

КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИНЖЕНЕРНЫЙ ИНСТИТУТ
Кафедра управления качеством

А.М. ШИХАЛЁВ, Д.П. ВОРОНЦОВ

**КОНТРОЛЬ КАЧЕСТВА ИЗДЕЛИЯ НА ОСНОВЕ
СТАТИСТИЧЕСКИХ ЗАКОНОВ
РАСПРЕДЕЛЕНИЯ**

Казань – 2021

УДК (075.8)311, 53.08

ББК 60.6я73 Г 86

*Принято на заседании учебно-методической комиссии ИИ КФУ
Протокол № 4 от 28 апреля 2021 года*

Под редакцией директора Инженерного института, зав. кафедрой технической физики и энергетики, доктора технических наук, профессора, члена-корреспондента АН РТ **Н.Ф. Кашапова**

Рецензенты:

Кандидат технических наук, доцент кафедры «Автомобильные двигатели и сервис» КНИТУ им. А.Н. Туполева – **КАИ О.Н. Рожко**.

Кандидат физико-математических наук, доцент кафедры «Математическая статистика» КФУ **Чебакова В.Ю.**

Шихалёв А.М., Воронцов Д.П.

Контроль качества изделия на основе статистических законов распределения /А.М. Шихалёв, Д.П. Воронцов. – Казань: Казан. ун-т, 2021. – 144 с.

Управление качеством как процесс предусматривает собой постоянный контроль исследуемых партий производимых изготовителем изделий и услуг, ориентированных на тот или иной сегмент рынка. Поэтому от выбора закона распределения такого рода случайных величин и зависит как оценка, согласно действующим стандартам и другим подзаконным актам, так и выбор того или иного набора формул, которыми пользуются при определении бракованных или не бракованных изделий при оценке уровня качества произведённой продукции.

Учебное пособие предназначено для подготовки и выполнения студентами самостоятельных работ по учебным дисциплинам управления качеством.

В заключение студентам предлагается выполнение контрольного домашнего задания по отдельным вариантам с последующей устной защитой полученных в процессе исследования результатов.

Таким образом, настоящее учебное пособие адресовано, в первую очередь, студентам таких специальностей, как «Управление качеством», «Инновационные технологии управления робототехническими системами» и т.д., а также широкому кругу читателей, интересующихся указанными проблемами.

© **Шихалёв А.М., 2021**

© **Воронцов Д.П., 2021**

© **Казанский университет, 2021**

ОГЛАВЛЕНИЕ

	Стр.
ВВЕДЕНИЕ	6
1. ЗАКОН РАСПРЕДЕЛЕНИЯ БЕРНУЛЛИ	9
1.1. Распределение Бернулли	9
1.2. Формула Бернулли	15
1.3. Наивероятнейшее число наступлений событий	19
1.4. Наиболее характерный пример распределения Бернулли	26
2. О СЛУЧАЙНЫХ ВЕЛИЧИНАХ И ИХ ЗАКОНАХ РАСПРЕДЕЛЕНИЙ	28
3. БИНОМИАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ	32
3.1. Вероятность при бросании игральных костей	34
3.2. Основные формулы комбинаторики	38
3.3. Определение биномиального распределения	41
3.4. Свойства биномиального распределения	45
3.5. Проверка применения нормального закона распределения	58
3.6. Расчёты критериев согласия	62
3.7. Результаты сравнения двух законов распределения	69
4. ЗАКОН РАСПРЕДЕЛЕНИЯ ПУАССОНА	73
4.1. Варианты творческого участия студентов-молодых специалистов	73
4.2. Распределение случайной величины согласно закону Пуассона	76
4.3. Процесс сравнения законов распределения биномиального закона, законов Пуассона и Гаусса	84
4.4. Общие рекомендации применению законов распределения случайных величин	96
5. ЗАКОН НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ГАУССА	101
5.1. Общие положения	101
5.2. Закон нормального распределения Гаусса	104
5.3. Порядок применения закона нормального распределения	110

6. ПРОВЕРКА ОЦЕНКИ Э. ДЕМИНГА КВАЛИФИКАЦИИ СПЕЦИАЛИСТОВ ПОДОПУЩЕННЫМ ИМИ ОШИБКАМ	118
КРАТКИЕ ВЫВОДЫ	132
ЗАДАНИЯ НА САМОСТОЯТЕЛЬНУЮ РАБОТУ	132
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	138

СПИСОК УСЛОВНЫХ ОБОЗНАЧЕНИЙ И СОКРАЩЕНИЙ

ВР – вариационный ряд;

СВ – случайная величина;

ЗР – закон распределения;

НЗР – нормальный закон распределения случайной величины;

БЗР – биномиальный закон распределения случайной величины;

ЗРП – закон распределения случайной величины Пуассона; (до 24 С.)

СLp – среднее значение в p-картах контроля качества;

UCLp – верхняя граница в p-картах контроля качества;

LCLp – нижняя граница в p-картах контроля качества;

χ^2 – критерий Пирсона;

Rom – критерий Романовского;

h – шаг для построения вариационного ряда;

R – размах выборки для построения вариационного ряда;

СС – статистическая совокупность;

УП – учебное пособие.

ВВЕДЕНИЕ

В данном пособии будут рассмотрены наиболее часто встречающиеся распределения случайной величины (СВ), что весьма актуально при организации контроля качества продукции и услуг, согласно предъявляемым им требованиям по действующим стандартам, ГОСТам и другим подзаконным актам.

Как требует утверждённая рабочая программа (учебной) дисциплины (РПД) «Всеобщее управление качеством», тема №8 посвящена изучению именно конкретным методам контроля качества, выбор из которых во многом определяется не только самим характером контроля качества, но и наличие исходной статистической информации при разработке стратегии контроля.

Основной характеристикой партии изделий (продукции) [1. - С. 84] при контроле по альтернативному признаку является генеральная доля дефектных изделий $q = D/N$, где D – число дефектных изделий в партиях объёмом N изделий. В практике статистического контроля генеральная доля q по понятным причинам НЕИЗВЕСТНА, и оценивается она по результатам контроля случайной выборки объёмом « n » изделий, из которых « m » штук являются дефектными. В этой связи необходимо кратко рассмотреть виды планов контроля качества.

Под планом статистического контроля понимается система правил, указывающих методы отбора изделий для проверки, и условия, при которых партию следует принять, забраковать или продолжить контроль наряду с выпуском продукции. При этом различают следующие **виды** планов статистического контроля партии продукции по альтернативному признаку:

1. Одноступенчатые планы. Если среди « n » случайно отобранных изделий число дефектных m окажется не больше приёмочного числа c ($m \leq c$), то партия принимается, иначе – бракуется.

2. Двухступенчатые планы. Если среди n_1 случайно отобранных изделий, число « m_1 » окажется не больше приёмочного числа c_1 ($m_1 \leq c_1$), то партия принимается. Если ($m_1 \geq d_1$), где d_1 – браковочное число, то партия бракуется. Если же ($c_1 < m_1 < d_1$), то принимается решение о взятии второй выборки объёмом « n_2 ». Тогда, если суммарное число дефектных изделий в двух выборках ($m_1 + m_2$) $\leq c_2$, то партия принимается, в противном случае – бракуется [там же. – С. 85].

3. Многоступенчатые планы также существуют как продолжение двухступенчатых, так и последовательный план (или последовательный контроль). Однако, важным является то, что при использовании методов выборочного контроля решение о качестве всей партии по данным выборочных наблюдений, оформленных, например, в традиционной форме контроля качества продукции и услуг в виде p -контрольных карт [4].

На данном этапе для нас важным является то, что создание рабочей выборочной контрольной совокупности, так или иначе связано с установлением закона распределения элементов выбранных для контроля единиц изделий (продукции). Тогда вопрос так или иначе возникнет о его характере. Чаще всего, это, уже знакомый всем, нормальный закон распределения (НЗР), биномиальный закон распределения (БЗР), или, для так называемых «редких событий», - закон распределения Пуассона (ЗРП) [3. – С. 213 - 222]. С другой стороны, контролируемые изделия, составляющие исходную, то есть исследуемую нами, выборочную совокупность, могут подчиняться чисто теоретически и другим законам распределения (равномерному, геометрическому, гипергеометрическому и другим), всё же рассмотрим часть из них в следующей последовательности: формы (формулы), но отнюдь не закон распределения СВ, Бернулли, что окажется, как сможете убедиться сами, окажется весьма полезным для последующего восприятия материала, и далее: биномиальный закон распределения, закон распределения Пуассона, а в заключение - и известный нам ещё со школьных времён нормальный закон распределения случайной величины (НЗР) К.Ф.

Гаусса. Поэтому последовательно и в меру кратко далее рассмотрим особенности каждого из предложенных здесь методов.

Отметим только, что в источнике [15] изложены самые современные методы управления системами и процессами. Такая новая парадигма управления людьми, системами и процессами по своей природе является, безусловно, новым заметным достижением как в деле управления производством, так и в деле контроля качества выпускаемой производством продуктов и услуг. И если многие исследователи делали акцент именно на процесс выявления бракованных изделий в контрольных партиях, то Эдвардс Деминг значительно расширяет само понятие качества контроля до уровня управленческой деятельности в целом, с чем, вне всякого сомнения, будет полезным ознакомиться и нашим студентам.

