т.е. *формулы преобразования*) на плоскости $X0Y$ *параллельно одной* из осей [10,56].

1. Если *формула преобразования* линейна $f(x) = a \cdot x + b$, то при значениях a=-1 и b=0 имеем $y = -x$: в существующей теории (пример 9 и пояснения к нему [29,32]) ее записывают в виде Z=X+Y, называют функцией «суммы» от случайных величин X и Y, и «доказывают теорему» $M(\mathbf{X}+\mathbf{Y}) = M(\mathbf{X}) + M(\mathbf{Y})$. 2. В приложениях теории вероятностей встречается формула математического ожидания $M(\mathbf{X}-\mathbf{Y}) = M(\mathbf{X}) - M(\mathbf{Y})$ «разности» $\mathbf{Z_1} = \mathbf{X} - \mathbf{Y}$ величин. В работах по теории вероятностей мы ее не видели, но она, как и отмеченная выше, следует из общего равенства $M(K\mathbf{X}+L\mathbf{Y}) = KM(\mathbf{X})+LM(\mathbf{Y})$ (свойство IV [9,60]). В принятой теории она соответствует записи формулы преобразования $y = x$ в виде $\mathbf{Z_1} = \mathbf{X} - \mathbf{Y}$.

II. *2-й способ преобразования* потребовался *для преобразования двумерных законов распределения произведения* и *суммы* системы (определения 9-10 [29,75]), *образованной совмещением* случайных величин X и Y к *одномерному закону*, ибо *первый* способ не определяет *ни того, ни другого* (п.6 примера 12 стр.38; вывод S3.12 [10,60]). Вообще говоря, это частный случай общего преобразования. Он связан с *применением* функции $f(x,\, y,\, a)$, *зависящей от параметра* a, в его простейшем виде – *линейной зависимости* переменных $x = z$ и $y = l_y \,/\, l_x \cdot z$ от параметра z [10,59], где l_x, l_y – области определения законов распределения величин X и Y: это определяет *одновременность изменения* аргументов *вероятностных* функций. Это и *позволяет преобразовать* эти *законы* распределения к *одномерному* распределению.

Формула преобразования, зависящая от параметра, применяется и в существующей теории: используется гиперболическая функция $y = \pm z/x$, зависящая от параметра $z > 0$. Ее чаще записывают в виде Z=X/Y и называют [1] *функцией распределения частного* «независимых» величин. Но в некоторых работах [13,270; 17,798] ее записывают в виде $\mathbf{Z} = \mathbf{X} \cdot \mathbf{Y}$ и называют *функцией распределения произведения* «независимых» величин.

***Замечание*. Трактовка *формулы преобразования*, как некоторой функции от случайных величин, *неверна*. Это привело не только к появлению отмеченных «теорем», но также и других «теорем» о *дисперсии суммы и разности*, о *математическом* ожидании *произведения* случайных величин и т.д.**

---

[34]Мы понимаем, какое значение придают в существующей теории нормальному распределению. Тем не менее подчеркиваем, что это одна из возможных аппроксимаций биномиального распределения: давая название своей выдающейся работе, А. Муавр прекрасно понимал, что им реально выполнено. Можно возмущаться, но это именно так: подробно эти вопросы обсуждаются при анализе числовых характеристик распределений (часть V исcсследований)

2.1. Непрерывное двумерное распределение

В работе [10] подробно рассмотрены оба способа в общем виде. Здесь ограничимся в основном 1-м способом *преобразования двумерного непрерывного* распределения с *линейной формулой преобразования* [10,59] и покажем его связь с *вычислением двойных* (**многократных**) интегралов. Во-вторых. В работе [29,62-63] показано, что для *двумерной* случайной $\mathbf{\Phi}\{(x_j, y_k) \to \mathbf{a_{j,k}}\}$ величины *ни суммы*, *ни произведения одномерных* случайных величин $\mathbf{\Phi}(x_j \to \mathbf{A_j})$ и $\mathbf{\Psi}(y_k \to \mathbf{B_k})$ *не существует.* Они определяются *только для системы* (**определения 8-10 [29,76]**), образованной *совмещением одномерных* случайных величин $\mathbf{\Phi}(x_j^1 \to \mathbf{a_j^1})$ и $\mathbf{\Psi}(x_k^2 \to \mathbf{a_k^2})$. При анализе покажем, что даже в этом случае этот *способ не определяет преобразование произведения* или *суммы одномерных* случайных величин: для *этого и потребовалась разработка* 2-го *способа преобразования.*

***<u>Пример</u>* <u>12.</u> Пусть *формула преобразования линейна* $y = -cx$ $(c > 0)$. Равенство $y = -cx$ означает, что в системе координат $X0Y$ все точки плоскости с координатами $(x',\ y' = -x')$ принадлежат данной прямой. Соответственно, неравенства $y < -cx$ и $y > -cx$ определяют все точки плоскости, которые лежат ниже или выше прямой $y = -cx$.**

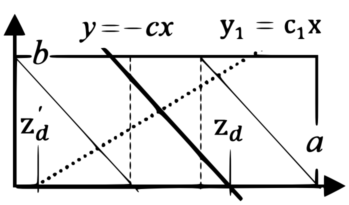

Рис. 13.

Выделим прямоугольную область $\{0 \le x \le a, 0 \le y \le b\}$ и положим $y = c(z - x)$, *где* $z \ge 0$ *действительное* число. При *изменении* параметра z от значения z_H до значения z_K функция $y = -cx$ *перемещается параллельно* самой себе вдоль оси X. При значениях $z_H = 0$, $z_1 = b\,/\,c$, $z_2 = a$ и $z_K = a + b\,/\,c$ прямая $y = -cx$ линия проходит через точки $(0, 0)$, $(0, b)$, $(a, 0)$ и (a, b) соответственно.

п.1. Рассмотрим положение прямой при промежуточных значениях $0 \le z_d \le z_K$ (рис.13, утолщенная сплошная линия). Линия перемещается в направлении оси X, следовательно, неравенство $y \le c(z - x)$ определяет все точки части прямоугольника (область D), которые лежат ниже прямой. Площадь S_D области D определяется двойным интегралом $S_\mathrm{D} = \iint dx\,dy$. Двойной интеграл вычисляется [24,149] приведением к повторному интегралу $S_\mathrm{D} = \int_0^{z_d} dx \int_0^{z_d - x} dy$ {a1}. Если прямая линия проходит через точку $(0, 0)$или (a, b), то все точки прямоугольника лежат выше или ниже прямой $y = c(z - x)$, т.е. $S_\mathrm{D} = 0$ или $S_\mathrm{D} = S = a \cdot b$ соответственно.

{A.14} Естественно ввести ось Z, совпадающую с осью X, на которой координаты точек определяются значением $z = x + y\,/\,c$.

При *изменении* параметра z_d от значения z_H до значения z_K формула (a1) определяет *одномерную зависимость изменения* площади $S_\mathrm{D}(z = x + y/c)$ (подмножества точек прямоугольника, лежащие ниже прямой $y = -cx$) от значения параметра z, т.е. от *изменения положения* функции $y = -cx$ при ее перемещении *параллельно оси* Y *в положительном* направлении оси X.

п.2. Если в области $\{0 \le x \le a, 0 \le y \le b\}$ определена действительная функция $\Phi(x, y)$, которая имеет непрерывную производную $\phi(x, y)$, то сумма значений функции W_D, которые лежат ниже прямой определяются двойным интегралом по области D $W_\mathrm{D} = \iint \phi(x, y) dx\,dy$. Приведением к повторному интегралу [23], получим: $W_\mathrm{D} =$

$\int_0^{z_d} dx \int_0^{z_d - x} \phi(x,y)\, dy$ **{a2}. Если функция $\varphi(x,y)$ представима в виде $\phi(x,y) = \psi_1(x) \cdot \psi_2(y)$, то формула (a2) принимает вид $W_{\mathbf{D}} = \int_0^{z_d} \psi_1(x) dx \int_0^{z_d - x} \psi_2(y) dy$ {a3}.**

При промежуточном значении параметра z_d формулы {a2}, {a3} определяют *одномерные зависимости изменения суммы* $W_{\mathbf{D}}(z = x + y\,/\,c)$ значений функции $\Phi(x,y)$ *от изменения положения* функции $y = -cx$ при ее *перемещении в положительном* направлении оси X, т.е. при *увеличении* параметра z_d.

п.3. Если в области $\{0 \le x \le a,\ 0 \le y \le b\}$ задана *плотность* $p(x,\ y)$ распределения *непрерывной двумерной* случайной величины $\mathbf{\Phi}$, то вероятность того (вывод W.27 [29,65]), что в области D появится *одно* из множества *элементарных* событий $\mathbf{a_{j,k}}$ (с двумя метками, определение 3 [19,29]), принадлежащих этой области, определяется двойным интегралом $P_{\mathbf{D}} = \iint p(x,y) dx\, dy$. Приведением к повторному интегралу [23], получим $P_{\mathbf{D}} = \int_0^{z_d} dx \int_0^{z_d - x} p(x,y)\, dy$ (5.a).

Если *вероятностная* функция есть *произведение* $p_{\Pi\,|\,\times}\{(x_j, y_k)\} = p(x_j) \cdot p(y_k)$ (раздел 4.1.2 [29,73]) *вероятностных* функций системы величин X и Y, образованная их *совмещением* X×Y, то формула {a4} имеет вид $P_{\mathbf{D}} = \int_0^{z_d} p(x) dx \int_0^{z_d - x} p(y) dy$ (5.b).

Хотя формулы (5.a) и (5.b) внешне отличаются, но не столь существенно, как отличается их смысл: по формуле (5.b) определяется вероятность события (вывод W.32 [29,78]), *заключающегося в одновременном* появлении *элементарных* событий «и $\mathbf{a}^1_\mathbf{j}$, и $\mathbf{a}^2_\mathbf{k}$» из множеств *элементарных* событий *каждого из опытов*, *принадлежащих области* ($x_{\mathbf{H}} \le x < x_{\mathbf{dan}},\ y_{\mathbf{H}} \le y < y_{\mathbf{dan}}$).

При промежуточном значении формулы (5.a), {5.b} определяют *одномерные зависимости изменения вероятности* $P_{\mathbf{D}}(z = x + y/c)$ *от изменения положения* функции $y = -cx$ при ее *перемещении в положительном* направлении оси X, т.е. при *увеличении* параметра z.

п.4. Можно рассматривать функцию $y_1 = c_1 x$ $(c_1 > 0)$ (рис.13, пунктирная линия), полагая $y_1 = c_1(x - z')$. При значениях $z'_H = -\,b\,/\,c_1$, $z'_1 = 0$, $z'_2 = a - b\,/\,c_1$ и $z'_K = a$ *прямая* $y_1 = c_1 x$ линия проходит через точки $(0,b)$, $(0,0)$, (a,b)и $(a,0)$ соответственно.

В этом случае, в отличие от функции $y \le c(z - x)$ (п.1), *неравенство* $y \le c(x - z)$ определяет все точки *части прямоугольника* (область D), которые *лежат выше* прямой. В остальном *преобразование* проводится также как и в пунктах 1-3: можно просто в формулах формально заменить функцию $y = c(z - x)$ функцией $y_1 = c_1(x - z')$.

<u>Замечание</u>. Зависимости $S_{\mathbf{D}}(z = x - y/c)$, $W_{\mathbf{D}}(z = x - y/c)$ **и** $P_{\mathbf{D}}(z = x - y/c)$, ***полученные перемещением* функции $y_1 = c_1 x$ будут *отличаться от зависимостей*, *полученных перемещением* функции $y = -c_1 x$. Это будет показано в примере 16 (стр.54).**

п.5. В пунктах 1-4 значение параметра z изменяется по оси X. Если перемещать обратную функцию $x = -\,y\,/\,c$ (или $x_1 = x\,/\,c_1$) вдоль оси Y (или X), то параметр z будет изменяться по оси Y, однако зависимости $S_{\mathbf{D}}(z)$, $W_{\mathbf{D}}(z)$ и $P_{\mathbf{D}}(z)$ не будут отличаться от зависимостей, определяемых функциями $y = -\,cx$ или $y_1 = c_1 x$.

Отличие появляется при *нелинейной формуле преобразования* $y = f(x)$.

п.6. Рассмотрим систему, образованную *совмещением* дискретных величин с вероятностными функциями $p(x_j)$, $p(y_k)$ (j=1,2,...,n1; k=1,2,...,n2), определенными на отрезках $(0 \le x \le a)$ и $(0 \le y \le b)$, где $x_j = x_{j-1} + \Delta_x \cdot j$ $(\Delta_x = a/n1)$, $y_k = y_{k-1} + \Delta_y \cdot k$ $(\Delta_y = b/n2)$.

На рис.13а. условно изображена таблица вероятностной функции двумерного распределения произведения значений функций $p(x_j),\ p(y_k)$. В 1-мварианте (п.I стр.35) вычисляются суммы значений в точках (x, y), расположенных на штриховой (при непрерывном распределении – точки ниже) прямой линии. При перемещении функции $y = -x$ параллельно оси Y в положительном направлении оси X, получим последовательность сумм: $P_1 = p_1^1 \cdot p_1^2$; $P_2 = p_1^1 \cdot p_2^2 + p_2^1 \cdot p_1^2$; $P_3 = p_1^1 \cdot p_3^2 + p_2^1 \cdot p_2^2 + p_3^1 \cdot p_1^2$; .... ((a))

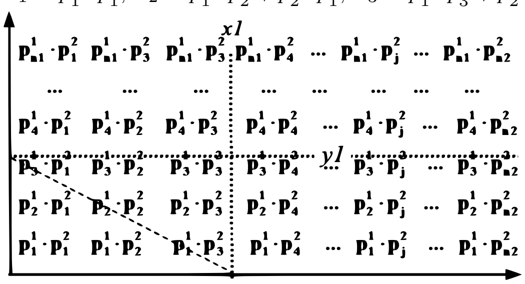

Рис. 13а.

Для системы, образованной *совмещением* величин **X** и **Y**, вероятности событий **A** и **B**, определяемых величинами равны: $P(A) = \sum_{v=j0}^{j1} P(x_v) = \sum_{v=j0}^{j1}\sum_{w=1}^{n1} P(x_v, y_w)$ и $P(B) = \sum_{w=i0}^{i1} P(y_w) = \sum_{v=1}^{n2}\sum_{w=i0}^{i1} P(x_v, y_w)$. Отсюда следует: *преобразование суммы* значений вероятностной функции системы, образованной *совмещением* вероятностных функций величин **X** и **Y**, к *одномерной* случайной величине возможно только для прямоугольной области $D(x_{j0} \le x < x_{j1}, y_{i0} \le y < y_{i1})$. Очевидно, что этому не отвечает последовательность сумм ((a)). Таким образом:

> **W.19.** ***Функция изменения вероятности, полученная 1-м вариантом преобразования, не имеет никакого отношения ни к распределению произведения, ни к распределению суммы*** **случайных** ***величин*** **X** ***и*** **Y** ***при их совмещении.***

Преобразование двумерной вероятностной функции системы, образованной *совмещением* вероятностных функций величин **X** и **Y** к одномерному виду достигается при введении 2-го варианта преобразования (п.II стр.35). Для этого полагаем, что переменные $x = z_d$ и $y = z_d \cdot b\,/\,a$ зависят от одного параметра z_d. В этом случае при *изменении параметра* z_d *переменные* $x = z_d$ и $y = z_d \cdot b/a$ *изменяются одновременно.*

В результате для области $D(0 \le x < x_{j1}, 0 \le y < y_{i1})$ имеем последовательность $P'_1 = p_1^1 \cdot p_1^2$; $P'_2 = (p_1^1 + p_2^1) \cdot (p_1^2 + p_2^2)$; $P'_3 = (p_1^1 + p_2^1 + p_3^1) \cdot (p_1^2 + \cdot p_2^2 + p_3^2) \cdot (p_1^3 + p_2^3 + p_3^3)$;..., которая соответствует *произведению вероятностных* функций при *совмещении* случайных величин **X** и **Y**. Чтобы преобразовать *сумму вероятностных* функций при *совмещении* случайных величин **X** и **Y** к *одномерному* виду, необходимо рассмотреть *сумму* событий в двух прямоугольных областях $D1\ (0 \le x < x_{j1},\ 0 \le y < y_{n1})$ и $D2\ (0 \le x < x_{n1},\ 0 \le y < y_{i1})$, учитывая совместимость этих событий.

Если **X** и **Y** непрерывные величины с плотностями $p(x),\ p(y)$, которые определены на отрезках $(0 \le x \le a)$ и$(0 \le y \le b)$ (полагаем $a > b$), то одномерные функции изменения вероятности *произведения и суммы* при *совмещении* величин имеют вид:

$$\mathbf{P}_{\mathbf{X}\cdot\mathbf{Y}}(0 \le z < z_d) = \int_0^{z_d} p(x)dx \cdot \int_0^{z_d \cdot b/a} p(y)dy; \qquad \textbf{(6.a)}$$

$$\mathbf{P}_{\mathbf{X}+\mathbf{Y}}(0 \le z < z_d) = \int_0^{z_d} p(x)dx + \int_0^{z_d \cdot b/a} p(y)dy - \mathbf{P}_{\mathbf{X}\cdot\mathbf{Y}}(0 \le z < z_d). \qquad \textbf{(6.b)}$$

> **W.20.** Таким образом, *преобразование двумерных* (многомерных) *законов* распределения *произведения и суммы системы*, образованной *совмещением* случайных величин, *к одномерному закону* определяет 2-й *способ преобразования* (п.II стр.35).

п.7. Отметим еще одну особенность *преобразования двумерных* (многомерных) распределений, которая существенно *отличает* его от *преобразования одномерных* распределений. Если *преобразование одномерных* распределений устанавливает *взаимное однозначное соответствие* между *исходным* и *преобразованным законами* распределения (вывод **W.11**, стр.28), то при *преобразовании*

двумерных **(многомерных) распределений в общем случае «обратного хода нет». По *одномерной зависимости* $P_D(z = x + y/c)$ *закон* распределения *двумерной* случайной величины *не определяется*.**

Однако если *закон* распределения есть *произведение системы* случайных величин, образованной *их совмещением*, то при *известном законе одной из исходных* величин определяется *закон* распределения другой.

Пусть $p(x_m, y_n) = p(x_m) \cdot p(y_n) = p^1_m \cdot p^2_n$ *вероятностная* функция *произведения дискретных* случайных величин X и Y. Для определенности положим $M \geq N$. На рис.13а. условно изображена таблица *вероятностной* функции для первых трех значений. В 1-м *варианте* определяются *суммы* значений в точках (x, y), расположенных на штриховой прямой. Отрезок прямой является диагональю прямоугольника (ограниченного сплошными прямыми), в котором расположены *суммы*, определяемые 2-м *вариантом преобразования*.

В 1-м способе (п.I стр.35) определяется *последовательность сумм*, которые вычисляются по формуле $\Sigma_w = \sum_1^w p^1_{w+1-v} \cdot p^2_v$ (w=1,2,...,N): значения p^2_w для данного значения w=1,2,...,N определяются *последовательно из равенств* $p^1_1 \cdot p_w = \Sigma_w - \Sigma'_w$ {*}, где $\Sigma'_w = \sum_2^w p^1_{w+1-v} \cdot p^2_v$ $(\Sigma'_1 = 0)$ – сумма, вычисленная по предыдущим значениям $p^2_1, p^2_2, \ldots, p^2_{w-1}$.

2-й способ. Определяется *последовательность сумм*, которые вычисляются по формуле $\Sigma_w = \sum_1^w p^1_v \times \sum_1^w p^2_v$: значения p^2_w для данного значения w=1,2,...,N определяются *последовательно из равенств* $p_w \cdot \sum_1^w p^1_v = \Sigma_w - \Sigma'_w$ {}, где $\Sigma'_w = \sum_2^w p^1_{w+1-v} \cdot p^2_v$ $(\Sigma'_1 = 0)$ – сумма, вычисленная по предыдущим значениям $p^2_1, p^2_2, \ldots, p^2_{w-1}$. Если значения $M < N$, то для значений w=1,2,...,M значения определяются, как и выше. Для значений w'=M+1,M+2,...,N-M следует определить соответствующие последовательности сумм, которые имеют небольшое отличие от предыдущих. Значения p^2_w вычисляются по формулам {*}-{**}.**

п.8. В теории действительных функций действительных аргументов обычно вычисляются интегралы $\xi(x) = \int_0^b \phi(x, y)\,dy$ и $\zeta(y) = \int_0^a \phi(x, y)\,dx$. Функции $\xi(x)$ и $\zeta(y)$ можно интерпретировать как зависимость *суммы значений функции* $\Phi(x, y)$ (площади, вероятности) *от изменения положения* прямой, параллельной оси Y (или X), при ее *перемещении в положительном* направлении оси X (или Y). Однако интегрирование можно выполнить и по прямой линии $y = cx$, вычисляя интеграл по площади, лежащей ниже прямой, перпендикулярной этому направлению: в этом случае интегрирование начинается в точке $(0, 0)$ и заканчивается в точке в точке (a, b). Этот способ интегрирования, как и способ, определенный в п.1-4 данного примера, сложнее: при вычислении интегралов (например, в формулах (5)-(6)) область $\{0 \leq x \leq a,\ 0 \leq y \leq b\}$ придется разделить на несколько участков.

Таким образом, показано: преобразование *двумерного* в *одномерное* распределение, соответствует *вычислению повторного* интеграла *по заданному направлению*. Применение такого *усложненного* способа интегрирования в теории *действительных* функций «не встречалось» (**оно «попадалось» в физике, механике, технике и т.п.**): скорее всего, в нем просто нет потребности. В теории вероятностей *преобразование двумерных* (***n*-мерных**) *законов* распределения необходимо в обоих вариантах (**пункты I-II, стр.35**): в том числе в ее приложениях [10,82-89,125-144].

Дифференцируя выражения **(5.a)** и **(5.b)** по параметру z [23,661], получим формулы для вычисления плотностей распределений:

$p(z) = \int_0^z p[x,\, c \cdot (z - x)]\, dx;$ **(a)** $\quad p(z) = \int_0^z p(x) \cdot p[c \cdot (z - x)]\, dx.$ **(b)** **(7)**

При значении $c = 1$ получим:

$g(z) = \int_0^z p[x,(z - x)]dx;$ **(a)** $\quad g(z) = \int_0^z p(x) \cdot p(z - x)]dx.$], **(b)** **(8)**

Формулы идентичны формулам, которые приводятся в литературе, например [1,8-9,11-20], и интерпретируются как *плотность* вероятности *суммы* **Z**=**X**+**Y** *зависимых* **(8.a)** и *независимых* **(8.b)** величин.

***Замечания* 11. 1.** **Область определения одномерных зависимостей во всех случаях, данных в примере 12, одинакова $0 \le z \le a + b / c$. 2. В литературе формулы (7-8) часто приводятся (в том числе – в справочникахе) в виде $g(z) = \int_{-\infty}^{\infty} p_\Phi[x,\, (z - x)]\, dx$. Такая запись не верна, ибо двойной интеграл в формулах вычисляется по области, лежащей ниже прямой $y = c(z - x)$ независимо от областей определения законов [23]. Следовательно, изменяется 0 на $-\infty$ только нижний предел интегрирования, в том числе, при неограниченных областях определения ($\infty < x, y < \infty$). 3. Подобным образом производится *преобразование двумерного закона* распределения в *одномерную зависимость* вероятности *от параметра* z с применением гиперболической функции $y = \pm m / x$ ($m > 0$ – число, которое и является параметром z. Т.е. имеем другой тип преобразования (замечание к пункту II, стр.35): при изменении параметра z изменяется *форма гиперболы*, т.е. имеем *зависимость распределения от изменения формы* гиперболы. Очевидно, что *ни к функции распределения частного* Z=X/Y, ни к *функции распределения произведения* Z=X·Y случайных величин X и Y преобразование никакого отношения не имеет. В принципе, этот вид преобразования применим к *любому двумерному* распределению (т.е. как к «независимым» так и «зависимым» величинам в существующей трактовке).**

Из анализа, данного выше, следует:

W.21. ***Преобразование не зависит*** **от того, что *преобразуется, действительная* функция $\Phi(x, y)$ *действительных* аргументов x и y или *законы* распределения: *двумерной* случайной величины Φ_{xy}, *произведения* $\Pi_\times = \Psi_x \Psi_y$ или *суммы* $\Sigma_\times = \Psi_x + \Psi_y$ системы, образованной *совмещением* $\Psi_x \times \Psi_y$ случайных величин.**

W.22. Функции $y = -(x-z)$ и $y = (x-z')$ можно записать в виде: $z = x+y$, $z' = x-y$. Очевидно, что их запись в виде $Z = \Psi_x + \Psi_y$, $Z = \Psi_x - \Psi_y$ *не имеют места и трактовка их как функции суммы или разности от случайных величин* Ψ_x и Ψ_y *недопустима*, соответственно функции $y = z / x$ и $z = x \cdot y$ *нельзя трактовать как частное* $Z = \Psi_1 / \cdot \Psi_2$ или *произведение* $Z = \Psi_1 \cdot \Psi_2$.

W.23. В результате *преобразования* определяются *одномерные* функции *изменения* вероятности $P_D(z = x + y / c)$, которые можно *трактовать, как законы* распределения *одномерных* случайных величин. По *одномерным законам* распределения *исходный двумерный* (многомерный) *закон* не определяется.

W.24. Вероятности *произвольных* событий вычисляются *по двумерному* (*n-мерному*) распределению. *Законы* распределения *одномерных* случайных величин, полученные *преобразованием, для этого непригодны*: определение *одномерных законов* имеет *смысл* только в случаях, необходимых для решения данной задачи.

W.25. Если *двумерное* распределение является *произведением системы* случайных величин, образованной их *совмещением*, то при *известном* распределении *одной из исходных* величин, определяется распределение другой.

Выводы справедливы для обоих способов преобразования. Они распространяются на преобразование законов распределения *n-мерной* величины, *произведения* или *суммы системы*, образованной *совмещением* n случайных величин.

{А.15} Подчеркнем еще раз (положение {А.4}, стр.12): здесь мы также говорим *о том*, что *неверна трактовка преобразования*, а не о том, что оно *выполнено неправильно*.