А если следовать так называемым «классическим» традициям контроля на производстве той или иной продукции, тогда будет необходимым признать и два рода принципиально противоположных случаев последствий контроля: не допустить при этом ошибку 1-го рода (не пропустить бракованное изделие) и всячески избегать ошибки 2-го рода (не забраковать объективно годное изделие).

Заметим также, что если основной материал данного пособия предполагается естественным образом построить на примерах чисто технологического контроля качества, то в заключительной его части предполагается предпринять попытку расширения предмета исследования до уровня социального и социально-правового статуса участвующих в производственном процессе работников отдельных категорий, то есть хотя бы частично расширить исследуемую проблему до уровня менеджмента качества. Тем более, что внимание проблеме распределения случайных величин и законов из распределения, в том числе и в социальной сфере, о чём также будет отмечено далее, в настоящее время уделяется весьма немалое, о чём свидетельствуют работы [29] – [40], приведённые в библиографическом списке в конце данного учебного пособия.

1. ФОРМЫ БЕРНУЛЛИ

Это весьма важная формула и последующие, вытекающие из неё соотношения. Метод Бернулли рассмотрим по источнику [2], тогда как далее метод Пуассона – по источнику [3. – С. 213 - 222]. А после ознакомления с этими ЗР СВ обратимся и к НЗР по источнику [3. -С. 106 - 139]. Приведённые здесь ссылки автора на указанные первоисточники определяются, главным образом, тем, что в них, на взгляд автора, приведённые выше методы изложены наиболее полно и в меру просто. Поэтому рассмотрим каждый метод по отдельности, а потом приведём необходимые, то есть полезные для послевузовской практики, соответствующие обобщения.

1. 1. Формы закона или схемы Бернулли

Сущность распределения элементов исследуемых статистических совокупностей методом Бернулли приводятся в [2] в двух главах. Раскрытие содержания метода Бернулли производится в двух главах 5 и 6 приведённого здесь первоисточника. При этом содержание приводимых здесь глав следующее. Глава 5: 1) примеры; 2) формула Бернулли; 3) наивероятнейшее число наступления события [2. - С. 49 – 62]. Глава 6: 1) содержание теоремы Бернулли; 2) доказательство теоремы Бернулли [2. - С. 62 – 71].

Раскрытие содержания метода в данном источнике, выдержавшего с 1976 г. за 30 лет 8 изданий с исправлениями [2. – С. 5] начинается с одного из примеров.

Для начала приведём лишь два, достаточно характерных, примера, что, на наш взгляд, окажется вполне достаточным, чтобы понять суть рекомендуемого здесь к рассмотрению метода.

Пример 1 (там же. - С. 59 – 60). Волокна хлопка такого-то сорта в среднем на 75% имеют длину, меньшую 45 мм, и на 25% - длину большую

(или равную) 45 мм. Найти вероятность того, что среди **трёх** наудачу взятых волокон **два** будут короче, а **одно** длиннее 45 мм.

Примечания: Уже известны доли того и другого сорта. Аи B – есть **ВЕРОЯТНОСТИ** выбора нами партию из 3-х волосков, взятыми нами при каждом испытании, и которое мы задаём для решения поставленной задачи **сами**. Далее так. Согласно [2. – С. 49 – 50] предлагается следующий ход рассуждений.

Обозначим в рамках отдельного события (каждый раз берём строго по 3 колоска) – выбор волокна с длиной < 45 мм – через A , а другое событие (волокно с длиной > 45 мм) через B . Тогда из самой постановки задачи вполне ясно, что вероятности тех событий P нам уже известны ещё на этапе постановки задачи (для нас они и есть т.н. **ИСХОДНЫЕ ДАННЫЕ**):

$$P(A) = \frac{3}{4} = 0,75; P(B) = \frac{1}{4} = 0,25.$$

Условимся далее схемой AAB обозначать сложное событие согласно условию задачи. Тогда должны иметь место и такие *сложные события*, как собственно AAB , BBA и ABA и т.п. То есть нужно вычислить вероятность события C , состоящего в том, что из трёх волокон два окажутся короче 45 мм. Для этого, очевидно, должна осуществиться одна из возможных схем:

$$AAB, ABA, BAA \tag{1}$$

Так как **ЛЮБЫЕ** два из этих трёх результатов **несовместимы** между собой, то по правилу сложения вероятностей можем составить следующее сложное, логическое (имеется в виду не чувственную триаду познания «ощущение-восприятие-представление», но её логическую часть «понятия-суждение-умозаключение») по своей природе, суждение:

$$P(C) = P(AAB) + P(ABA) + P(BAA),$$

то есть в каждом наблюдении может быть **ИЛИ** событие AAB , **ИЛИ** событие ABA , **ИЛИ** событие BAA (схема логического сложения по известной схеме «ИЛИ»).

Отметим также попутно, что в целом и главным задачей науки является получение *новых знаний* не как чисто технического (механического) явления, но и как философской категории. Поэтому в данном месте изложения будет уместным напомнить, что именно новые знания (см. предыдущую формулу) можно подчас получить даже не на научном, а на умозрительном (то есть обыденном) уровне познания. Так, по крайней мере, нам демонстрирует содержание примера 1.

С другой стороны, вероятность каждой из схем (1) по правилу умножения вероятностей для независимых событий представится как произведение трёх множителей, из которых два равны $P(A) = 3 / 4$, а один $P(B) = 1 / 4$. Таким образом, вероятность каждой из ранее предложенных нами трёх схем (1), полученных нами на даже умозрительном уровне равна

$$P(AAB) = P(ABA) = P(BAA) = \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right) = \frac{9}{64},$$

и, следовательно, по известной теореме сложения вероятностей несовместных событий по схеме «ИЛИ 1-е, ИЛИ 2-е, ИЛИ 3-е»:

$$P(C) = \left(\frac{9}{64}\right) + \left(\frac{9}{64}\right) + \left(\frac{9}{64}\right) = 3 \frac{9}{64} = \left(\frac{27}{64}\right) \approx 0,422 \text{ (или 42,2\%)},$$

что и является решением поставленной задачи.

В данном случае, согласно терминологии [13. – С.5], применяется *комбинаторный принцип умножения вероятностей*, если искомая вероятность события $P(C)$ может осуществляться, как будет показано ниже, C_n^k способами – по числу сочетаний из n (здесь $n = 3$) по k (здесь $k = 2$), – см. формулу (3а). Действительно, по полученной ранее формуле (3а) получим: $C_3^2 = (1 \cdot 2 \cdot 3) / [1 \cdot 2 \cdot (3 - 2)!] = 3$ (комбинации) – уже в результате так называемого «научного познания», в отличие от «познания обыденного (умозрительного)», что для нас в методологическом плане весьма важно.

Для справки, если кто подзабыл, то: $1! = 1$; $0! = 1$; (любое число)⁰ = 1.

Таким образом, по данным и решению примера сделаем следующее обобщение: 1) вероятности наступления ожидаемого события A и не

наступления его в виде события B в виде значений их вероятностей $P(A) = 0,75$ и $P(B) = 0,25$ – соответственно, нам были *известны априорно*, то есть до нашего опыта (иными словами, из нашего ретроспективного, то есть предшествующего опыта. Иными словами, мы, оставаясь лишь на обыденном (умозрительном) уровне познания, :1) представили и формализовали (эксплицировали) число возможных комбинаций исследуемых величин; 2) сами подсчитали число возможных комбинаций, которое затем, правда, *верифицировали* (уточнили) ход решения по приведённой ниже формуле (3а), полученной кем-то до нас уже в результате именно *научного познания* в области математики под названием «комбинаторика».

Если же сразу воспользоваться приведённой также ниже формулой Бернулли (4) для общего числа отобранной партии из волосков числом $n = 3$, а также наступлением при этом двух событий A ($k = 2$), то легко получим ту же самую вероятность:

$$P(C) = C_n^k p^k (1 - p)^{n-k} = C_3^2 p^2 (1 - p)^{3-2} = 3 (0,75)^2 (1 - 0,75)^1 = 0,422.$$

Пример 1 характерен тем, что мы получили один и тот же результат $P(C) = 0,422$ как на умозрительном, так и на научном уровнях познания – приобретения новых знаний как философской категории. Первый раз мы на умозрительном уровне представили все три возможные комбинации из отобранной партии в три нити – см. выражение (1), а затем то же получили уже на научном уровне по формуле (3а):

$$C_3^1 = \frac{3!}{1! (3 - 1)!} = 3$$

возможные комбинации, согласно исходному уравнению (1).

Пример 2 (там же. - С. 50 – 52), который по своей природе окажется немного посложнее 1-го примера. Итак, за много лет наблюдений: на каждые 1000 новорожденных приходится 515 мальчиков и 485 девочек. В некоторой семье шестеро детей. Найти вероятность того, что среди них не более 2-х девочек. Здесь же заметим, что исходные вероятности известны. А 6 детей, среди которых НЕ БОЛЕЕ 2-х девочек – уточнили условие задачи *мы сами*.