***Предупреждение*. Однако в последующих исследованиях *неверная трактовка* может привести к *неверным* заключениям: *правильность* выполнения *преобразования* не дает гарантии того, что этого не будет. Здесь предупреждение более актуально, чем при преобразовании одномерных распределений (положение А.4, стр.12). Это будет показано далее.**

2.2. Дискретное многомерное распределение

Преобразование многомерных дискретных распределений рассмотрим для частного случая: *повторение N=1,2,...* раз *одномерного опыта* (**с множеством элементарных событий $\mathbf{a_k}$, $\mathbf{P}(\mathbf{a_k}) = p(x_k) = p_k$ $(k = 1, 2, ..., K \geq 2)$**) *при неизменных* условиях его проведения (**расширенная схема последовательных независимых испытаний Бернулли**)

{А.16.1} Здесь и далее обозначение $p(x_k) = p_k$ применяется только для сокращения записи.

Существующее решение задачи получено в теории событий: **для вероятности того, что в N «*независимых*» *опытах* появится «и m_1 событий $\mathbf{a_1}$, и m_2 событий $\mathbf{a_2}$, ..., и $m_\mathbf{K}$ событий $\mathbf{a_K}$ », получена формула:**
$$\mathbf{P_N}(m_1, m_2, ..., m_K) = \{N!/(m_1!m_2!, ..., m_K!)\}p_1^{m_1}p_2^{m_2}...p_K^{m_K}. \qquad \mathbf{(9)}$$

{А.16.2} *Условие* соответствует *совмещению N опытов* (определение 5 [19,31]) или N случайных величин (определение 5 [29,67]), определяющее *N-мерную вероятностную* функцию. Т.е. *число повторений опыта* определяет *размерность промежуточного* распределения.

Формулу **(9)** называют *полиномиальным* распределением. В работах оно рассматривается реже *биномиального* (**некоторые аспекты которого, связанные с его выводом, рассмотрены в [29,11-14]**) распределения: вывод *полиномиального* распределения дается редко (**из цитируемых работ – в [16,65; 20,23], но в некотором сокращенном варианте**), а обычно говорится, что он проводится по аналогии с выводом *биномиального* распределения.

***Замечание* 12.** **1.** **Формула (9) записана для *элементарных* событий в нашем понимании (положение {A.11} [29,41]). Вообще говоря, с позиций, развитых в настоящей работе (замечания 15-16 [29,49-50]), ее следовало бы записать в другом виде. Однако она получается громоздкой (в отличие от (9)) и не является общей.** **2.** **В существующей теории чаще используются *сложные* события $\mathbf{A_k}$: но не совсем понятно в каком понимании – классическом или аксиоматическом – применяется это понятие. Наиболее последовательно аксиоматическая теория применяется в работе [6,184]: рассматриваются $k = 1, 2, ..., K \geq 2$ *возможных* исходов, которым «приписываются» вероятности p_k удовлетворяющие условию нормировки $p_1 + p_2 + ... + p_K = 1$ $\{\mathbf{b}\}$, а «. . . полиноминальное потому, что выражение совпадает с общим членом разложения полинома $(p_1 + ... + p_K)^N$». При этом подчеркивается, что m_k – произвольные неотрицательные числа, удовлетворяющие очевидному условию $m_1 + m_2 + ... + m_K = N$ $\{\mathbf{b^*}\}$. Условие {b} иногда записывают и для событий $\mathbf{A_k}$: из него следует, что события $\mathbf{A_k}$ *несовместимы* (определение 7 [19,36]). Условие {b*}, хотя и очевидное, существенно влияет на решение задачи: об этом попозже.**

Красивая и, в принципе, *правильная* формула, полученная «по образу и подобию» (**как утверждается в существующей теории**) с биномиальным распределением: но правильно, подобно с выводом формулы Бернулли (**об этом – в п.1 примера 15, стр.49**). Однако если формула Бернулли $\mathbf{N!/(m_1!m_2!)}$ $(\mathbf{m_1 = m, m_2 = N - m})$ $(\mathbf{m = 0, 1, ..., N})$ дает строгий порядок расположения коэффициентов, а, следовательно, и вероятностей $P_N(m_1, m_2)$, то *полиномиальные* коэффициенты никакого порядка не обеспечивают.

***Замечания* 13.** **1.** **В работах ничего не говорится о порядке расположения полиномиальных коэффициентов $\mathbf{N!/(m_1!m_2!, ..., m_K!)}$, да и об их вычислении[35] тоже. Например, если $K = 3$ и $N = 2$, то по формуле, учитывая, что условию {b*} соответствуют две группы чисел (2,0,0) и (1,1,0), из которых можно образовать еще по две комбинации (0,2,0), (0,0,2) и (1,0,1), (0,1,1), получим шесть полиномиальных коэффициентов: p_j^2 (j=1,2,3), $2p_1p_k$ (k=2,3) и $2p_2p_3$. *В каком порядке их надо разместить*? Разложение полинома $(p_1 + ... + p_K)^N$ $\{\alpha\mathbf{1}\}$ при значении $n = 2$ вроде бы «дает порядок» p_1^2, $2p_1p_2$, $2p_1p_3$, p_2^2, $2p_2p_3$, p_3^2 $\{\mathbf{b1}\}$, однако он условный, ибо зависит от порядка приведения подобных членов. Во-вторых, даже при этом «порядке», например, при значениях $p_1 = p_2 = p_3 = 1\,/\,3$ получим $1\,/\,9$, $2\,/\,9$, $2\,/\,9$, $1\,/\,9$, $2\,/\,9$, $1\,/\,9$. Варьируя значения вероятностей p_k, получим распределения с другими значениями и порядками расположения вероятностей. Читатель может легко это проверить: например, положить вероятности равными $1\,/\,2$, $1\,/\,4$, $1\,/\,4$ или $1\,/\,2$, $1\,/\,3$, $1\,/\,6$. При $n = 3$ имеем 10 коэффициентов: p_j^3 (j=1,2,3), $3p_1^2p_k$ и $3p_1p_k^2$ (k=2,3), $3p_2^2p_3$, $3p_2p_3^2$ и $6p_1p_2p_3$: полином $\{\alpha\mathbf{1}\}$ «дает порядок»[36] p_1^3, $3p_1^2p_2$, $3p_1^2p_3$, $3p_1p_2^2$, $6p_1p_2p_3$, $3p_1p_3^2$, p_2^3, $3p_2^2p_3$, $3p_2p_3^2$ и p_3^3 $\{\mathbf{b2}\}$. При значениях $n = 4$ и $n = 5$ – 15 и 21 соответственно. Очевидно, что увеличение числа n повторений *опыта*, конечно же, не способствует *улучшению* порядка. При увеличении числа K *элементарных* событий становится только хуже.**

2. **Полином $\{\alpha 1\}$ записывают также в виде [1,74] $(p_1x_1 + p_2x_2 + ... + p_Kx_K)^N$ $\{\alpha\mathbf{2}\}$, т.е. степени полинома K *переменных*. По сути – в N-*мерном* пространстве: реально это не приводит к улучшению порядка, но *именно* эта *трактовка* распределения обычно дается в современной литературе, что *совершенно не верно*. Об этом в п.2.3 примера 14 (стр.46).**

---

[35] Из просмотренных работ, об их вычислении говорится в примере а) [6,60]: метод удобен: он использован в [26,25-26], хотя первоначально мы вычисляли их по-другому

[36] В работе [17,198] они записаны в другом порядке: член p_2^3 стоит перед последним в выражении {b2}, что только подчеркивает отсутствие порядка

При анализе существующего построения *биномиального* распределения [29,11-14] показано: **построить его на основе существующей *теории случайных* величин невозможно, ибо *в ней не существует правил* для такого построения** (**положение {A.2} [29,14]**). Естественно, *не существует* таких *правил* и для построения *полиномиальных* распределений, на основе теории случайных величин. А это приводит к некоторым курьезам, которые определяют неверную трактовку и неправильное понимание сути распределений.

***<u>Пример</u>* <u>13.</u>** **«Существует множество ситуаций, в которых совместное распределение *<u>трех</u>* случайных величин задается *полиномиальным* распределением $\mathbf{P_n(X_1 = k_1, X_2 = k_2, X_3 = k_3)} = \frac{\mathbf{n!} p_1^{k_1} p_2^{k_2} p_3^{k_3} (1-p_1-p_2-p_3)^{n-k_1-k_2-k_3}}{\mathbf{k_1!k_2!k_3!(n-k_1-k_2-k_3)!}}$ (1.7); здесь k_1, k_2, k_3 – целые числа, такие что $k_1 + k_2 + k_3 \leq n$. Например, если $\mathbf{X_1, X_2, X_3}$ представляют собой числа выпадений единицы, двойки и тройки при бросаниях правильной кости, то их совместное распределение задается формулой (1.7) с $p_1 = p_2 = p_3 = 1/6 \ldots$.**

Чтобы получить (маргинальное) распределение величин $(\mathbf{X_1, X_2})$, нужно зафиксировать k_1 и k_2, и просуммировать вероятности 1.7) по всем возможным значениям k_3, т.е. $k_3 = 0, 1, ..., n - k_1 - k_2$. Используя формулу Ньютона, получим триномиальное распределение[37] $P_n(\mathbf{X_1} = v_1, \mathbf{X_2} = v_2) = \frac{\mathbf{n!} {p'}_1^{v_1} {p'}_2^{v_2} (1-p'_1-p'_2)^{n-k_1-k_2}}{\mathbf{v_1!v_2!(n-v_1-v_2)!}}$ (1.8). Суммируя по $v_2 = 0, 1, ..., n - v_1$, находим распределение величины $\mathbf{X_1}$: оно сводится к биноминальному распределению» [6,229].

п.1. Обратим внимание на следующее. *Левая* часть равенства (1.7) соответствует в формуле (9) появлению *трех элементарных* событий, а *правая* – *четырех* (при бросании игральной кости – три *элементарных* и одно *сложное* событие $\mathbf{A = a_4 + a_5 + a_6}$ – выпадение *или* «4», *или* «5», *или* «6»). Аналогично, *левая* часть равенства (1.8) соответствует в формуле (9) появлению *двух элементарных* событий, а *правая* – *трех* (при бросании игральной кости – два *элементарных* и одно *сложное* событие $\mathbf{A_1 = a_3 + a_4 + a_5 + a_6}$). *Можно ли между этими частями ставить знак равенства*? Прежде всего, отметим:

> **{A.17} В приведенном примере (бросание игральной кости) значение $p_4 = 1 - p_1 - p_2 - p_3 = 1/2$ – это вероятность $\mathbf{P(A)}$ *сложного* события A (или $\mathbf{A_1}$). Однако, то, что *его вероятность* можно выразить через вероятности других *элементарных* событий, *вовсе не означает*, что оно *не происходит в испытаниях* (или экспериментах): кстати, событие A (или $\mathbf{A_1}$) будет появляться чаще остальных.**

п.2. Против равенства $p_4 = 1 - p_1 - p_2 - p_3$ (или $p_3 = 1 - p_1 - p_2$) возражений нет: оно *правильно* определяет вероятность *сложного* события A (или $\mathbf{A_1}$). С точки зрения математики *неравенство* $k_1 + k_2 + k_3 \leq n$ вроде бы тоже *правильно*, однако *более точный смысл дает равенство* $m_1 + m_2 + m_3 + m_4 = n$ (условие {b*}). Из него определяется число $m_4 = n - m_1 - m_2 - m_3$ появлений события A (или $\mathbf{A_1}$).

п.3. Таким образом, в существующей теории утверждается: формула (1.7) определяет «*совместное*» распределение 3-х, формула (1.8) – 2-х случайных величин, а биномиальное – распределение одной случайной величины. В соответствии с определениями 'D, D1' [29,16-17], *биномиальное*

---

[37] **По отношению к оригиналу в формуле (1.8) приняты обозначения чисел v_j и вероятностей p'_j, чтобы подчеркнуть ее отличие от формулы (1.7)**

распределение трактуется как *одномерное* (определяется одной координатой), *триномиальное* – как *двумерное*, а *исходное* распределение, определяемое формулой (1.8), – как *трехмерное* распределение.

Из всех *утверждений*, приведенных в примере 3, *справедливо только одно*: *биномиальное* распределение является *одномерным* распределением. Формула (9) содержит *произведение всех* значений *вероятностной* функции, которое характерно *только для системы*, образованной *совмещением* N случайных величин, определяемых *повторением* N раз *опыта с одной одномерной* случайной величиной.

Полиномиальный коэффициент определяет *сумму произведений* значений *вероятностной* функции, имеющих *одинаковые степени*. *Полиномиальные* коэффициенты определяются на основе 1-го *способа преобразования многомерного* распределения *к одномерному* виду с применением *линейной формулы преобразования*: оно рассмотрено ниже. Заодно покажем *связь преобразования сложного* опыта, образованного *совмещением одномерных опытов*, с *преобразованием* распределений *дискретных* случайных величин.

***Пример* 14. Пусть $p(x_k) = P(a_k)\ x_k = \Delta \cdot k$ (k=1,2,...,K) – распределение одномерной случайной величины X, определенной на множестве элементарных a_k событий опыта. Как и в существующей теории событий, рассмотрим повторение опыта N=1,2,... раз: в соответствии с новой исходной системой – экспериментов со случайной величиной X.**

п.1. 1. Будем определять вероятность *одновременного* появления «*и элементарного* события $\mathbf{a_1}$, *и элементарного* события $\mathbf{a_2}$, ..., *и элементарного* события $\mathbf{a_K}$» (т.е. их одновременного появления) в N опытах. В терминах случайных величин: вероятность появления «*и значения* $p(x_1)$, *и значения* $p(x_2)$, ..., и *значения* $p(x_K)$ в N *экспериментах*. *Логическая* связка соответствует *совмещению* N *опытов* (определение 5 [19,31]) или N случайных величин (определение 5 [29,67]).