В общем, НАДО, чтобы в семье было либо 0, либо 1, либо 2 девочки. Вероятность этих ЧАСТНЫХ событий обозначим как P_0 , P_1 , P_2 . Тогда по правилу сложения вероятностей искомая вероятность

$$P = P_0 + P_1 + P_2 \quad (2)$$

Проще всего начать с вычисления P_0 . Если рассматривать рождение мальчика или девочки как независимое от рождения остальных детей, то по правилу умножения вероятностей того. Что все 6 детей окажутся мальчиками, равна произведению шести множителей, равных 0,515:

$$P_0^{\text{рожд } 0 \text{ дев}} = (0,515)^6 \approx 0,019.$$

Понятно, что число таких комбинаций будет единственно возможной:

$$P_0 = 1 \cdot P_0^{\text{рожд } 0 \text{ дев}} = P_0^{\text{рожд } 0 \text{ дев}} = (0,515)^6 \approx 0,019.$$

А вот вычисление P_1 может произойти, что вполне очевидно, уже не одним, а ШЕСТЬЮ различными способами. Рассмотрим один из них, например, когда девочка рождается, скажем, например, 4-м ребёнком (или 1-м, или 2-м ребёнком – какая, в сущности, разница?). Тогда по правилу умножения вероятностей очевидно независимых друг от друга событий получим:

$$P_1^{\text{рожд } 1 \text{ дев}} = (0,515)(0,515)(0,515)(0,485)(0,515)(0,515) \approx 0,018.$$

Но девочка в семье может появиться 1-й или 2-й, или ... 6-й – всего шесть возможных вариантов, и полученную нами вероятность $P_1^{\text{рожд } 1 \text{ дев}}$ надо сделать в 6 раз большей:

$$P_1 = 6 \cdot P_1^{\text{рожд } 1 \text{ дев}} = 6 \cdot 0,018 = 0,108.$$

А поскольку девочек по условию должна быть «не менее двух», то теперь ещё в дополнение к предыдущим нашим прежним соображениям, в том числе и формализованным, с настоятельной необходимостью, надо вычислить ещё к тому же и вероятность наступления события P_2 . Понятно, какими бы по порядку девочки не родились, вероятность P_2 по правилу умножения вероятностей независимых событий для одной из возможных комбинаций будет равна следующему значению (на сей раз округлим до

тысячных для достижения более точного итогового результата), снова используя результаты своего изначального умозрительного по своей природе познания, получим очевидное (причём, неважно, какими по очереди девочки в исследуемой семье родились):

$$P_2^{\text{рожд } 2 \text{ дев}} = (0,515)^4 (0,485)^2 \approx 0,0165.$$

Но полученное число, в свою очередь, надо умножить на число всех возможных *разновидностей* рассматриваемого типа (2 девочки из 6-ти). Итак, предварительно подытожим: мы с настоятельной необходимостью должны помнить ещё и про число разновидностей. Следовательно, для этого надо учесть число различных способов выбора двух детей из 6-ти имеющихся. Однако, число возможных комбинаций типа «два из шести» на *умозрительном уровне* познания, интересующего нас объекта исследования уже на простом умозрительном уровне познания, оценить, согласимся, уже непросто. Поэтому воспользуемся ранее, до нас накопленными и формализованными средствами, поскольку объективно требуется уже уровень *научного познания*. Например, той же формулой (3а). Вот и воспользуемся ей, ибо ранее установлено, что искомое число есть ЧИСЛО СОЧЕТАНИЙ из *шести по два*, то есть по формуле (3а) число сочетаний при $k = 2$ из $n = 6$ случаев (или $C_n^k = C_6^2 = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 / [1 \cdot 2 \cdot (6 - 2)!] = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 / [1 \cdot 2 \cdot (1 \cdot 2 \cdot 3 \cdot 4)] = 15$ (комбинаций). Вероятность каждой из которых нами уже вычислена и равна 0,0165.

Тогда P_2 из формулы (2) равна:

$$P_2 = 15 \cdot P_2^{\text{рожд } 2 \text{ дев}} = 15 \cdot 0,0165 \approx 0,248.$$

В конечном итоге по правилу сложения вероятностей независимых событий по схеме «ИЛИ» в итоге получим:

$$P = P_0 + P_1 + P_2 = 0,019 + 0,108 + 0,248 = 0,375 \text{ (или 37,5\%)}.$$

То есть немного реже, чем в 4-х случаях из 10-ти (40,0% или 0,400), с вероятностью $P \approx 0,375$, в таких семьях будет ТОЧНО не более двух девочек из шести детей (при условии наблюдения данного процесса, правда, много и

много раз, т.е для которого число испытаний n будет стремиться в пределе к бесконечности (или $n \rightarrow \infty$). Задача решена.

Рассмотрим далее известной формулы или формы Бернулли на предмет их применения в процессе контроля качества производимых изделий и услуг.

1.2. Формула Бернулли

Описание и вывод формулы Бернулли приведён в том же источнике на стр. 52 – 54 как обобщение рассмотренных двух примеров, и может быть выведена следующим образом.

Итак, в примерах 1 и 2 приводились примеры ИСПЫТАНИЙ. Испытание – это осуществление некоторых условий, при наличии которых может наступить интересующее нас событие. Теперь – о одной из главных схем ТВ как математической науке. Эта схема состоит в том, что рассматривается ПОСЛЕДОВАТЕЛЬНОСТЬ *взаимно независимых* испытаний, в каждом из которых событие A может наступить, а может и не наступить (швейцарский учёный Яков Бернулли, 1654 – 1705 гг.). Итак, обобщающее примеры 1 и 2 имели следующее обобщённое задание: «При некоторых условиях вероятность появления события A в каждом испытании равна p ; найти вероятность того, что серия из n независимых испытаний даст k появлений и $(n - k)$ непоявлений события A ».

Событие, вероятность которого надо найти, тем временем (на материалах примеров 1 и 2), в свою очередь, распадается на ряд разновидностей: каждая из них требует наступления n определённых результатов – k появлений и $(n - k)$ непоявлений события A . По правилу умножения вероятность каждой определённой разновидности:

$$p^k(1 - p)^{n-k}.$$

В свою же очередь, число различных возможных разновидностей равно числу различных групп по k испытаний в каждой, которые можно составить из n испытаний, то есть равно C_n^k , что уже при решении примеров 1 и 2

применяли ранее. То есть, применяя правило сложения и известную формулу из комбинаторики для числа сочетаний C_n^k :

$$C_n^k = \frac{n(n-1)\dots(n-(k-1))}{k(k-1)\dots 2 \cdot 1},$$

мы находим, что искомая вероятность k появлений события A при n независимых испытаниях равна

$$P_n(k) = \frac{n(n-1)\dots(n-(k-1))}{k(k-1)\dots 2 \cdot 1} p^k(1-p)^{n-k}, \quad (3)$$

что и решает поставленные задачи, отражённые в примерах 1 и 2.

Обозначающий нами число сочетаний дробный множитель в формуле (3) может быть преобразован в более удобный для вычислений вид (3а):

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (3a)$$

что для искомой вероятности $P_n(k)$ вычисления вероятностей в канонической формуле Бернулли примет окончательный вид:

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k(1-p)^{n-k}. \quad (4)$$

Поэтому и формула (3), и формула (4) так и называются – *формулами (или формами) Бернулли*.

Однако, ещё раз, уже более подробно, для примера 2 вычислим число возможных комбинаций C_6^2 , где $k = 2$, а $n = 6$ по формуле (3а):

$$C_n^k = C_6^2 = \frac{6!}{2!(6-2)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}{1 \cdot 2 \cdot (1 \cdot 2 \cdot 3 \cdot 4)} = \frac{30}{2} = 15 \text{ (комбинаций)}.$$

Затем по формуле (4) с учётом результатов расчётов числа возможных комбинаций по формуле (3а) получим:

$$P_n(k) = C_6^2 p^2 (1-p)^{6-2} = 15 \cdot (0,485)^2 \cdot (0,515)^4 = 0,248.$$

Тогда для условий примера 2 с вероятностью P_2 , полученной в конечном итоге путём именно не простого, обыденного, но *научного познания*, сложим уже известные нам из прежних обыденных умозаключений P_0 , P_1 и P_2 – из научного познания. Получим: $P = P_0 + P_1 + P_2 = 0,019 + 0,108 + 0,248 = 0,375$. То есть, по формуле (4) мы отыскали в примере 2 именно $P_2 =$

$P_6 (k = 2) = P_6^2$. Заметим кстати, что если бы вероятности P_0, P_1 и P_2 были вычислены также по формуле (4), результаты наших вычислений остались бы теми же самыми. Вот если бы условием задачи было «рождение РОВНО 2-х девочек из 6 родившихся», тогда бы только лишь выражением (3) или (4) для вычисления вероятности $P(2)$ были бы просто такими: $P(2) = 0,248$ (или 24,8%).

Ну а теперь приведём из сделанного нами один такой пример.

Пример для выпускников (на основе примера 2) для оценки контроля качества по следующему условию на будущем их месте работы. Вы предлагаете начальнику производства по своей инициативе следующее предложение, которое не заменяет существующий процесс контроля качества изделий, но *дополняете* его свои следующим управленческим предложением: «Известно из предыдущего по предприятию опыта (а его ещё надобно вам, будущим специалистам, как-то предварительно добыть!), что доля годных изделий (оценочная вероятность) $p = 0,8$; вероятность появления брака $(1 - 0,8) = 0,2$. Если произведём отбор из 5-ти образцов готовой продукции в день в течение недели (5 рабочих дней), какова вероятность появления хотя бы одного бракованного изделия?».

Решение производится по той же формуле (4) и для $C_5^1 = 5$ по формуле (3а). За день получим $P_5(1) = 0,064$ или 6,4 %; за 5 дней 32,0%, то есть шансы обнаружить среди готовых изделий случаи брака составили бы примерно 1/3 от объёма всех произведённых предприятием изделий (при условии, конечно, в пределе - бесконечного числа проводимых вами опытов). Таким образом, поставленная вами задача вами же и решена.

Пример 3. Тоже достаточно характерная задача по непосредственному применению правила Бернулли [там же. - С. 54 – 55] в виде формулы (4).

Как и стало обычным, нам было уже кое-что про общую вероятность известно (как и здесь), менялись только условия наступления (ненаступления) события A , что следует из условия следующей задачи:

«Вероятность того, что расход воды на предприятии окажется нормальным (не более столько-то литров в сутки, чем принято), равна: $p = 3/4$. Найти вероятность того, что в ближайшие 6 дней расход воды будет нормальным в течение 1, 2, 3, 4, 5 и 6 дней». Для решения обозначим $P_6(k)$ вероятность того, что в течение $k=6, 5, 4, 3, 2, 1$, Одней из 6-ти расход воды будет нормальным. По формуле (3) или (4), где $p = 3/4$ (округляем до сотых):

$$P_6(6) = \frac{6!}{6!(6-6)!} \cdot \left(\frac{3}{4}\right)^6 \cdot \left(\frac{1}{4}\right)^{6-6} = 1 \cdot \frac{3^6}{4^6} = 0,18. \quad (4a)$$

$$P_6(5) = \frac{6!}{5!(6-5)!} \cdot \left(\frac{3}{4}\right)^5 \cdot \left(\frac{1}{4}\right)^{6-5} = 6 \cdot \frac{3^5}{4^6} = 0,36.$$

$$P_6(4) = \frac{6!}{4!(6-4)!} \cdot \left(\frac{3}{4}\right)^4 \cdot \left(\frac{1}{4}\right)^{6-4} = 15 \cdot \frac{3^4}{4^6} = 0,30.$$

$$P_6(3) = \frac{6!}{3!(6-3)!} \cdot \left(\frac{3}{4}\right)^3 \cdot \left(\frac{1}{4}\right)^{6-3} = 20 \cdot \frac{3^3}{4^6} = 0,13.$$

$$P_6(2) = \frac{6!}{2!(6-2)!} \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^{6-2} = 15 \cdot \frac{3^2}{4^6} = 0,03.$$

$$P_6(1) = \frac{6!}{1!(6-1)!} \cdot \left(\frac{3}{4}\right)^1 \cdot \left(\frac{1}{4}\right)^{6-1} = 6 \cdot \frac{3^1}{4^6} \approx 0,0043 \approx 0.$$

$$P_6(0) = \frac{6!}{0!(6-0)!} \cdot \left(\frac{3}{4}\right)^0 \cdot \left(\frac{1}{4}\right)^{6-0} = 1 \cdot \frac{1 \cdot 1}{4^6} \approx 0,0002 \approx 0.$$

То есть вероятность каждый день из 6-ти иметь перерасход, составляет всего $1/4^6$, или $P_6(0) \approx 0$, как и $P_6(1)$, впрочем. Остальные вероятности будут такими: $P_6(6) \approx 0,18$; $P_6(5) \approx 0,36$; $P_6(4) \approx 0,30$; $P_6(3) \approx 0,13$; $P_6(2) \approx 0,03$; $P_6(1)$ и, формально, для «нулевого» дня $P_6(0) \approx 0,00$. В сумме получается: $P = 1,00$, поскольку перечисленные выше вероятности составляют *полную группу событий*, что вполне очевидно. Здесь обозначение $P_6(0)$ означает вероятность буквально каждый день из 6-ти иметь перерасход, что составляет ($1/4^6 = 1/4096 = 0,00024$ или $0,024\%$, что близко к нулю).

Из наших вычислений видно, что наиболее вероятен перерасход воды в течение одного или двух дней из шести и что вероятность перерасхода в течение пяти или шести дней $[P_6(1) + P_6(0)]$ практически равна нулю.

Таким образом, на основе примеров 1, 2 и общей формулы Бернулли, апробированной на примере 3, формально можно было бы переходить к следующему подразделу. Однако, сама постановка задачи в примере 3, да и полученные результаты, могут получить следующее обобщение: необходимо вычислить не только вероятности наступления тех или иных событий, но и суметь как-то определять именно *наивероятнейшее* число наступления событий.

1.3. Наивероятнейшее число наступлений событий

Для большей наглядности визуализируем результаты решения примера 3 на рис. 1 в виде полигона распределения (в отличие от гистограммы, ординаты выражаются не столбиками, а воображаемыми вертикальными линиями) для тех же шести дней ($n = 6$) и $p = 3/4$ [2. – С. 55 – 62]. По оси абсцисс будем откладывать значения k , а по ординате – найденные нами значения вероятностей наступления события A , обозначенного нами как символ $P_6(k)$, рассчитанных по формулам (4а). Отметим для себя ещё раз: числа n и p полагаем для решения поставленной задачи считаем нам заданными.

Для практики иногда надо знать, какое число наступления события является *наивероятнейшим*, то есть при каком числе k вероятность $P_n(k)$ наибольшая. Согласно рис. 1 видим, что с ростом числа дней k вероятность нормального расхода воды сначала возрастает, а затем, достигнув своего наибольшего значения, начинает убывать. Для практики же иногда требуется знать, какое число наступления событий является наивероятнейшим, то есть при каком числе k вероятность $P_n(k)$ наибольшая. Наиболее простое

достижение поставленной нами цели может обеспечить уже знакомая нам формула Бернулли, которой на этот раз воспользуемся следующим образом.

С тем, чтобы впоследствии как-то формализовать процесс нахождения наиболее вероятного события, попробуем сначала, скажем, формированием такого соотношения между числителем и знаменателем, исходя из мысленных результатов оценки такого соотношения, как $P_n(k + 1) > P_n(k)$. То есть последующие вероятности как бы больше предыдущих. А для удобства реализации этого строгого неравенства попытаемся сначала вычислить величину следующего отношения, имеющего непосредственное отношение к только что упомянутому нами приведённому выше строгому неравенству:

$$\frac{P_n(k + 1)}{P_n(k)},$$

которое распишем в терминах закона Бернулли в виде (4), а затем проведём необходимые сокращения в удобном для восприятия виде (2. – С. 56 – 57):

$$\begin{aligned} \frac{P_n(k + 1)}{P_n(k)} &= \frac{n!}{(k + 1)! (n - k - 1)!} p^{k+1} (1 - p)^{n-k-1} : \\ &: \frac{n!}{k! (n - k)!} p^k (1 - p)^{n-k} = \frac{n! (p^{k+1} (1 - p)^{n-k-1}) k! (n - k)!}{(k + 1)! (n - k - 1)! n! p^k (1 - p)^{n-k}} \\ &= \frac{p^{k+1} (1 - p)^{n-k-1}}{p^k (1 - p)^{n-k}} \cdot \frac{k!}{(k + 1)!} \cdot \frac{(n - k)!}{(n - k - 1)!} = \frac{n - k}{k + 1} \cdot \frac{p}{1 - p}. \end{aligned} \quad (5)$$

Соотношение (5) сравнивается с «1»: оно может быть > 1 , < 1 или $= 1$:

Если мы хотим выяснить, при каких значениях k выполняется неравенство $P_n(k + 1) > P_n(k)$, то для этого нужно узнать, при каких значениях k имеет место неравенство, когда правая часть (5) > 1 или $(n - k) p > (k + 1) (1 - p)$, откуда после преобразований получаем

$$np - (1 - p) > k.$$

То есть, пока k , возрастая, не достигнет величины в левой части неравенства, будет соблюдаться $P_n(k + 1) > P_n(k)$, хотя $P_n(k)$ будет возрастать, и как только достигнет величины $np - (1 - p)$, далее при увеличении k будет уменьшаться.

Кстати, на рис. 1 такая картина вполне просматривается. Действительно, при наших исходных данных (величин n и p) по только что приведённой формуле (левая её часть): $6 \cdot (3/4) - (1 - 3/4) = 18/4 - 1/4 = 17/4 = 4,25 = k$.

$P_6(k)$



Рис. 1. Вероятности наступления события A (расход будет нормальным) от числа испытаний k (дни) – по формулам (4а).

То есть, как только число k перешагнёт через грань 4,25 – вероятность $P_{n=6}(k = 4,25)$ начнёт чётко убывать. Однако, $k = 4,25 > 4$. Тогда в натуральном ряде чисел k необходимо увеличить до следующей единицы (параметр k является по природе своей именно целым числом, это мы помним), то есть $k = 5$. Тогда вероятность достигнет своего максимального значения: $P_{n=6}(k = 5) = 0,36 = \max$ – см. рис. 1, а после будет убывать, что и подтверждает диаграмма на рис. 1.