$\mathbf{a_1a_3}$	$\mathbf{a_2a_3}$	$\mathbf{a_3^2}$	p_1p_3	p_2p_3	p_3^2
$\mathbf{a_1a_2}$	$\mathbf{a_2^2}$	$\mathbf{a_3a_2}$	p_1p_2	p_2^2	p_3p_2
$\mathbf{a_1^2}$	$\mathbf{a_2a_1}$	$\mathbf{a_3a_1}$	p_1^2	p_2p_1	p_3p_1

Рис.14

В результате получим K^N *произведений* вида $\mathbf{a_1^{m_1}a_2^{m_1}\cdots a_K^{m_K}}$ (верхний индекс – показатель степени) или $p_1^{m_1}p_2^{m_2}\cdots p_K^{m_K}$, где m_k – целые числа, такие что $m_1+m_2+...+m_K=N$ (условие {b*}).

При $N=2$ они представляются в виде *двумерной* таблицы или *двумерного* распределения *совмещенной* случайной величины в координатах $X0Y$. Для значений $K=3$, $N=2$ они условно изображены[38] на рис.14 (обозначение $p_k = p(x_k)$ принято для сокращения записи) в виде произведений $\mathbf{a_ka_j}$ (k,j=1,2,3) или значений p_kp_j в точках, определяемых парами $(x_k,\ y_j)$ координат. При $N=3$ – *трехмерной* таблицы или *трехмерного* распределения *совмещенной* случайной величины: положение *произведений* $p_kp_jp_m$ будет определяться *тройками* $(x_k,\ y_j,\ z_m)$ координат соответственно. Его, в принципе, еще можно изобразить, но при $N>3$ – такой «картинки» нам не построить, да что-то она нигде (не только

[38]Расположение столбцов и строк принято в соответствии с возрастанием закона $P(x_n)=\sum_{k=1}^{n}p(x_k)$ распределения величины X (или Y) в положительном направлении оси X (или Y)

в теории вероятностей) и не попадалась ...

{A.18} По сути поставленной задачи, K^N *произведений элементарных* событий представляют собой *сложный* N*-мерный* опыт, образованный *совмещением одного и того же* (простого) *опыта* N раз: в терминах случайных величин – N*-мерное* распределение *системы* случайных величин, образованной *совмещением* N раз *одной и той же одномерной* случайной величины.

2. В принятой теории, из так называемой «теоремы о математическом ожидании произведения случайных величин **X** и **Y**» следует нечто другое: «Теорема верна при любом числе независимых между собой множителей. Требование независимости весьма существенно. Если оно не соблюдено, заключение теоремы вообще несостоятельно; было бы неправильным, например, такое преобразование $M\mathbf{X^2} = M(\mathbf{X}\cdot\mathbf{X}) = M\mathbf{X}\cdot M\mathbf{X} = (M\mathbf{X})^2$, так как оба сомножителя одинаковые, они зависимые» [12,169]. Объяснения, на каком основании одинаковые сомножители «зависимы» и откуда появилась эта теорема, в работах мы не нашли: остаются догадки. Возможно, первое связано с равенством [12,87] $\mu_2(\mathbf{X}') = D(\mathbf{X}') = M(\mathbf{X^2}) - [(M(\mathbf{X})]^2 = \nu_2(M\mathbf{X}) - m_{\mathbf{X}}^2$ **{b3}**; второе – с равенством [12,170] $\mu_{1,1}(\mathbf{X}',\mathbf{Y}') = K(\mathbf{X}',\mathbf{Y}') = M(\mathbf{XY}) - M(\mathbf{X})\cdot M(\mathbf{Y}) = \nu_{1,1}(\mathbf{X},\mathbf{Y}) - m_{\mathbf{X}}m_{\mathbf{Y}}$ **{b4}**. Напомним: если начало координат *не совпадает* с математическим ожиданием, то числовые характеристики $M\mathbf{X^k} = \sum_{i=1}^{K} x_j^k p(x_j) = \nu_k(\mathbf{X})$ называют *начальными* [12,85], а если *совпадает*, то характеристики $M(\mathbf{X} - m_{\mathbf{X}})^k = \sum_{i=1}^{K}(x_j - m_{\mathbf{X}})^k p(x_j) = \mu_k(\mathbf{X}')$ называют *центральными* моментами [12,87] порядка k.

{A.19} *Равенства* {b3}-{b4} определяют *только связь* между *начальными* и *центральными* моментами *одного* порядка, определенными в *разных системах* координат: x и $x' = x - m_{\mathbf{X}}$ в первом случае; x, y и $x' = x - m_{\mathbf{X}}, y' = y - m_{\mathbf{Y}}$ – во втором.

***Есть ли связь равенства* {b3} «*с условием независимости*» (раздел 1.3, стр.30) $p(x_j, y_k) = p(x_j)\cdot p(y_k)$? *Какое отношение имеет равенство* {b4} *к математическому ожиданию произведения случайных величин* X *и* Y?**

Отметим, что такие *трактовки равенств* определяются *отождествлением понятия* случайной величины *с координатами*, которым присваиваются вероятности, и возникшей отсюда коллизией обозначений. Об этом при анализе числовых характеристик случайных величин (часть V исследований).

п.2. То, о чем говорилось в п.1 – это напоминание уже пройденного (раздел 4.1.2 [29,73]), а теперь, собственно, о *преобразовании.*

1. Точки на осях X и Y расположены на равных расстояниях, следовательно, *линейную формулу преобразования* следует взять в виде[39] $y = (z - x)$. При увеличении параметра z прямая линия $y = -x$ последовательно проходит через точки, определяемых парами (x_k, y_j) координат, *сумма номеров* которых отвечает *последовательности* чисел k+j=w=2,3,...,6 **{*}**. Из рисунка 14 видно, что этой же *последовательности* соответствует *сумма номеров элементарных* событий. Соответственно, *суммируются все произведения* вероятностей $p_k p_j = p(x_k)\cdot p(y_j)$ в точках (x_k, y_j), *принадлежащих прямой* линии $y = -z$ *для данного* числа w, определяемого *последовательностью* {*}. Значения *сумм* вероятностей присваиваются точкам $z_{w=k+j} = x_k + y_j$ на оси Z (положение {A.14}, стр.36).

[39] Если $x_k = \Delta_x \cdot k$ и $y_j = \Delta_y \cdot j$, то следует использовать функцию $y = -(\Delta_y/\Delta_x)(z - x)$

{A.20} При числе K *элементарных* событий и *повторении опыта* N раз, *сумма номеров* координат (вероятностей в сокращенном обозначении или элементарных событий) в *произведениях* $p_1, p_2, ..., p_K$ *вероятностных* функций определяется *последовательностью* чисел $w = N, N+1, N+2, ..., NK$ (10). Число членов *полиномиального* (одномерного) распределения – $N(K-1)+1$ (11).

2. Коэффициенты триномиального распределения при значениях $N = 2, 3, 4$ равны: $p_1^2|2p_1p_2|2p_1p_3+p_2^2|2p_2p_3|p_3^2$ **{b5}**; $p_1^3|3p_1^2p_2|3p_1^2p_3+3p_1p_2^2|6p_1p_2p_3+p_2^3|3p_1p_3^2+3p_2^2p_3|3p_2p_3^2|p_3^3$ **{b6}**. $p_1^4|4p_1^3p_2|6p_1^2p_2^2+4p_1^3p_3|12p_1^2p_2p_3+4p_1p_2^3|12p_1p_2^2p_3+6p_1^2p_3^2+p_2^4|12p_1p_2p_3^2+4p_2^3p_3|6p_2^2p_3^2+4p_1p_3^3|4p_2p_3^3|p_3^4$ **{b7}**. Коэффициенты разделены вертикальными линиями. Различие видно уже в числе членов полученного распределения и числе полиномиальных коэффициентов, например, для $N = 2, 3, 4, 5$ имеем **5,7,9,11** и **6,10,15, 21** соответственно. Легко проверить, что полученное распределение содержит все полиномиальные коэффициенты. Во-вторых, число членов полиномиальных (в том числе – биномиального) распределений определяется по одной формуле.

3. В-третьих, есть также определенная аналогия в *расположении* членов *полученного полиномиального* и *биномиального* распределений: *суммы полиномиальных* коэффициентов, определяющих *члены полученного* распределения, *симметричны* относительно *центрального* члена, имеющего *максимальную сумму*. Особенно наглядно это проявляется, если положить вероятности *одинаковыми*. Для иллюстрации в таблицах 1-3 приведены коэффициенты при значениях $K = 2.3.4$ (вероятности элементарных событий $1/K$) и числах $N = 1, 2, \ldots, 5$ *повторений опытов*: коэффициенты для *данного* значения N расположены по строкам таблиц.

		1		1		
		1	2	1		
	1	3		3	1	
	1	4	6	4	1	
1	5	10		10	5	1

Табл.1

				1	1	1				
			1	2	3	2	1			
		1	3	6	7	6	3	1		
	1	4	10	16	19	16	10	4	1	
1	5	15	30	45	51	45	30	15	5	1

Табл.2

						1	1		1	1						
					1	2	3	4	3	2	1					
			1	3	6	10	12		12	10	6	4	1			
		1	4	10	20	31	40	44	40	31	20	10	4	1		
1	5	15	35	65	101	135	155		155	135	101	65	35	15	5	1

Табл.3

Во всех случаях *сумма коэффициентов* по строкам определяется значением K^n, соответственно, значения коэффициентов по строкам следует умножить на значение $1\,/\,K^N$: вероятность $1\,/\,K^N$ следует также из *произведения* (вывод **W.29 [19,41]**) *элементарных* событий при *совмещении* N *опытов* в *предположении*, что каждому *элементарному* событию *опыта* соответствует *один возможный* исход. Каждое из распределений вероятностей *симметрично* относительно *математического* ожидания, которое в *центральной* системе координат равно 0.

Обратим внимание на одну особенность, которая видна из таблиц: в табл.2 есть *одно максимальное* значение *для каждого* N, *совпадающее с математическим* ожиданием, а в табл.1 и 3 – *одно максимальное* значение имеется только *для четных* значений N. Это *отличие характерно* для распределений с *четным* и *нечетным* числом *элементарных* событий.

Выше показано (выводы **W.23-W.24**, стр.40), что *преобразование* определяет *одномерные* распределения независимо от значений чисел K и N, следовательно, следует использовать степень полинома $(p_1 + p_2z+, ... + p_Kz^{K-1})^N$ одной переменной. При $K = 2$ имеем степень $(p_1 + p_2z)^N$ бинома, общий член разложения. которого совпадает с биномиальным распределением: еще одна аналогия.

{A.21} Представление *полиномиального* распределения в виде *разложения* полинома $(p_1 + p_2 z + ... + p_K z^{K-1})^N$ (12) по степеням z *возможно именно потому*, что *исходным* для его построения является N-*мерное* распределение вероятностей *произведения одинаковых одномерных* распределений *системы* N величин, образованной их *совмещением*.

Можно сказать по-другому: разложения полинома (12) *правильно отражает суть преобразования* и определяет *одномерность полиномиального* распределения. Однако по разложению *невозможно* определить точки, которым *принадлежат полученные суммы* вероятностей: они определяются только из преобразования.

Убедиться в правильности положения {A.21} можно умножением членов полинома (рис.14) $p_1 + p_2 \cdot z + p_3 \cdot z^2$ на последовательность p_1, $p_2 \cdot z$ и $p_3 \cdot z^2$: если формально, то просто приписав степени z^{j-1}, z^j, z^{j+1} произведениям в строках и столбцах таблицы с номерами j=1,2,3. Коэффициенты разложения полинома при степенях z^w полностью соответствуют последовательности чисел w=2,3,...,6 (положение {A.20}, стр.46). Отметим следующее: N произведений полиномов $\{\alpha\}$, как и полиномов $\{\alpha^*\}$ тоже определяют N-мерную таблицу. Например, при значениях $K = 3$ и $N = 2$, умножив члены 1-го полинома p_1x_1, p_2x_2, p_3x_3 последовательно на члены 2-го p_1x_1, p_2x_2, p_3x_3 и расположив их произведения в том же порядке, что и на рис.14.

4. Если *элементарные* события *равновероятны*, то имеется аналогия *вычисления коэффициентов* по таблицам с *вычислением биномиальных коэффициентов* по треугольнику Паскаля. Пусть для данного числа K событий известны коэффициенты для строки с номером N: примем для них обозначение N_j^n $(j = 1, 2, ..., n(K-1)+1)$. Для строки с номером $N+1$ имеем $N(K-1)+K$ коэффициентов, которые вычисляются в следующем порядке: первые K – последовательностью возрастающих сумм $N_k^{N+1} = \sum_{j=1}^{k} N_j^N$ (k=1,2,...,K) **{b8}**; последующие $N(K-1)-K$ – суммами $N_{K+k-1}^{N+1} = \sum_{j=k}^{K+j-1} N_j^N$ [k=2,3,..., N(K-1)-K+1] **{b9}**; последние K – последовательностью убывающих сумм $N_{K+k-1}^{N+1} = \sum_{j=k}^{2n+1} N_j^n$ [k=N(K-1)-K+2, N(K- 1)- K+3,..., N(K-1)-K+1] **{b10}**.