С другой стороны, обозначив наивероятнейшее значение k_0 числа k , получим, следуя простой (обыденной) логике, следующее нестрогое по своему смыслу отношение:

$$P_n(k_0 + 1) \leq P_n(k_0), \quad (6)$$

откуда уже по известному нам соотношению $np - (1 - p) > k$ для прежних условий рассуждений с учётом (6) запишем

$$k_0 \geq np - (1 - p).$$

Опять же, с другой стороны, мы уже упоминали о том, что

$$P_n(k_0 - 1) < P_n(k_0),$$

для чего, по предыдущему, должно быть

$$k_0 - 1 \leq np - (1 - p) \text{ или } k_0 \leq np - (1 - p) + 1 = np + p.$$

Таким образом, наивероятнейшее значение k_0 числа k должно удовлетворять следующему двойному нестрогому неравенству:

$$np - (1 - p) \leq k_0 \leq np + p, \quad (7)$$

или в принятых нами ранее числовых данных: $4,25 \leq k_0 \leq 5,25$ - см. рис. 1, где при $k_0 = 5$ вероятность $P_{n=6}(k=5) = 0,36 = \max$.

Казалось бы, что полученное двойное нестрогое неравенство (7) предоставляет исследователю какие-то новые, дополнительные возможности, что будет далее продемонстрировано в рамках примера 3. Однако, прежде хотя бы отчасти немного порассуждаем о его содержании для лучшего его понимания в деле практического применения.

Во-первых, промежуток от $np - (1 - p)$ – левая часть выражения (7), до $(np + p)$ – правая часть того же выражения, в котором должно лежать интересующее нас число k_0 , имеет величину: $(np + p) - [np - (1 - p)] = 1$, что очевидно. Итак, величина промежутка получается строго равной единице. Следовательно, если, например, левая граница выведенной нами формулы (7), не есть целое число, то между левым и правым концом промежутка будет обязательно лежать *одно и только одно* целое число. Вот тогда только и число k_0 будет однозначно определено.

Действительно, ведь почти всегда в относительных единицах $p < 1$. Значит, величина $[np - (1 - p)]$ сможет быть целым числом лишь в очень исключительных случаях. Но и в этих воображаемых исключительных случаях двойное нестрогое неравенство (7) дают для числа k_0 именно два

значения: $[np - (1 - p)]$ и $(np + p)$, отличающиеся друг от друга ровно на единицу. Тогда эти два числа и будут *наивероятнейшими*. Тогда их вероятности должны будут быть равными между собой и превышать вероятность всех других значений числа k . Впрочем, случай почти исключительный для простоты восприятия. Что и покажем на простом вычислительном примере.

Как известно, в качестве исходных данных нам, по формуле Бернулли нам нужны изначально *всего два* параметра: число анализируемых ситуаций или возможных случаев ($n = 15$), а также вероятность наступления каждого из них $p = 1/2$. Тогда, даже не приводя соответствующего графика, при $n = 15$ и $p = 1/2$ наши прежние соотношения будут выглядеть так: левая часть нестрогого двойного неравенства (7) $np - (1 - p) = 15 (1/2) - (1 - 1/2) = 7$; $np + p = 15 (1/2) + 1/2 = 8$. Тогда наиболее вероятными значениями числа наступлений событий служат числа 7 и 8.

Заметим только, что их вероятности не только равны между собою и равны (легко подсчитать) приближённо $0,196$: это максимум для $k = 7$ и одинаковой (равной) ординаты $P_{15}(k)$ также и для $k = 8$. Что и было бы отображено на графике, если он был построен, что не только подтверждает вышеприведённые суждения, но и (мысленно) визуализирует их.

Действительно, такой бы график в заявленных координатах возрастал бы при последовательном возрастании k от 1 до 15-ти до $k = 7$ (причём $P_{15}(7) = P_{15}(8) \approx 0,196 = P_{max}$ — это при $k=7$ и $k=8$ — соответственно).

Итак, то, что выведенное нами нестрогое двойное неравенство (7), сделанное на основе формулы Бернулли (4), может оказаться продуктивным в решении ряда практических задач, проиллюстрируем на задаче, приведённой в примере 3, что является методически весьма важным для выпускников нашего вуза.

Кстати, в отличие от рис. 1, более яркую картину дают полигон распределения изменения величины $P_n(k)$ именно с ростом k , когда число n становится больше. Поскольку график, приведённый на рис. 5 [2. – С. 57] не

показался авторам достаточно убедительным, покажем проявление того же самого сравнительного подхода при $n = 10$ и $p = 1/2$. Для этого проведём отдельные расчёты по той же формуле (3) или (4) как её модифицированного преобразования в последовательности, подобной (4а), и сведём их результаты в отдельную таблицу, где k –по-прежнему число появлений событий; $P_n(k)$ – искомая вероятность k -появлений событий A с округлением до тысячных долей при n независимых испытаний:

Таблица результатов расчёта $P_n(k)$ при $n = 10$

К	$P_n(k)$	К	$P_n(k)$
1	2	1	2
0	0,001	6	0,205
1	0,010	7	0,117
2	0,044	8	0,044
3	0,117	9	0,010
4	0,205	10	0,001
5	0,246	-	-

Из приведённой здесь таблицы видно, что максимума вероятность $P_n(k)$ достигает при значении $k = 5$ (выделено шрифтом) максимального значения, а до него и после него значения являются симметричными от максимума, что и продемонстрируем на рис. 1а:

$P_n(k)$

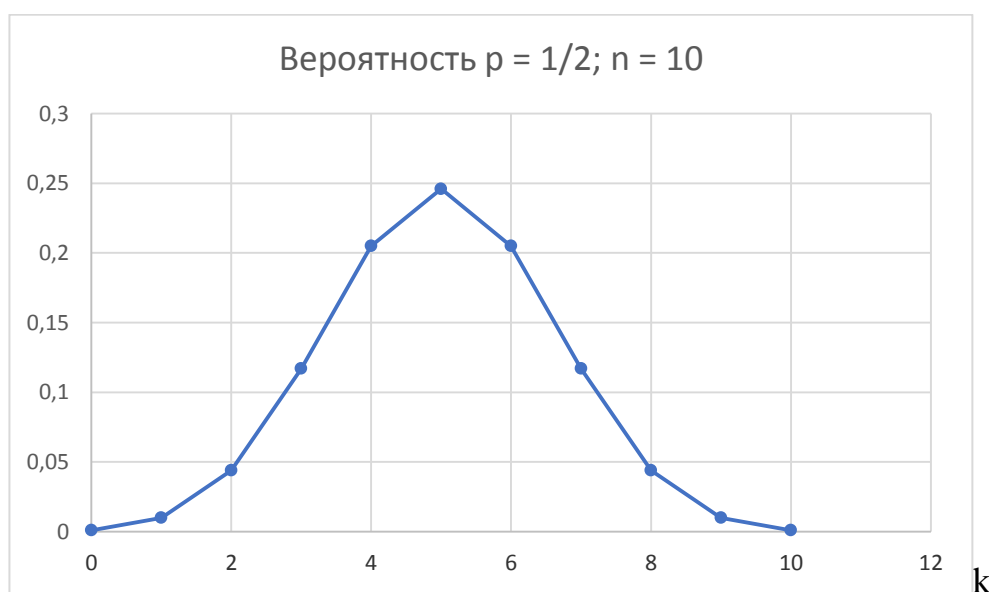


Рис. 1а. Вероятности k -появлений события A от числа испытаний n .

Из сравнения видов графиков на рис. 1 и родственного ему рис. 1а видно, что своей симметричностью график обязан в большей степени вероятности наступления k -события $p = 1/2$, хотя и при собственно росте параметра k даже при иных значениях вероятности p график в основном сохраняет приведённую здесь тенденцию, в чём легко убедиться, например, при проведении вычислений на практических занятиях, либо при проведении самостоятельных (здесь – исследовательских по своему характеру) расчётов по тем же формулам.

Пример 4[2. - С. 59]. «В результате многолетних наблюдений для некоторой местности было выявлено, что вероятность того, что в течение 1 июля пройдёт дождь, равна $4/17$. Найти *наивероятнейшее* число дождливых дней 1 июля за ближайшие 50 лет». То есть, то есть эксплицируем (формализуем), что в нашем примере $n = 50$ и $p = 4/17$ (то есть главными двумя обязательными параметрами *пир* для использования формулы Бернулли исследователи *реально* располагают). Надо только нашим будущим выпускникам постараться запомнить как постановку, так и ход решения поставленной здесь задачи, которая в их предполагаемой послевузовской деятельности может занимать немаловажное место.

Итак, решение осуществляется по выведенной нами же недавно формуле (7), то есть по знакомой уже формуле:

$$50 \frac{4}{17} - \left(1 - \frac{4}{17}\right) \leq k_0 \leq 50 \frac{4}{17} + \frac{4}{17}, \text{ то есть}$$

$$11,765 - 0,765 \leq k_0 \leq 11,765 + 0,235, \text{ или}$$

$$11 \text{ дней} \leq k_0 \leq 12 \text{ дней.}$$

В итоге получим, что *наивероятнейшее* событие под названием «число дождливых дней 1 июля в ближайшие 50 лет» успешно решена: это равновероятные между собой числа - 11 и 12 дней. Следовательно, наивероятнейшим значением числа дождливых дней за ближайшие 50 лет будут РАВНОВЕРОЯТНЫЕ между собой числа 11 и 12 дней июля. Таким

образом, задача, приведённая в примере 4, успешно решена. Заметим только, что перед началом наших (и простых по своей природе вычислений) мы располагали только статистикой «дождей на 1 июля предшествующих лет». Понятно, что в роли вероятности их прохождения именно 1 июля выступала лишь её статистическая оценка в виде «средних случаев дождя на 1 июля», наблюдаемая прежде за много лет как оценка в $4/17$.

В итоге мы получили, что за будущие 50 лет «1-го июля дождь пройдёт в 11-ти или 12-ти случаях». Значит, в ближайшие 50 лет 38 – 39 раз в тот же день следует ожидать без дождей. Другими словами: мы не знаем, какой в течение 50-ти лет день 1 июля будет дождливым, но вполне можем по нашим расчётам утверждать, что от 38-ми до 39-ти раз этот день, 1 июля то есть, будет без дождей. Напомним себе также, что каждое наше решение должно заканчиваться обязательным нашим умозаключением (логический аспект познания в триаде «понятия-суждения-умозаключение»), выраженными простыми словами. То есть и эта задача в примере 4 тоже нами успешно решена.

Ну а теперь рассмотрим не менее распространённый, но наиболее часто встречающийся случай, достойный применения формы (формулы) Якова (Якоба) Бернулли (XVII–XVIII век, Россия).

1.4. Наиболее характерный пример применения формулы Бернулли

Бывает, что число n весьма велико, тогда и np будет также большим, если не слишком мала p . Тогда в выражении (7) обе части нестрогого неравенства, что вполне очевидно, будут близки к $k_0 \approx np$. О том и следующий пример 5.

Немного предварим его содержание следующими рассуждениями такого характера. Из практики известно, что встречаются ситуации, когда число n весьма велико (массовое производство изделий, например). Тогда и произведение np будет очень большим числом, если только вероятность p не

чрезвычайно мала. А так как в выражениях левой и правой части формулы (7), между которыми заключена *наивероятнейшее* число появлений ожидаемого нами события, их вторые члены меньше единицы по определению. Тогда оба эти числа (а значит и заключённое между ними *наивероятнейшее* число появления событий k_0), будет близко к np , хотя бы справа в формуле (7). Далее – сама проблема и её решение в примере 5.

Пример 5 [2. - С. 60 - 61]. Пусть на телефонной станции (ТС) вероятность p соединения абонентов за $t < 15$ секунд равна $0,74$ ($p = 0,74$). Тогда из каждой 1000 вызовов, поступающих на ТС можно принять $1000 \cdot 0,74$ за вероятнейшее число соединений абонентов за срок, меньший 15 сек., или $k_0 = 1000 \cdot 0,74 = 740$ (соединений за $t < 15$ с). Это из каждой тысячи вызовов, поступающих на ТС.

Но данному выводу можно придать более точную форму: если *наивероятнейшее* число появления события при n испытаниях, то «доля» их появлений будет k_0/n (7) превратится в неравенство (8):

$$p - \frac{1 - p}{n} \leq \frac{k_0}{n} \leq p + \frac{p}{n}. \quad (8)$$

А теперь при той же $p = 0,74$ для отдельного испытания будем увеличивать их число n (данное обстоятельство повлечёт за собой и увеличение k_0 , что понятно). Тогда дроби $(1 - p)/n$ и p/n будут необходимо УМЕНЬШАТЬСЯ. При этом из формулы (8) следует, что $p \approx k_0/n$. Отсюда и вывод: «Таким образом, *наивероятнейшая* доля появления события при большом числе испытаний практически равна вероятности появления события при отдельном испытании» [2. – С. 61].

Краткие выводы. В качестве таких выводов по разделу можно сформулировать следующее.

Итак, мы в меру подробно ознакомились с содержанием схемы Бернулли, точнее, лишь с его формулой. Действительно, ряд прикладных приёмов (в том числе и задач контроля качества продукции и услуг) могут оказаться полезными не только в понимании процесса оценки качества в

целом, но и в овладении одной из практических методик использования исходной статистической информации для достижения поставленной конечной цели: или считать данную продукцию (услугу) качественной, или посчитать её браком.

Получив после приведённых выше преобразований формулу (7), мы тем самым приходим к весьма важному ПРАКТИЧЕСКОМУ выводу [2. – С. 64]: если мы производим серию из большого числа испытаний n , то с вероятностью, близкой к единице, можем ожидать, что число k появлений ожидаемых событий A будет весьма близко к своему наивероятнейшему значению k_0 , отличаясь от него лишь на незначительную долю общего числа n произведённых нами испытаний.

Поэтому именно это предложение и принято называть теоремой Бернулли, предложенную им в начале XVIII столетия, представляет собой один из важнейших законов ТВ, доказанной впоследствии великим русским математиком П.Л. Чебышевым (возможно известные читателям-студентам и фамилии его знаменитых учеников – А.А. Маркова и А.М. Ляпунова) [14. – С. 140].

Однако, нас в изучаемом контексте всё же наибольший интерес представляет не столько отдельная теорема (формула) Бернулли, весьма полезная для отдельных практических расчётов (в чём мы с вами уже имели возможность убедиться, и не раз), но именно ЗАКОНЫ РАСПРЕДЕЛЕНИЯ (ЗР) исследуемых СЛУЧАЙНЫХ ВЕЛИЧИН (СВ), что далее рассмотрим отдельно. Но тогда следует несколько слов сказать как о случайных величинах, так и законах их распределения.

2. О СЛУЧАЙНЫХ ВЕЛИЧИНАХ И ИХ ЗАКОНАХ РАСПРЕДЕЛЕНИЯ

Как уже нами отмечалось, закон распределения случайных величин (ЗР СВ) исследуемого статистического процесса далеко не исчерпывается приведённым выше и разобранный нами подробно формулой Бернулли, так

как не менее важным является уточнение самих исходных понятий «случайная величина» и «закон её распределения» в целом [27], [28].

Случайная величина. Это величина меняется под влиянием случайных воздействий [2. – С. 72 - 74], численное значение которой не может быть нами раз навсегда определено. Как и число мальчиков из тысячи случаев рождений, как и длина волокон хлопка, зависящая также от многих-многих причин (сорта, местности произрастания, от расположения «коробочки» хлопка на каждом кусте и пр.). И список таких примеров может быть сколь угодно расширен (число метеоритов, падающих на землю за некоторый промежуток времени, неизбежное рассеяние снарядов по дальности и направлению от стреляющей пушки по одной и той же цели, и вес зёрен пшеницы, взятые даже из одного колоса и масса других примеров). Такого рода величины мы и называем в ТВ *случайными величинами*. Отсюда и несомненная важность при её изучении.

Поэтому знание данной СВ вовсе не означает знание её, объективно существующего конкретного численного значения. Пример: наблюдаемый нами конденсатор до пробоя под штатным электрическим напряжением проработал, скажем, 5324 часа. В данном примере время работы конденсатора уже не является СВ, поскольку оно нам уже доподлинно известно. Другое дело: мы собрали множество наблюдений таких численных и непохожих друг на друга значений (хотя и не исключаем отдельных случаев совпадений при оценке «времени жизни» конденсаторов данного типа). Отметим только, что даже имеющийся ряд численных значений ещё не даёт нам таких сведений об исследуемой нами СВ, которые могли бы служить нам материалом для практически необходимых оценок.

Тогда поставим интересующий нас вопрос по-другому: если мы хотим оценить, например, время жизненного цикла исследуемого конденсатора того или иного типа, то зададимся такой проблемой: а как часто наблюдается то или иное время жизни конденсатора. Иными словами, мы хотели бы знать

вероятности этих, фиксируемых нами, моментов времени. Величины, подобные здесь рассмотренной нами, и называются *случайными*.

Вполне закономерный и логически вытекающий вопрос из предыдущего: а как распределены вероятности отдельных измерений (в широком смысле – результатов наблюдений исследуемой нами СВ)?

Понятие о законе распределения СВ. Итак, из своих наблюдений за отдельными случаями проявления СВ, как бы подразумевается, что их мы оформили (эксплицировали, формализовали) в виде некоторой СС с именем X в виде её элементов $\{x_i\}$, $i = 1, n$, где i –текущая переменная (счётчик наших наблюдений), а выражение $i = 1, n$ мы записали лишь для того, чтобы показать возможному читателю наших с вами экзерциций, что в сформированной нами СС содержится ровно n элементов, которые начинаем считать с самого начала, то есть с «единицы». Тогда сама СС примет вид: $X = \{x_i\}$, $i = 1, n$.

Ясно, что предыдущие наши рассуждения мы сводим теперь не только к понятию СВ (то есть к её элементам x_i числом n), но и к *частотам* их появления как отдельных, случайных по своей природе величин. Иными словами, мы хотим знать, как часто следует ожидать появления выявленных нами ранее отдельных элементов x_i числом n этой самой, сформированной нами же самими, СС с именем X . А для её характеристики [там же. – С. 76] необходимо и достаточно знать:

- 1) перечень всех её возможных значений x_i ;
- 2) вероятность p_i появления каждого из этих значений x_i .