Теперь можно начинать с $N = 1$: при переходе к $N = 2$ применяются только формулы {b8} и {b10}. При переходе от значения $N = 2$ к значению $N = 3$ одно значение вычисляется по формуле {b9}: исключение для *опыта* с 2-мя *элементарными* событиями, в котором формула {b9} начинает применяться при 3-ем переходе.

п.3. В принципе, то о чем говорилось в п.2, *не просто аналогия*, а *закономерность*, являющаяся результатом *правильного* решения *одной и той же* задачи: *преобразования* N-*мерного* распределения в *одномерное*. Одна досада: есть «мелкая неприятность», приводящая к серьезной проблеме. Она связана с числами $(m_1, m_2, ..., m_K)$, отвечающими условию {b*}, которые будем называть *допустимыми группами* чисел (определение 1 [26,5] – *допустимые группы* чисел шаров, которые можно составить из N одинаковых шаров).

Как и для решения задачи, рассмотренной в работе [26], для *построения полиномиального* распределения необходимо решить три задачи[40]:

I. Найти все решения уравнения $m_1 + m_2 + ... + m_K = N$ {b*}, т.е. все *допустимые группы* чисел $(m_1, m_2, ..., m_K)$, отвечающих условию {b*}.

II. Определить число *возможных комбинаций* для каждой *допустимой группы* чисел.

III. Для каждой *допустимой группы* определить $1 \leq m_k \leq N$ число

[40] Есть существенное отличие, связанное с постановкой задач, но в данном случае оно не имеет значения

возможных комбинаций, **с учетом различимости групп с одинаковым числом *элементарных* событий** (**координат значений** $p(x_k)$ (**k=1,2,...,K**) **вероятностной функции**), ***без учета порядка их расположения*** **в группе.**

Вычисление *допустимых групп* чисел и определение числа *возможных комбинаций из допустимых групп* решается просто только *при значении* $K = 2$, **ибо в этом случае *допустимы группы* состоят *из пары* чисел** $(m_1,\ m_2)$, **для каждой из которых возможны только 2-е комбинации** $(m,\ N - m)$ **и** $(N - m,\ m)$. **Последовательности** $m = N\,, N - 1\,, N - 2, ..., 0$ **и** $N - m$ **определяют число *возможных комбинации допустимых групп* и порядок их расположения.**

Очаровательная «картинка». Однако, к огромному сожалению, при значениях $K > 2$ **«чудная картина куда-то улетучивается».**

***Полиномиальные* коэффициенты** $N!\,/\,(m_1!m_2!..., m_K!)$ ***непосредственно* решают только *третью* задачу. В работе [6,60]** (**сноска 34, стр.40**) **дан удобный способ решения 2-й задачи: *второй раз* применяется формула вычисления *полиномиальных* коэффициентов в виде** $N!/(k_1!k_2!...k_Z!)$, **где** k_z – **числа *одинаковых* групп чисел** m_k **(включая число групп, состоящих из нулей), входящих *в данную допустимую комбинацию.* Но при рассмотрении полиномиального распределения об этом, почему-то даже не упоминается. Дело в следующем:**

> **{A.22} Решение задач II-III возможно только тогда, когда известны все *допустимые группы* чисел** $m_1, m_2, ..., m_K$ **для данного** N, **т.е. *реальные решения* уравнения** $m_1 + m_2 + ... + m_K = N$ (**условие {b*}**).

Но в теории вероятностей об не говорится: по-видимому, *даже задачи такой не ставилось* – ни в прошлом, ни в настоящем. Об этом говорилось в [26,24]. Хотя в [6,58] и говорится о числе различных решений уравнения, но реально решается задача о числе *различных размещений* при известных числах $m_1, m_2, ..., m_K$. **Для небольших чисел** K **и** N **они легко подбираются, но даже при данном значении** K **для произвольного числа** N **задача не решена. Мы писали [26], что есть идея осуществления решения и даже начато ее воплощение, однако до сих пор так и не продолжили. Казалось, что работу [29] *по созданию новой исходной системы теории* случайных величин, «быстренько закончим», однако ... слишком много вопросов осталось « за кадром, так что конца пока что не видать».**

Из анализа, данного в примере 14, следует:

> **{A.23} *Полиномиальное*** (**в том числе – биномиальное**) **распределение определяется *произведением одинаковых одномерных* распределений системы, образованной *совмещением*** $N \geq 2$ **случайных величин. Оно является *исходным* для построения *одномерного полиномиального* распределения: построение осуществляется 1-м *способом линейного преобразования*** N***-мерного*** **распределения к *одномерному* распределению.**

При *преобразовании одномерных*, а также *непрерывных двумерных* (**многомерных**) распределений (**положения {A.4}, {A.15} стр.12, 41**) мы подчеркивали, что в существующей теории *преобразования* выполнены *правильно*, однако их *трактовка* абсолютно *не соответствует сути преобразования.* На основе анализа, данного в примере 14, очевидно, что этого *невозможно* сказать о *преобразовании многомерных дискретных* распределений с числом *элементарных* событий (**значений вероятностной функции**) *больше двух*, ибо *неверна не только трактовка, но и само преобразование*[41].

Это приводит не только к «курьезам» (**пример 13, стр.43**), но и к серьезным последствиям: *утверждениям* в существующей теории, которые *не имеют отношения к реальности.*

***<u>Пример</u>* <u>15.</u> В [1,79] утверждается: «По существу ... можно доказать аналогичную[42] ... теорему и для полиномиального распределения». Из теоремы следует, что *полиномиальное* распределение *аппроксимируется произведением* K *нормальных* законов. Самого доказательства нет, но оно «встречалось» в литературе: сведения о полиномиальном распределении были необходимы еще при написании работы [26].**

Что-то здесь «не склеивается»: в нашем понимании *произведение нормальных* законов *предопределяет существование произведения биномиальных* распределений. Но в этом случае одно непонятно: 1. *Как из произведения двучленов, получить произведение хотя бы трехчленов*? Можно *перемножить* «сколь душе угодно» биномов, но как ни старайся, получишь *только произведение биномов.*

Однако все по порядку.

Цитата, предваряющая постановку общей интегральной теоремы: п.1.1. «Мы перейдем теперь к формулировке интегральной предельной теоремы в общем случае схемы независимых испытаний. По-прежнему μ_k означает число появления событий $\mathbf{A}_\mathbf{k}^\mathbf{n}$ в n последовательных испытаниях. В зависимости от случая μ_k могут принимать лишь значения $0, 1, 2, ..., n$, причем, так как в каждом испытании возможны k исходов и эти исходы несовместимы, то должно иметь место равенство $\mu_1 + \mu_2 + ... + \mu_k = n$ (1).

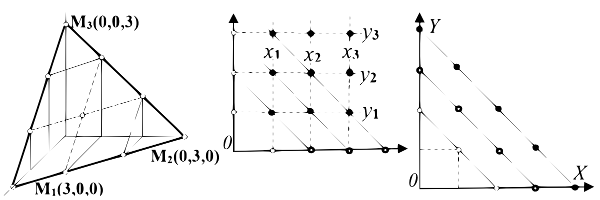

Рис. 15

п.1.2. Станем теперь на величины $\mu_1, \mu_2, ..., \mu_k$ смотреть как на прямоугольные координаты точки в K-мерном евклидовом пространстве. При этом результаты n испытаний изобразятся точкой с целочисленными координатами, не меньше нуля и не большими n; будем называть такие точки целочисленными. Равенство (1) показывает, что результаты испытаний изобразятся не произвольными целочисленными точками в гиперкубе $0 \leq \mu_k \leq n$, а лишь теми из них, которые находятся в гиперплоскости (1). На рис.11 (*примечание автора*: рис.15 слева почти полностью повторяет его: нет только заштрихованной фигуры, изображенной на рис.11) изображено положение возможных результатов в гиперплоскости (1) для случая $n = 3$, $K = 3$. Проведем преобразование координаты по формулам $x_k = (m_k - n \cdot p_k) / \sqrt{(np_k q_k)}$, $q_k = 1 - p_k$» [1,89]. Затем дается формулировка теоремы, а далее следует: «Доказательство и по идее и по осуществлению является почти полной копией рассуждений, проведенных при доказательстве интегральной теоремы Муавра-Лапласа» [1,90].

2. Пока речь идет о событиях (п.1.1) вроде бы все понятно, но п.1.2 вы-

---

[41] Утверждение справедливо и при преобразовании общих дискретных распределений (с разными вероятностями и числами элементарных событий), ибо они преобразуются по рассмотренной схеме

[42] Локальной теореме Муавра-Лапласа

зывает в основном вопросы, например: **2.** ***Почему число μ_k появления событий $\mathbf{A}^{\mathbf{n}}_{\mathbf{k}}$ в n испытаниях, связывают с координатами?*** **3.** ***Какое отношение к координатам имеет равенство*** (1)? **4.** ***Почему координаты преобразуются по формулам, применяемым при преобразовании биномиального распределения?***

Из рассуждений, приведенных в п.1.2 цитаты, следует: число координат связывают с числом событий, т.е. при числе событий $K = 2$ имеем координаты $X0Y$ на плоскости (рис.15а посредине), а равенство $\mu_1 + \mu_2 = n$ определяет прямые линии $y = -x + (n+1)$. Следовательно, на линиях $y = -x+(n+1)$ определены и соответствующие биномиальные распределения. Это противоречит некоторым положениям принятой теории (например, теоремам о математическом ожидании и дисперсии биномиального распределения).

3.Рассмотрим, как определяется формула Бернулли в классической теории.

Будем следовать работе [13,59-66]. Пусть производится $n = 1, 2, 3, ...$ независимых выстрелов по мишени: вероятность попадания (событие $\mathbf{A}^{\mathbf{n}}_{\mathbf{1}}$) в мишень при выстреле в испытании с номером n равна $\mathbf{P}(\mathbf{A}^{\mathbf{n}}_{\mathbf{1}}) = p_1$, а вероятность промаха (событие $\mathbf{A}^{\mathbf{n}}_{\mathbf{2}}$) – $\mathbf{P}(\mathbf{A}^{\mathbf{n}}_{\mathbf{2}}) = p_2$ $(p_1 + p_2 = 1)$. При значении $n = 2$ возможны 4 случая появления событий: 1. Попадание при обоих выстрелах. 2. Попадание при 1-м и промах при 2-м выстреле. 3. Промах при 1-м и попадание при 2-м выстреле. 4. Промах при обоих выстрелах. Получим 4-е произведения [13,60-61] $\mathbf{A}^{\mathbf{1}}_{\mathbf{1}}\mathbf{A}^{\mathbf{2}}_{\mathbf{1}}|\mathbf{A}^{\mathbf{1}}_{\mathbf{1}}\mathbf{A}^{\mathbf{2}}_{\mathbf{2}}$, $\mathbf{A}^{\mathbf{1}}_{\mathbf{2}}\mathbf{A}^{\mathbf{2}}_{\mathbf{1}}$ $|\mathbf{A}^{\mathbf{2}}_{\mathbf{1}}\mathbf{A}^{\mathbf{2}}_{\mathbf{1}}$. Вероятности появления событий равны во всех испытаниях, соответственно, получим 3-х членную $p_1^2|2p_1p_2^2|p_2^2$ формулу Бернулли. Значение $n = 3$ дает 8 произведений, которые определяют 4-х членную $p_1^3|3p_1^2p_2|3p_1p_2^2|p_2^3$ формулу Бернулли. При значении $n = 4$ получим 16 произведений и 5-ти членную $p_1^4|4p_1^3p_2|6p_1^2p_2^2|4p_1p_2^3|p_2^4$ формулу Бернулли, а при произвольном значении n получим 2^n произведений и $n + 1$ член в формуле Бернулли.

Если $n = 2$, то в соответствии с п.1.2 цитаты, результат представляется в 3-х точках M1(2,0), M2(1,1) и M3(0,2) (рис.15а посредине) на линии $y = -x + 2$). Соответственно, при $n = 3$ результат представляется в 4-х точках M1(3,0), M2(2,1), M3(1,2) и M4(0,3) на линии $y = -x + 3$), а при $n = 3$ – в 5-и точках M1(4,0), M2(3,1), M3(2,2), M4(1,3) и M5(0,4) на линии $y = -x + 4$).

На основе таких же рассуждений при числе событий $K = 3$ получим: при значении $n = 3$ – 27 произведений вида $A_j^w A_i^r A_m^t$ (j,i,m=1,2,3) (w,r,t=1,2,3), которые определяют 10 полиномиальных коэффициентов и 7 членов триномиального распределения, а при значении $n = 4$ – 81 произведение вида $A_j^w A_i^r A_m^t A_z^n$ (j,i,m,n=1,2,3) (w,r,t,z=1,2,3,4), определяющие 15 полиномиальных коэффициентов и 9 членов триномиального распределения(п.2.2 примера 4 стр.46).

Вместо «просветления» (т.е. ответов на поставленные вопросы) из анализа следует еще вопрос: **5.** ***Куда в рассматриваемых координатах поместить точки, соответствующие вероятностям произведений событий?***

Вопрос не так прост, ибо разместить произведения надо так, чтобы получить на *основе размещения* вероятностей *произведений полиномиальные* коэффициенты и *полиномиальное* распределение.

{А.24} Очевидно, что геометрическое построение, рассмотренное в работе [1], такой возможности не предоставляет. Построение, позволяющее правильно решить задачу и ответить на все вопросы, основано на положение А.23 (стр.45), и подробно рассмотрено в примере 14 (стр.42).

Построение является общим для полиномиальных (в том числе – биномиального) распределений. Для значений $K = 3$ и $n = 2$ положение произведений, определяемое парами (x_j, y_k) $x_j, y_k = 1, 2, 3)$ координат, обозначено кругами (рис.15а справа), а положение исходных вероятностных функций – тонкими окружностями на осях X и Y. Показаны три (из 5-ти) положения линий $y = -x+m$ (m=2,3,4), которые проходят через точки, определяемые парами: 1) (x_1, y_1); 2) (x_2, y_1) и (x_1, y_2); 3) (x_3, y_1), (x_2, y_2) и (x_1, y_3). Точкам соответствуют произведения вероятностей: 1) p_1^2; 2) p_1p_2 и p_2p_1; 3) p_3p_1; p_2^2 и p_1p_3.

Суммы вероятностей по точкам, принадлежащих данной прямой, определяют коэффициенты триномиального распределения: p_1^2, $2p_1p_2$, $p_1p_3 + p_2^2$, $2p_2p_3$ и p_3^2. Координаты точек, в которой оно определено, равны: $\xi_2 = x_1 + y_1$, $\xi_3 = x_2 + y_1 = x_1 + y_2$, $\xi_4 = x_3 + y_1 = x_2 + y_2 = x_1 + y_3$, $\xi_5 = x_3 + y_2 = x_2 + y_3$ и $\xi_6 = x_3 + y_3$.

Для значений $K = 3$ и $n = 3$ положение произведений $p_ip_jp_k$ определяется тройками (x_i, y_j, z_k) $(x_i, y_j, z_k = 1, 2, 3)$ координат. Положение исходных вероятностных функций на осях X, Y и Z обозначено (рис.15b) тонкими окружностями.

На рис.15b, для примера, также показаны 3 плоскости (из 7-ми), обозначенные треугольниками с штрихпунктирными и штриховыми линиями, и шестиугольником, обозначенным утолщенными штриховыми линиями. Плоскостям принадлежат точки: 2) (x_2, y_1, z_1), (x_1, y_2, z_1) и (x_1, y_1, z_2) – обозначены черными квадратами; 3) (x_3, y_1, z_1), (x_2, y_2, z_1), (x_1, y_3, z_1), (x_1, y_2, z_2), (x_2, y_1, z_2) и (x_1, y_1, z_3) – обозначены черными квадратами; 4) (x_2, y_3, z_1), (x_3, y_2, z_1), (x_1, y_3, z_2), (x_2, y_2, z_2), (x_1, y_3, z_2), (x_1, y_2, z_3) и (x_2, y_1, z_3) – обозначены черными кругами. Точкам соответствуют произведения вероятностей: 1) p_1^3; 2) $p_1^2p_2$, $p_1^2p_2$ и $p_1^2p_2$; 3) $p_1^2p_3$, $p_1^2p_3$, $p_1^2p_3$, $p_1p_2^2$, $p_1p_2^2$ и $p_1p_2^2$, $p_1^2p_3$; 4) $p_1p_2p_3$, $p_1p_2p_3$, $p_1p_2p_3$, $p_1p_2p_3$, $p_1p_2p_3$, $p_1p_2p_3$ и p_2^3; 5) $p_1p_3^2$, $p_1p_3^2$, $p_1p_3^2$, $p_2^2p_3$, $p_2^2p_3$ и $p_2^2p_3$; 6) $p_2p_3^2$, $p_2p_3^2$ и $p_2p_3^2$; 7) p_3^3. Суммы вероятностей по точкам, принадлежащие данной плоскости, определяют коэффициенты полиномиального распределения: 1) p_1^3; 2) $3p_1^2p_2$; 3) $3(p_1^2p_3 + p_1p_2^2)$; 4) $(6p_1p_2p_3 + p_2^2p_3)$; 5)$3(p_1p_3^2 + p_2^2p_3)$; 6) $3p_2p_3^2$ и 7) p_3^3.

В принципе, *правильное геометрическое* построение и *верное* решение задачи можно получить и с применением существующей теории *на основе* понятия *прямого* (декартова) *произведения* в теории вероятностей которое хорошо описано В. Феллером: «Два данных мысленных эксперимента с пространствами элементарных событий $\mathbf{\Omega_1}$ и $\mathbf{\Omega_2}$ можно рассматривать одновременно или последовательно. Это позволяет рассматривать пары возможных исходов, т.е. ввести прямое произведение $(\mathbf{\Omega_1}, \mathbf{\Omega_2})$ в качестве нового пространства элементарных событий» [6,147]. «Для определенности рассмотрим два пространства элементарных событий $\mathbf{\Omega_1}$ и $\mathbf{\Omega_2}$ с точками $\alpha_1, \alpha_2, ...$ и β_1, β_2, ..., имеющими вероятности $p_1, p_2, ...$ и $q_1, q_2, ...$ соответственно. Мы интерпретируем произведение пространств $(\mathbf{\Omega_1}, \mathbf{\Omega_2})$ как пространство элементарных событий, описывающих последовательность двух экспериментов, соответствующих $\mathbf{\Omega_1}$ и $\mathbf{\Omega_2}$» [6,149].

Далее положить: «... вероятности на пространстве определены по правилу $\mathbf{P}(\alpha_{\mathbf{j}}, \beta_{\mathbf{k}} = p_jq_k$» (условие, не имеющее отношение к независимости случайных величин). Т.е. «... два события "первый исход есть $\alpha_{\mathbf{j}}$" и "второй исход есть $\beta_{\mathbf{k}}$ " стохастически независимы». Прямое произведение N случайных величин с одинаковыми вероятностными функциями и определяет N-мерную таблицу произведений, т.е. систему $N \geq 2$ величин, образованную их совмещением. Затем применить к полученному N-мерному распределению то, что в существующей теории называют [12,188; 13,272] «суммой случайных величин», которое реально является 1-м способом преобразования многомерного распределения к одномерному распределению и не имеет никакого отношения к сумме случайных величин (п.6 примера 12 стр.37-38; вывод W.19 стр.38).

п.2. Вернемся к заданным вопросам (стр.49-50). В каких-то из просмотренных работ[43] нашлись предпосылки «доказательства». Пусть *полиномиальное* распределение определяется *повторением* n раз *опыта с одномерной* случайной величиной, *вероятностная* функция которой равна $p(x_k) = \mathbf{P}(\mathbf{a_k})$, $\sum_{k=1}^{K} p(x_k) = 1$, $x_k = \Delta \cdot k$ (k=1,2,...,K), где K – число значений (*элементарных* событий) *вероятностной* функ-

[43] Мы не помним, в какой работе, а искать снова нет смысла

ции. *Полиномиальное* распределение представляется в виде «совместного» распределения «независимых» случайных величин X_j (j=1,2,...,K) (где K – число *элементарных* событий), каждая из которых имеет *биноминальное* распределение с математическим ожиданием np_{2j} и дисперсией $np_{2j-1}p_{2j}$, $(p_{2j-1}+p_{2j}=1)$. По сути, рассматривается *повторение опыта с двумя* событиями: *конкретного* событие $\mathbf{A_j}$ и *противоположного* ему события $\mathbf{\bar{A}_j}=\sum_{k=1}^{j-1}\mathbf{A_k}+\sum_{k=j+1}^{K}\mathbf{A_k}$ (из суммы исключается событие $\mathbf{A_j}$). Пологая[44] $\mathbf{P(A_j)}=p_{2j-1}$ и $\mathbf{P(\bar{A}_j)}=p_{2j}$ (в нашем понимании – опыт с 2-мя элементарными событиями) можно *повторять опыты* с этими случайными величинами. **1.** В таком рассмотрении нет ничего «противозаконного»: в результате *при повторении каждого из этих опытов* n раз, получим K *биномиальных* распределений *с отличающимися* вероятностями. **2.** Опять же, не совершая «никаких противоправных деяний» мы можем рассматривать *вероятностные* функции *произведения или суммы системы*, образованной *совмещением* K случайных величин.

Распределения случайных величин X_j и служат основой для «доказательства выше озвученной теоремы». Остается выяснить два вопросика:

1. *Что же доказывается в теореме*? **2.** *Какое отношение имеет доказанная теорема к полиномиальному* (т.е. при значениях $K>2$) *распределению*?

п.2. Чтобы ответить на 1-й нам даже не надо обращаться к *новой исходной системе* понятий: *достаточно* рассмотреть схему *повторения опытов* с 2-мя событиями $\mathbf{A_j}$ $\{\mathbf{P(A_j)}=p_{2j-1}\}$ и $\mathbf{\bar{A}_j}$ $\{\mathbf{P(\bar{A}_j)}=p_{2j}\}$ $(p_{2j-1}+p_{2j}=1)$ (j = 1,2,...,K) с вероятностями, *изменяющимися от опыта к опыту*, впервые подробно исследованной еще Пуассоном. Иногда эта схема рассматривается в работах, например, [12,68; 13,62]. Далее применяем разложение *произведения биномов* $\Pi_1^n(p_{2j-1}+p_{2j}\xi)$ (n=1,2,...) {**} по степеням ξ. Его называют «формальным» приемом, однако он не такой уж формальный: 1) мы сразу можем сказать, что получим *одномерное* распределение вероятностей (положение {A.21} стр.47); 2) он соответствует преобразованию многомерного распределения в одномерное. Конечно же, при выводе на основе существующей (классической, не говоря уж об аксиоматической) теории событий прием «смотрится» чисто формальным. Для наглядности приведем коэффициенты при значении $n=2,3$ $p_1p_3|p_2p_3+p_1p_4|p_2p_4$ и $p_1p_3p_5|p_2p_3p_5+p_1p_4p_5+p_1p_3p_6|p_2p_4p_5+p_2p_3p_6+p_1p_4p_6|p_2p_4p_6$ {11}: число произведений и число их сумм в коэффициентах соответствует *биномиальному* распределению: при значениях $p_1=p_3=\cdots=p_{2K-1}=q$ и $p_2=p_4=\cdots=p_{2K}=p$ сводится к нему. Вторая закономерность: *сумма номеров* координат (вероятностей, в сокращенном обозначении, номеров элементарных событий) в *произведениях вероятностных* функций определяется *последовательностью* чисел $w=n^2,n^2+1,n^2+2,...,n^2+n$ {***}.

Постановка задачи в п.1 определяет отличие от этой схемы: при *разных* вероятностях p_{2j-1} и p_{2j} мы можем провести *только серию опытов длиной* K, а далее *повторять эту серию.* Выражение {11} определяют коэффициенты при значении $K=3$. При продолжении опытов, коэффициенты определяются последовательным умножением произведения $\Pi_1^3(p_{2j-1}+p_{2j}\xi)$ на биномы $p_1+p_2\xi$, $p_3+p_4\xi$ и $p_5+p_6\xi$: в результате получим *произведение 3-х биномов* $\{(p_1+p_2\xi)(p_3+p_4\xi)(p_5+p_6\xi)\}^2$ во 2-й степени. При повторении серии n раз они будут определяться *произведением 3-х биномов* $\{(p_1+p_2\xi)(p_3+p_4\xi)(p_5+p_6\xi)\}^n$ **{12}** в степени n.

Формула {12} записана в таком виде, чтобы показать: *при повторении серии опытов* n раз *коэффициенты* распределения можно определить *произведением полиномов 3-ей степени*, *коэффициенты* разложения которо-

[44]Обозначения p_{2j-1} и p_{2j} приняты вместо привычных обозначений q_j и p_j потому, что они позволяют определить определенные закономерности, в том числе, связь с преобразованием распределений

го определяются формулами {11}. При *любом* значении n распределение вероятностей будет *одномерным* с числом членов $3n+1$. *Структура* членов будет *подобна* членам *биномиального* распределения, а при значениях $p_1 = p_3 = p_5 = p$ и $p_2 = p_4 = p_6 = q$ *сводится к биномиальному* распределению.

Однако, записав формулу {12} в виде $\Pi_{j=1}^{3}(p_{2j-1}+p_{2j}\xi)^n$, получим произведение 3-х биномиальных распределений, члены которых определяются формулами $C_n^m p_{2j-1}^{n-m} p_{2j}^m$. Если $p_1 = p_3 = p_5 = p$ (j=1,2,3) и $p_2 = p_4 = p_6 = q$, то из формулы {11} получим бином $(q+p\xi)^{n'}$: его разложение соответствует биномиальному распределению при значении $n' = 3n$.

> **{A.24} В отличие от *произведения 3-х биномов*, *произведение 3-х биномиальных* распределений определяет *трехмерную* таблицу *произведений членов* распределений, которую *необходимо преобразовать к одномерному* распределению.**

$p_1^2 p_4^2$	$2p_1 p_2 p_4^2$	$p_2^2 p_4^2$
$2p_1^2 p_3 p_4$	$4p_1 p_2 p_3 p_4$	$2p_2^2 p_3 p_4$
$p_1^2 p_3^2$	$2p_1 p_2 p_3^2$	$p_2^2 p_3^2$

Рис. 16.