То есть задать исследуемую СВ легче всего в виде следующей таблицы, состоящей всего из двух строк, верхняя из которых будет содержать в каком-нибудь порядке возможные значения (элементы СВ) x_i , а нижняя – их вероятности p_i . Тогда исходные данные типа «элемент случайной величины – вероятность её появления» образует некоторую пару:

Таблица вероятностей

	Значения пар « X_i » - « p_i »			
Значения элементов X_i	X_1	X_2	...	X_n
Значения вероятностей p_i	p_1	p_2	...	p_n

Задав такую таблицу, мы тем самым задаём закон распределения исследуемой СВ x_i . Однако созданием лишь такой таблицы исследование на данном этапе далеко не заканчивается, поскольку к настоящему времени известно немалое число именно ЗАКОНОВ РАСПРЕДЕЛЕНИЯ исследуемой случайной величины (элементов СВ с данным нами именем X). Элементы СВ могут быть распределены по какому угодно закону распределения: нормальному ЗР К.Ф. Гаусса (НЗР), который известен ещё со школьного курса (та самая знаменитая «шапка Гаусса» в координатах плотности вероятности $f(x)$ – по ординате графика и исследуемой СВ с именем X по оси абсцисс; подробнее см. её график на рис. 5), Релея (та самая «шапка», но как бы скошенная вправо)[16], Вейбулла [17], равномерный ЗР [18] и ряд других.

Для справки. Попутно заметим только, что перечисленные выше ЗР широко используются в самых разных аспектах практики. Так, ЗР СВ Релея обычно применяется при анализе надёжности автоматизированных и других подобных технических систем. ЗР СВ Вейбулла – при исследовании плотности вероятности безотказной работы исследуемых объектов. Тогда как равномерный закон используется лишь на отдельных отрезках ЗР СВ с постоянными плотностями распределения. О других ЗР СВ, таких, как модальный, антимодальный ЗР СВ и ряде других можно подробнее узнать из соответствующих источников, их описывающих.

Однако, несмотря на то, что, что перечисленные ЗР ориентированы на изучение поведения в основном *непрерывных* СВ (более подробно о ЗР Гаусса – см. ниже), нас, судя по приведённым выше примерам, всё же больше

интересуют случайные события, носящие *дискретный* характер, столь типичных для современной т.н. «турбулентной экономике» [19,20]. Причём из всего разнообразия ЗР СВ подробнее рассмотрим такие ЗР СВ, которые носят именно дискретный характер, что, кстати, широко используются в р-картах контроля качества готовых изделий на предприятиях, что и является для нас особенно актуальным.

Далее, прежде всего, речь пойдёт подробнее о двух таких дискретных ЗР СВ, как биномиальный закон распределения и закон распределения Пуассона. Но и нормальный закон распределения Гаусса также не должен быть нами оставлен в стороне для более подробного рассмотрения, хотя изначально он создан был для исследования именно непрерывных СВ. Но то обстоятельство, что большинство элементов СВ подвержены действию множеству, в том числе и неконтролируемыми никем причин, НЗР легко адаптируется и к исследованию *дискретных* СВ, в чём легко убедимся с вами в последующем изложении.

Затем рассмотрим подробнее также широко используемый в контроле качества биномиальный закон распределения.

3. БИНОМИАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ

В дополнение к предыдущему материалу (а до сих пор мы с вами лишь конкретно использовали формулу Бернулли для решения отдельных практических и, как преобразование её - путь к отысканию наивероятнейшего значения того или иного ожидаемого нами события А. Но, как заметили в предыдущей главе, нас может интересовать не только отдельные решения, но и исследование общего вероятностного поведения элементов СВ с точки зрения распределения между ними тех или иных значений вероятностей согласно таблице, объединяющей в себе как наступление событий, так и вероятностей из наступления. И которую мы

назвали отражением именно РАСПРЕДЕЛЕНИЯ отдельных значений (элементов) СВ с присущими им вероятностями.

Таким образом, поставленное нами во взаимно однозначное соответствие собственно появления отдельного события с вероятностью его наступления мы и условились называть ЗАКОНОМ РАСПРЕДЕЛЕНИЯ исследуемой СВ или, сокращённо, ЗР СВ.

В качестве такого ЗР элементов СВ далее рассмотрим с позиции так называемого закона о дискретном распределении элементов СВ биномиального закона распределения (БЗР) [3. – С. 194 – 212], который весьма часто применяемого при контроле качества продукции (ресурсов, услуг).

Ранее нами рассмотрен в качестве полезного примера формула, предложенная Яковом Бернулли. Однако в более широком смысле мы ввели понятие *распределения* элементов СВ как функции, которая определяет долю случаев реализации каждого из всех возможных результатов при воображаемых ими т.н. МНОГОКРАТНЫХ измерениях. Сюда же можно отнести и понятие *предельного распределения*, которое определяет вероятность того, что единственное измерение приведёт к одному из возможных значений исследуемого явления.

Из множества возможных распределений распределение Гаусса (нормальный закон распределения – НЗР) отличается тем, что описывает распределение результатов любых измерений в случае, если эти измерения подвержены действию множества небольших и случайных ошибок (что бывает, понятно, в общем не всегда). Поэтому НЗР Гаусса является, например, для физиков, наиболее важным из всех предельных распределений. Но при этом вполне могут представить практический и теоретический интерес и ряд других распределений, которые в причинной связи как с отдельными факторами-причинами, так и их совокупностями, могут представлять собой закон распределения несколько иного набора свойств, которые присущи именно нормальному распределению их.

То есть прежде всего речь должна идти, прежде всего, о таких видах дискретных законах распределения случайных величин, как биномиальный ЗР, и ЗР Пуассона, наиболее часто применяющиеся в задачах контроля качества.

Поэтому мы с вами сначала рассмотрим биномиальное распределение или, сокращённо, БЗР, которое хоть и имеет известное значение как для экспериментальной физики, так и для контроля качества изготовленной на предприятии продукции в целом, однако из-за своей ОТНОСИТЕЛЬНОЙ ПРОСТОТЫ именно этот закон распределения может послужить «отличным введением к пониманию многих характеристик распределений» [3. – С. 195]. Поэтому и начнём его изучение, а также других ЗР с простых модельных примеров.

3.1. Вероятность при бросании игральных костей

Биномиальный ЗР лучше показать на таком хрестоматийном примере, как бросание трёх игральных костей [3. – С. 195 - 196] (кости шестигранные, грани пронумерованы от 1 до 6-ти). Считаем ВОЗМОЖНЫМИ результаты – это 0, 1, 2 или 3 единички – всего 4 варианта. Ну, и как всегда: если будет бросать кости ОГРОМНОЕ число раз, то «единички» выпадут k раз, где $k = 0, 1, 2$ или 3 раза, как обозначено в ряде источников. Далее рассмотрим содержание возможных вариантов наступления (или не наступления) интересующих нас событий.

Сначала рассмотрим вариант наступления всех четырёх событий. Итак, вероятность выпадения одной «единички» (как и другого числа от 1 до 6-ти), равна ровно строго $(1/6)$. А теперь будем бросать, скажем, ТРИ кости и найдём вероятность выпадения всех трёх «единичек» ($k = 3$). Тогда вероятность выпадения трёх «единичек» при бросании *сразу трёх* костей может быть определена (по правилу в ТВ – «выпала одна «единичка», «И» ещё одна, «И» ещё одна» – данное естественное логическое умножение

заменяется на буквально СООТВЕТСТВУЮЩЕЕ ему алгебраическое, или как произведение перечисленных нами ОТДЕЛЬНЫХ событий следующего вида:

$$P(\text{три "1" из 3 бросаний}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \left(\frac{1}{6}\right)^3 = 0,46\% \approx 0,5\%. \quad (9)$$

А вот рассчитать вероятность выпадения уже ДВУХ «единичек» ($k = 2$) явно несколько ТРУДНЕЕ, поскольку мы можем получить этот результат уже РАЗЛИЧНЫМИ способами. Например, 1-я и 2-я кости могли бы выпасть на «единички», а третья – нет ($A, A, \text{не-}A$). Или первая и третья могли бы выпасть «единичками», а вторая – нет ($A, \text{не-}A, A$) и т.д. Поэтому, вследствие такой сложности, и не сразу видной, будем вычислять искомую вероятность, в отличие от формулы (9), в два очевидных приёма:

1. Сначала вычислим вероятность выпадения двух «единичек» при какой-то заданной определённой комбинации, например, « $A, A, \text{не-}A$ » (вероятность, что третий раз не выпадет «единичка» равна $5/6$, что весьма понятно даже на интуитивном уровне):

$$P(A, A, \text{не} - A) = \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^1 = 0,0231 \approx 2,3\%. \quad (10)$$

Но и вероятность P для двух других возможных комбинаций будет *точно такой же*, как это показано в формуле (10). Поэтому объективно существуют только ТРИ различные комбинации, когда мы могли бы получить из трёх бросаний кости именно ДВЕ «единички»: ($A, A, \text{не-}A$), ($A, \text{не-}A, A$) и ($\text{не-}A, A, A$).

2. Таким образом, ПОЛНАЯ ВЕРОЯТНОСТЬ получения двух «единичек» в какой угодно комбинации можно вычислить по формуле (10) для одной кости (а у нас их три):

$$P(\text{две "1" из 3 бросаний}) = 3 \cdot \left[\left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right) \right] \approx 6,9\%. \quad (11)$$

Для справки: подобные же расчёты по формуле (11) дают значения вероятности выпадения ОДНОЙ «единички» в трёх бросаниях (34,7%) и НИ ОДНОЙ «единички» (57,9%), что далее и проверим.