В результате преобразования (пример 14, стр.44) получим члены распределения, которые полностью совпадают с коэффициентами, полученные разложением произведения биномов {12} по степеням ξ. «Картинку» *трехмерного* распределения *произведения 3-х биномиальных* распределений изобразить, мягко говоря, сложно: она будет еще и «запутанной». Учитывая, что биномиальное распределение – частный случай полиномиального, для наглядности, на рис.16 изображена *двумерная* таблица *произведения вероятностных* функций биномиальных распределений $(p_1^2 \,|\, 2p_1p_2 \,|\, p_2^2)$ и $(p_3^2 \,|\, 2p_3p_4 \,|\, p_4^2)$.

В общем случае, при K элементарных событиях, получим произведение K биномов $\Pi_{j=1}^{K}(p_{2j-1}+p_{2j}\xi)^n$, которое определяет произведение K биномиальных распределений, члены которых определяются формулами $C_n^m p_{2j-1}^{n-m} p_{2j}^m$ (j=1,2,...,K) при значении $n' = Kn$.

> **{A.25} Из анализа следует: выше «озвученная теорема» доказывает, что при неограниченном увеличении n *произведение* K *биномиальных* распределений аппроксимируется *произведением* K *нормальных* законов.**

По нашему мнению, это очевидно без доказательства, а после преобразования произведения 3-х нормальных законов получим нормальный закон с математическим ожиданием $M(X) = \sum_1^3 M(X_j)$ и дисперсией $D(X) = \sum_1^3 D(X_j)$: соответственно, в общем случае $M(X) = \sum_1^K M(X_j)$ и $D(X) = \sum_1^K D(X_j)$.

Ответ на 2-й вопрос следует из положения **{A.25}**: **«доказанная теорема» *не имеет никакого отношения к полиномиальному* распределению.**

Ответ простой, и дать его можно было сразу после того, как был задан вопрос. А для подтверждения *правильности* ответа, вообще говоря, достаточно сравнить коэффициенты в выражениях {11} с коэффициентами в выражениях {b6}-{b8} (п.2.2 примера 14, стр.45). Можно также сравнить коэффициенты распределений при равных вероятностях, приведенных в табл.1-3 (стр.46).

Мы не знаем, когда доказана эта теорема, но это, не имеет какого-либо особого значения. Более интересны рассуждения и доводы, которые послужили основанием для проведенного «доказательства», однако в просмотренной литературе они нам «не повстречались».

Доказано то, что доказано, однако не совсем то, что утверждается: такое, хотя и редко, но бывает. Анализ, данный в примере 15, показал *полную несостоятельность*, как *доказательства*, так и самой «*теоремы об аппроксимации полиномиального* (**т.е. при значении** $K > 2$) распределения *произведением K нормальных* законов».

2.3. *Математические* ожидания и *дисперсии одномерных* распределений, полученные *обоими способами преобразования двумерного* распределения

В замечании к пояснению 2-х спососов преобразования (**стр.33**), были упомянуты «теоремы» о *сумме математических* ожиданий и *дисперсий*, которые в существующей теории трактуются как «теоремы о *математическом* ожидании и *дисперсии суммы*» (**для дисперсий – «независимых»**) случайных величин. С другой стороны, в примере 12 показано (**выводы W.19-W.20, стр.38**), что *преобразование двумерных* (**многомерных**) *законов* распределения *произведения и суммы* системы, *образованной совмещением* случайных величин, к *одномерному закону* определяет 2-й *способ преобразования*: 1-й *способ к этому отношения не имеет*. Тем не менее, *теоремы о сумме* математических ожиданий и дисперсий *связаны именно* с 1-м *способом преобразования*: для 2-го *способа эти свойства не присущи*. Покажем на простых примерах, что они относятся *именно к преобразованию двумерных* (**многомерных**) распределений.

<u>Пример</u>* <u>16.</u> Вычислим значения функций** $v = x \cdot y$ **и** $w = x + y$ **в точках** $x, y = 0, 1, 2, 3, 4$ **. Нормируем (множитель** $N = 1/100$**) и представим данные в виде таблиц 4 и 5 на рис.16. По сути, табл.4 (раздел 4.1.2 [29,73])) определяет *двумерную вероятностную* функцию *произведения $p_{\Pi}|_{\times}\{(x_j, y_k)\} = p(x_j) \cdot p(y_k)$ **(j,k=0,1,2,3,4) *системы*, образованной *совмещением одномерных* случайных величин** $\mathbf{X}^{(1)}$ **и** $\mathbf{Y}^{(1)}$**. Распределения *одномерных* величин** $\mathbf{X}^{(1)}$ **и** $\mathbf{Y}^{(1)}$ **одинаковы и равны** $p^{(1)}(x_j) = p^{(1)}(y_k) = 0.00; 0.10; 0.20; 0.30; 0.40$**.**

0.00	0.04	0.08	0.12	0.16
0.00	0.03	0.06	0.09	0.12
0.00	0.02	0.04	0.06	0.08
0.00	0.01	0.02	0.03	0.04
0.00	0.00	0.00	0.00	0.00

Табл.4

0.04	0.05	0.06	0.07	0.08
0.03	0.04	0.05	0.06	0.07
0.02	0.03	0.04	0.05	0.06
0.01	0.02	0.03	0.04	0.05

Табл.5

Рис. 17.

Табл.5 определяет (замечание к выводу W.27 [29,64]) *дискретную вероятностную* функцию $p\{(x_j,\, y_k)\}$ ***двумерной* величины** $\Phi(\{(x_j, y_k) \;\rightarrow\; a_{j,k}\}$**. *Вероятностные* функции величин** $\mathbf{X}^{(2)}$ **и** $\mathbf{Y}^{(2)}$ **(верхний индекс – номер величины), полученных *суммированием* значений по *столбам* или *строкам* таблицы, также одинаковы и равны** $p^{(2)}(x_j) = p^{(2)}(y_k) = 0.10; 0.15; 0.20; 0.25; 0.30$**.**

Они приведены и для того, чтобы еще раз можно было легко убедиться в том, что распределения величин $\mathbf{X}^{(2)}$ **и** $\mathbf{Y}^{(2)}$ **не характеризуют распределение (выводы 1-3 [29,62]) *двумерной* величины. В существующей теории**

величины $\mathbf{X}^{(2)}$ и $\mathbf{Y}^{(2)}$ именуют «зависимыми».

Замечание 14. В данном случае *размерность* распределения определяется *числом независимых переменных* (аргументов).

1. Математические ожидания (условно изображены точками на осях) и дисперсии величин $\mathbf{X}^{(1)}$, $\mathbf{Y}^{(1)}$ и $\mathbf{X}^{(2)}$, $\mathbf{Y}^{(2)}$ $(x_j^{(\mathbf{v})}, y_k^{(\mathbf{v})} = 0, 1, 2, 3, 4;\ v = 1, 2)$ равны $M_{\mathbf{X}}^{(\mathbf{1})} = M_{\mathbf{Y}}^{(\mathbf{1})} = 3$, $D_{\mathbf{X}}^{(\mathbf{1})} = D_{\mathbf{X}}^{(\mathbf{1})} = 1$ и $M_{\mathbf{X}}^{(\mathbf{2})} = M_{\mathbf{Y}}^{(\mathbf{2})} = 2.5$, $D_{\mathbf{X}}^{(\mathbf{2})} = D_{\mathbf{Y}}^{(\mathbf{2})} = 1.75$ соответственно.

	$y=\xi-x$									$y_1=(x-\xi')$								
V	0.00	0.00	0.01	0.04	0.10	0.20	0.25	0.24	0.16	0.00	0.04	0.11	0.20	0.30	0.20	0.11	0.04	0.00
	$M_{\mathbf{Z}}^{1} = 6$, $D_{\mathbf{Z}}^{1} = 2$									$M_{\mathbf{Z}}^{2} = 6$ $D_{\mathbf{Z}}^{2} = 2$								
W	0.00	0.02	0.06	0.12	0.20	0.20	0.18	0.14	0.08	0.04	0.08	0.12	0.16	0.20	0.16	0.12	0.09	0.04
	$M_{\mathbf{Z}}^{1} = 5$, $D_{\mathbf{Z}}^{1} = 3{,}5$									$M_{\mathbf{Z}}^{2} = 5$, $D_{\mathbf{Z}}^{2} = 3{,}5$								
Λ	0	1	2	3	4	5	6	7	8	-4	-3	-2	-1	0	1	2	3	4

Табл.6.

При значениях $z = 0, 1, ..., 8-4$ (или $z' = z-4$) прямая $y = z - x$ проходит через точки (x, y) (перемещение функций вдоль оси X), *сумма* j+k=m номеров которых равна *последовательности* чисел $m = 2, 3, 4, 5, 6$ (для функции $y_1 = x - z - m = 4, 8, 12, 8, 4$). Весьма условно прямые $y = z - x$ изображены на рис.16 сплошными линиями. Значения *сумм* вероятностей в точках даны в таблице 6 (2-я строка для табл.4 и 4-я – для табл.5) и представляют распределения *одномерных* величин V и W, полученных *преобразованием двумерных* распределений.

Небольшое отступление. В существующей теории при рассмотрении преобразования непрерывных *двумерных* распределения иногда «попадаются» рисунки (например, рис.16 [1,138] и рис.42 [12,188]), подобные рис.13 (стр.36). Правда, на них прямую линию $y = z - x$ обозначают z=x+y, где z, x и y – «значения» случайных величин Z, X и Y: т.е. обозначение связывают с существующей *трактовкой* случайной величины как *действительной* функции. Однако, при доказательстве «теорем о математических ожиданиях *суммы* и *произведения*» случайных величин подобных рисунков, но с изображением математических ожиданий величин $\mathbf{X}^{\mathbf{v}}$ и $\mathbf{Y}^{\mathbf{v}}$ (v=1,2), а также точек, обозначающих *результат применения* «теорем» нет ни в одной работе. Возможно потому, что возникает вопрос: *где находятся точки, определяемые «теоремами»*? Хотя из пояснений к рис.44 [12,191] можно догадаться, что точка находится на оси X: правда с объяснением «как-то не очень». Из преобразования легко следует: если прямая линия $y = z - x$ проходит через точку $(M_{\mathbf{X}}^{(\mathbf{v})}, M_{\mathbf{Y}}^{(\mathbf{v})})$ (v=1,2), то ее пересечение с осью X определяет точку $M_{\mathbf{Z}}^{(\mathbf{v})} = M_{\mathbf{X}}^{(\mathbf{v})} + M_{\mathbf{Y}}^{(\mathbf{v})}$ на оси Z, в которой находится математическое ожидание величины Z, полученной *преобразованием двумерного* распределения.

Распределения, полученные с применением *линейных формул преобразования* $y = (z - x)$ и $y_1 = (x - z)$ *различны*, но *математические* ожидания и *дисперсии* величин V и W (3-я и 5-я строки табл.6) в обоих случаях *одинаковы* и определяются *суммами* математических ожиданий и дисперсий величин $\mathbf{X}^{(\mathbf{v})}$ и $\mathbf{Y}^{(\mathbf{v})}$ (v=1,2).

2. *Действительные* функции $y = (z - x)$ и $y_1 = (x - z)$ *действительного* переменного, записанные в виде Z=X±Y, в принятой теории трактуются как «*сумма*» и «*разность*» X±Y величин. Теоремы об их математическом ожидании и дисперсии имеют вид $M(X \pm Y) = M(X) \pm M(Y)$ и $D(X \pm Y) = D(X) + D(Y) \pm K(\mathbf{XY})$ (где $K(\mathbf{XY})$ – ковариация; при независимых величинах $K(\mathbf{XY}) = 0$). Результаты, полученные в п.1, противоречат в частности теореме $M(\mathbf{X} - \mathbf{Y}) = M(\mathbf{X}) - M(\mathbf{Y})$, а если величины X и Y «*зависимы*» – то и теореме $D(\mathbf{X} \pm \mathbf{Y}) = D(\mathbf{X}) + D(\mathbf{Y}) \pm K(\mathbf{XY})$. Но и в теоремах о *суммах* характеристик справедлива только *правая часть* равенств: *левые части* $M(\mathbf{X} + \mathbf{Y})$ и $D(\mathbf{X} + \mathbf{Y})$ *не имеют к ним никакого отношения*.

3. В принципе, рассматривать *теорему* $M(\mathbf{XY}) = M(\mathbf{X})M(\mathbf{Y})$ *нет смысла*: *произведения математических ожиданий в природе вообще не суще-*

ствует. Дело в том, что распределение *двумерной* величины *произведения «независимых»* величин $\mathbf{X}^{(1)}$ и $\mathbf{Y}^{(1)}$ является *двумерной* функцией (табл.4, рис.17), а математические ожидания $M_{\mathbf{X}}^{(1)}$ и $M_{\mathbf{Y}}^{(1)}$ характеризуют *положение его центра группирования.* Но *пояснить*, *куда* *поставить* точку, определяющую «математическое ожидание» $M(\mathbf{X}^{(1)}\mathbf{Y}^{(1)}) = M(\mathbf{X}^{(1)})M(\mathbf{Y}^{(1)})$ нам *не* *представляется* *возможным*.

4. Рассмотрим *дискретное* распределение *двумерной* случайной величины с *вероятностной* функцией $p(x_k, y_m) = p_{k,m}$ (k=1,2,...,K; m=1,2,...,M). Вероятностные функции одномерных величин **X** и **Y** определяются выражениями $p(x_k) = \sum_{m=1}^{M} p_{k,m}$ и $p(y_m) = \sum_{k=1}^{K} p_{k,m}$, а их математические ожидания соответственно равны: $M(\mathbf{X}) = \sum_{k=1}^{K} x_k \sum_{m=1}^{M} p_{k,m}$, $M(\mathbf{Y}) = \sum_{m=1}^{M} y_m \sum_{k=1}^{K} p_{k,m}$. Пусть K=3 и M=2. 1-й способ преобразования, определяет коэффициенты вероятностной функции: $P_1 = p_{1,1}$, $P_2 = p_{2,1} + p_{1,2}$, $P_3 = p_{3,1} + p_{2,2}$ и $P_4 = p_{3,2}$, а значения координат равны $\xi_1 = x_1 + y_1$, $\xi_2 = x_2 + y_1 = x_1 + y_2$, $\xi_3 = x_3 + y_1 = x_2 + y_2$ и $\xi_4 = x_3 + y_2$. Умножим коэффициенты коэффициенты P_v на значения координат ξ_v, соответствующие точкам в которых определены коэффициенты, получим: $p_{1,1}(x_1 + y_1) + p_{2,1}(x_2 + y_1) + p_{1,2}(x_1 + y_2) + p_{3,1}(x_3 + y_1) + p_{2,2}(x_2 + y_2) + p_{3,2}(x_3 + y_2)$.

Сложим все значения сумм, относящиеся к координате X или Y. Получим $M(\mathbf{G}) = \sum_{k=1}^{K} x_k \cdot \sum_{m=1}^{M} p_{k,m} + \sum_{m=1}^{M} y_m \cdot \sum_{k=1}^{K} p_{k,m} = M(\mathbf{X}) + M(\mathbf{Y})$.

5. Рассмотрим распределение *трехмерной* случайной величины с *вероятностной* функцией $p(x_k, y_m, z_n) = p_{k,m,n}$ (k,m,n=1,2,3). Вероятностные функции одномерных величин **X**, **Y** и **Z** определяются выражениями $p(x_k) = \sum_{m=1}^{3}\sum_{n=1}^{3} p_{k,m,n}$, $p(y_m) = \sum_{k=1}^{3}\sum_{n=1}^{3} p_{k,m,n}$ и $p(z_n) = \sum_{k=1}^{3}\sum_{m=1}^{3} pk,m,n$, а их математические ожидания равны: $M(\mathbf{X}) = \sum_{k=1}^{3} x_k p(x_k)$, $M(\mathbf{Y}) = \sum_{m=1}^{3} y_m p(y_m)$ и $M(\mathbf{Z}) = \sum_{n=1}^{3} z_n p(z_n)$. Значения коэффициентов вероятностной функции равны: $P_1 = p_{1,1,1}$, $P_2 = p_{1,1,2} + p_{1,2,1} + p_{2,1,1}$, $P_3 = p_{1,1,3}+p_{1,3,1}+p_{3,1,1}+p_{1,2,2}+p_{2,1,2}+p_{2,2,1}$, $P_4 = p_{1,2,3}+p_{2,1,3}+p_{2,3,1}+p_{1,3,2}+p_{3,1,2}+p_{3,2,1}+p_{2,2,2}$, $P_5 = p_{1,3,3}+p_{3,1,3}+p_{3,3,1}+p_{2,2,3}+p_{2,3,2}+p_{3,2,2}$, $P_6 = p_{2,3,3}+p_{3,2,3}+p_{3,3,2}$ и $P_7 = p_{3,3,3}$. Значения координат точек $\xi_1 = 0$, $\xi_2 = z_2 = y_2 = x_2$, $\xi_3 = z_3 = y_3 = x_3 = y_2 + z_2 = x_2 + z_2 = x_2 + y_2$, $\xi_4 = y_2 + z_3 = x_2 + z_3 = x_2 + y_3 = y_3 + z_2 = x_3 + z_2 = x_3 + y_2 = x_2 + y_2 + z_2$, $\xi_5 = y_3 + z_3 = x_3 + z_3 = x_3 + y_3 = x_2 + y_2 + z_3 = x_2 + y_3 + z_2 = x_3 + y_2 + z_2$, $\xi_6 = x_2 + y_3 + z_3 = x_3 + y_2 + z_3 = x_3 + y_3 + z_2$, $\xi_7 = x_3 + y_3 + z_3$. Умножим коэффициенты коэффициенты P_v на значения координат ξ_v, соответствующие точкам в которых определены коэффициенты, и группируем члены сумм, как и выше, получим: $M(\mathbf{G}') = M(\mathbf{X}) + M(\mathbf{Y}) + M(\mathbf{Z})$.

Формулы для суммы дисперсии получается подобным образом. Легко показать, что при замене *формулы преобразования* $y{=}{-}x$ функцией $y{=}x$ формулы не изменятся.

> **{A.26}** ***Теоремы о сумме математических ожиданий и дисперсий одномерных*** **случайных величин «*жестко связаны*» с 1-м *способом преобразования двумерного* (n-мерного) распределения к *одномерному* распределению: *они не имеют никакого отношения к сумме* случайных величин.**

Связь числовых характеристик распределений *произведения* и *суммы*, полученных 2-м *способом преобразования системы*, образованной *совмещением* случайных величин, *с исходными* характеристиками *не определяется.* Некоторые *качественные* выводы следуют из анализа формулы, которая определяет *преобразования закона* распределения *произведения* величин **X** и **Y**: $P_{\mathbf{X}\cdot\mathbf{Y}}(0 \le z < z_d) = \int_0^{z_d} p(x)\,dx \cdot \int_0^{z_d \cdot b/a} p(y)\,dy$ [10,60].

***Пример* 17.** **На рис.17 условно изображены законы распределения $P_1(z)$ и $P_2(z)$ случайных величин (штриховая и штрих пунктирная линии). Сплошной линией изображена функция $P(z)$ *произведения* законов. Так как функции $P_1(z)$ и $P_2(z)$ монотонно возрастающие и ограничены $0 \leq P_1(z), P_2(z) \leq 1$, то при любом значении $0 < z < 1$ выполняются неравенства $P(z) < P_1(z)$ и $P(z) < P_2(z)$.**

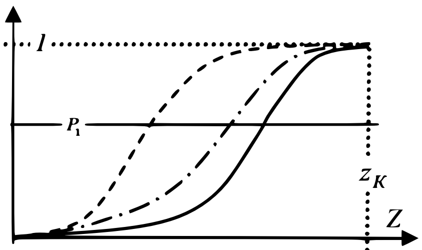

Рис. 18.

Для данного значения $P = P_1$ (сплошная прямая), координата $z = z(P_1)$ больше любого из значений $z_1 = z_1(P_1)$ и $z_2 = z_2(P_1)$ (функции $z(P)$, $z_1(P)$ и $z_2(P)$ обратные к функциям $P(z)$, $P_1(z)$ и $P_2(z)$). Следствие: начальная область медленного изменения функции $P(z)$ шире, а область быстрого изменения уже аналогичных областей каждой из функций $P_1(z)$ и $P_2(z)$.

Соответственно, расстояния между значениями вероятностей в области медленного изменения увеличиваются, а в области быстрого изменения уменьшаются. Т.е. значения вероятностей *произведения* теснее группируются к математическому ожиданию $M(\mathbf{G})$, чем вероятности каждой из функций $P_1(z)$ и $P_2(z)$ к своим математическим ожиданиям $M(\mathbf{X})$ и $M(\mathbf{Y})$. Следовательно, математическое ожидание функции $P(z)$ больше каждого из математических ожиданий функций, а дисперсия - меньше каждой из дисперсий.

Таким образом, из анализа следует:

W.27. ***Математическое ожидание* распределения, полученного 2-м *способом преобразования произведения* величин X и Y *системы*, образованной их *совмещением*, больше *максимального из математических* ожиданий $M(\mathbf{X})$ и $M(\mathbf{Y})$, а его *дисперсия – меньше минимального* из значений $D(\mathbf{X})$ *и* $D(\mathbf{Y})$.**

Из анализа выражения $P_{\mathbf{X+Y}}(0 \leq z < z_d) = \int_0^{z_d} p(x)\,dx + \int_0^{z_d \cdot b/a} p(y)\,dy - P_{X \cdot Y}(0 \leq z < z_d)$ *закона* распределения *суммы* величин [10,61] придем к противоположному результату:

W.28. ***Математическое ожидание распределения, полученного 2-м вариантом преобразования суммы величин* X *и* Y *системы, образованной их совмещением, меньше минимального из математических ожиданий* $M(\mathbf{X})$ *и* $M(\mathbf{Y})$, *а его дисперсия – больше максимального из значений* $D(\mathbf{X})$ *и* $D(\mathbf{Y})$.**

А как же быть с «доказательством озвученных теорем» данным в существующей теории? Как и в примере 15 (**п.1, стр.49**), необходимо разобраться, что *они доказывают* и какое *отношение* имеют эти *доказательства* к *утверждениям* в «теоремах».

***Замечание* 15.** **Нам неизвестно, когда и как появилось представление о сумме и произведении случайных величин, принятое в существующей теории: известно только то, что теоремы об их математических ожиданиях впервые доказаны П. Чебышевым. Доказательства теорем о математических ожиданиях, основаны на равенствах: $M(\mathbf{X} \pm \mathbf{Y}) = \sum_{k=1}^{K} \sum_{n=1}^{N} (x_k \pm y_n) p_{k,n}$, $M(\mathbf{X} \cdot \mathbf{Y}) = \sum_{k=1}^{K} \sum_{n=1}^{N} (x_k \cdot y_n) \cdot p_{k,n}$ (*). Вероятно-**

сти $p_{k,n} = p(x_k, y_n)$ представляются в виде *двумерной* таблицы, т.е. определяют *двумерное* распределение. Суммы по столбцам или строкам таблицы $\sum_{n=1}^{N} p_{k,n} = p_k = p(x_k)$, $\sum_{k=1}^{K} p_{k,n} = p_n = p(y_n)$ (**) определяют распределение *одномерных* величин X и Y с математическими ожиданиями $M(\mathbf{X}) = \sum_{k=1}^{K} x_k \cdot p_k$ и $M(\mathbf{Y}) = \sum_{n=1}^{N} y_n \cdot p_n$ соответственно. С учетом формул (**), выражения для вычисления математических ожиданий принимают вид $\sum_{k=1}^{K}\sum_{n=1}^{N} x_k \cdot p_{k,n}$ и $\sum_{n=1}^{N}\sum_{k=1}^{K} y_n \cdot p_{k,n}$ (***).

С другой стороны, учитывая *сочетательное* свойство *умножения* относительно *суммы* или *разности* $(a \pm b) \cdot c = a \cdot c \pm b \cdot c$ *действительных* чисел, имеем $(x_j \pm y_n)p_{j,n} = x_j \cdot p_{j,n} \pm y_n \cdot p_{j,n}$. Следовательно, *сумма в правой* части 1-го из равенств (*) равна $\sum_{k=1}^{K}\sum_{n=1}^{N}(x_k \pm y_n) \cdot p_{k,n} = M(\mathbf{X}) \pm M(\mathbf{Y})$. Представив вероятности $p_{k,n}$ в виде $p_{k,n} = p_k p_n$, в формулах (***) получим произведения $x_k \cdot p_k \cdot p_n$ и $y_n \cdot p_k \cdot p_n$. Учитывая сочетательное свойство умножения, имеем $(x_k \cdot y_n)(p_k \cdot p_n) = (x_k \cdot p_k)(y_n \cdot p_n)$. Следовательно, произведение в правой части 2-го из равенств (*) равно $\sum_{k=1}^{K}\sum_{n=1}^{N}(x_k \cdot y_n) \cdot p_{k,n} = M(\mathbf{X}) \cdot M(\mathbf{Y})$.

Таким образом, «теоремы доказывают», что *распределительное и сочетательное свойства умножения* для правых частей равенств «работают» и «работают» правильно. Вообще-то, это то, что принимается без доказательств[45] и опять, мягко говоря, «не совсем то что нужно», ибо требовалось доказать, что левая часть равна правой части равенства. Однако с этим, как мы видели выше, возникают проблемы. О проблеме, связанной с *произведением* случайных величин, поговорим при рассмотрении числовых характеристик распределений (**часть V исследований**).

В примере 16 рассмотрены *одинаковые вероятностные* функции случайных величин $\mathbf{X^v}$ и $\mathbf{Y^v}$ (**где v=1,2 – номер случайной величины**). Однако можно взять *две любые разные последовательности* чисел и, проведя операции, выполненные в примере, получить множество других *двумерных дискретных* распределений. В любом случае *математические* ожидания и *дисперсии преобразованных* распределений *определяются суммами математических* ожиданий и *дисперсий* величин $\mathbf{X^v}$ и $\mathbf{Y^v}$.

Можно показать, что при *преобразовании 2-х непрерывных* распределений, получим такие же результаты. В случае *преобразования n-мерных* распределений, *математическое* ожидание и *дисперсия одномерного* распределения, полученного *преобразованием*, будут определяться *суммами математических* ожиданий и *дисперсий одномерных* величин $\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_n}$.

Список литературы

1. Б.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

[45] **Есть очень большие сомнения в том, что это вообще можно «доказать»**

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. Теории событий: критика сцществующих понятий и создание новой исходной системы понятий.

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. И.И. Бондарчук. 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.