Так, выпадение ОДНОЙ «единички» в трёх бросаниях по аналогии с расчётом трёх бросаний – это равноправные варианты $(A, не-A, не-A)$, $(не-A, A, не-A)$ и $(не-A, не-A, A)$, например, вероятность P для первого варианта:

$$P(A, не - A, не - A) = \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 = 11,6\%.$$

Это в одном бросании, но подобных ситуаций получается целых три: $(A, не-A, не-A)$, $(не-A, A, не-A)$ и $(не-A, не-A, A)$:

$$P(\text{одна "1" из 3 бросаний}) = 3 \cdot \left[\left(\frac{1}{6}\right) \cdot \left(\frac{5}{6}\right)^2 \right] \approx 34,7\%. \quad (12)$$

Ну и так же просто, как и для формулы (9), P (3 «единички» из 3 бросаний), составит формула (12), а также для наступления события с вероятностью (для $k=0$) P (0 «единичек» из 3 бросаний):

$$P(\text{нет "1" из 3 бросаний}) = \left(\frac{5}{6}\right)^3 \approx 57,9\%. \quad (13)$$

Теперь по результатам (13) при $k=0$, (12) при $k=1$, (11) при $k=2$ и (9) при $k=3$ можно представить на графике как примере биномиального распределения (см. рис. 2). Заметим, кстати, что все наши предыдущие рассуждения строились на так называемом «умозрительном уровне» или на уровне простого здравого смысла (т.е. так называемого обыденного познания). Но, к сожалению, подобное возможно лишь в достаточно простых, хотя и по природе своей весьма нетривиальных, задачах.

Итак, после рассмотрения построенного нами графика на рис. 2 для данного частного примера уже можно сделать некоторые предварительные обобщения. Прежде всего, глядя на выражения (9) – (13), можно весьма просто прийти к выводу, что в основе действий наших лежало простое правило: перебрать все возможные комбинации, особенно для менее

наглядных, по отношению к (9) и (13), выражения (10), (11) и (12), так как они связаны с *перебором* всех возможных случаев.

Далее - немного просто порассуждаем. С одной стороны, в алгебре отношений для матрицы типа «объект-объект» - те же квадратные корреляционные матрицы, например, - для симметричных отношений, лежащих выше главной диагонали. Если они равны единице, то отношения называются *рефлексивными*: каждый объект подобен сам себе. Так, все технические системы являются рефлексивными, а вот экономические, социологические и пр. – далеко не всегда.

Пусть, например, матрица типа «объект-объект» будет у нас размером $(n \times n)$, когда размер (порядок) матрицы в $n = 3$ элемента. То есть бросаем игральную кость 3 раза, и «ждём» появления «единичек» (или какого другого числа – это уж точно неважно): выражения (9) – (13) годятся для буквально любого выпавшего числа – от 1 до 6-ти (по числу граней игральной кости).

Эта функция представляет собой *биномиальный закон распределения* исследуемой СВ, который символически записывается следующим образом: $b_{n,p}(k) \text{ с } n = 3 \text{ и } p = 1/6$ по формулам (10), (11), (12) и (13):

Для более продуктивного понимания последующего учебного материала, необходимо обратиться к ранее пройденному ещё в средней школе материалу из теории множеств, а именно – к так называемой «комбинаторике», точнее, к её основным для нас формулам, определяющими такие важные для понимания понятия её аппарата, как «размещения», «перестановки» и, особенно важное для нас, понятие «сочетание». Попутно отметим, что «размещениями из n элементов по m называются такие выборки, которые, имея по m элементов, выбранных из числа данных n элементов, отличаются одна от другой либо составом элементов, либо порядком их расположения» [7. – С. 26]. И там же: если $m = n$, то число таких размещений, которые отличаются только порядком расположения элементов, но не самими элементами. Такие размещения называются перестановками; их число равно $n!$

Вероятность, %

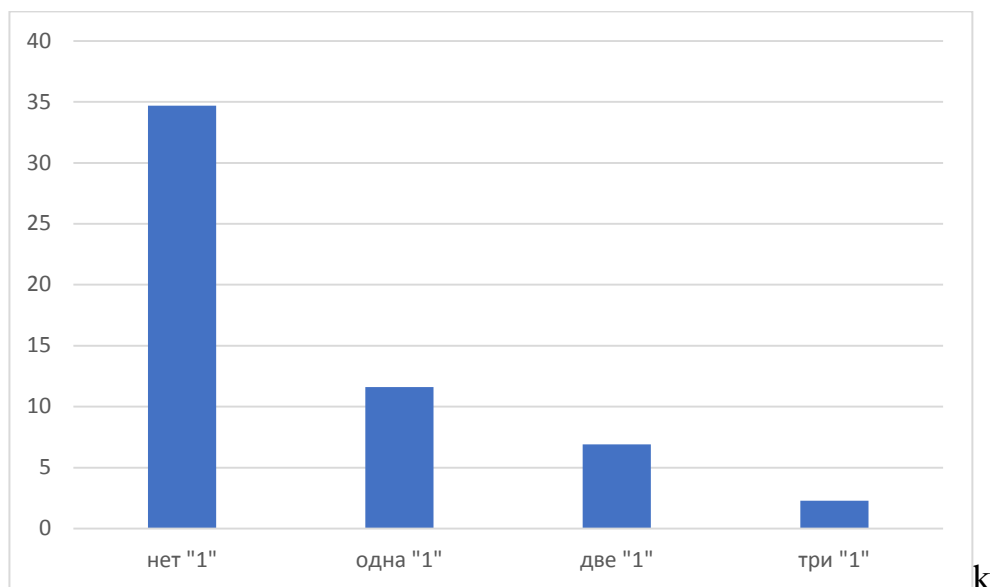


Рис. 2. Вероятность получения k «единичек» при бросании 3-х игральных костей.

Относительно же возможности применения формул биномиального закона распределения по части того или иного подхода к формированию нулевой гипотезы о его возможном действии, в литературе определённно не сказано, что тем не менее вызывает следующее понимание. Действительно, если известна средняя вероятность наступления интересующего исследователя события (здесь – случаев брака) с известным числом общих испытаний, то применение формулы Бернулли с предшествующим числом сочетаний из комбинаторики будет вполне уместной.

3.2. Основные формулы комбинаторики

Итак, при вычислении выражений (10) – (12) для $k = 1, 2$ (число появления «единичек» при бросании 3-х костей) мы видим, что число парных сравнений z для симметричных рефлексивных матриц [26] может быть (всего наддиагональных элементов рефлексивных симметричных матриц типа «объект-объект»):

$$z = \frac{n(n-1)}{2}.$$

Для $n = 3$ (как у нас) может быть $z_3 = 3 \cdot (3 - 1) / 2 = 3$ (комбинации). Ну да: первая со второй (или $1 - 2$); $1 - 3$; $2 - 3$. А если бы $n = 4$? – Было бы уже $z_4 = 4 \cdot (4 - 1) / 2 = 6$ (комбинаций): $1 - 2$; $1 - 3$; $1 - 4$; $2 - 3$; $2 - 4$; $3 - 4$. Ведь так? И наши предыдущие вычисления при шести бросании кости с ожиданием тех же «единичек» существенно возросли бы (да и запутаться в комбинациях шансы возрастают). Поэтому и существует такой раздел алгебры или теории множеств, как **комбинаторика** [7. – С. 24 – 27; 13]. Это область математики, в которой изучаются вопросы о том, сколько различных комбинаций, *подчинённых определённым условиям*, можно составить из элементов, принадлежащих анализируемому нами множеству. Поэтому методы комбинаторики дают возможность осуществлять подсчёт числа возможных исходов и числа благоприятных исходов в различных конкретных случаях. При этом при изучении ТВ важно усвоить такие основные понятия комбинаторики, применительно к исследуемой статистической совокупности, как *размещения, перестановки и сочетания*.

Пример размещения. Из элементов a , b и c можно организовать такие выборки-размещения по 2 элемента: ab , ba , ca , ac , bc и cb . Всего их получается шесть, и одна от другой отличается либо элементами, либо их порядком. Число размещений из n элементов по m будем обозначать A_n^k . Пусть имеется n элементов. Первый можно выбрать n способами. Вторым элементом можно выбирать из оставшихся $(n - 1)$ способами. Тогда пары двух элементов, что очевидно, можно образовать $n(n - 1)$ способами и т.д.

Итак, для примера применения понятия «размещения» решим следующую задачу.

Задача 1. Пусть в высшей лиге по футболу 18 команд, которые борются за золотые, серебряные и бронзовые медали [7. – С. 26]. Сколькими способами медали могут быть распределены между командами? – Иначе говоря, нам нужно найти число РАЗМЕЩЕНИЙ A_{18}^3 . Для общего случая

пусть имеем n элементов, а число способов размещения по k элементам будет:

$$A_n^k = n(n-1)(n-2) \dots [n-(k-1)]. \quad (14)$$

Тогда наше решение по формуле (14) будет таким (при $k=3$):

$$A_{18}^3 = 18(18-1)(18-2) = 18 \cdot 17 \cdot 16 = 4896,$$

или на умозраительном уровне:

$$A_{18}^3 = 18(18-1)(18-2)(18-3) = 18 \cdot 17 \cdot 16 = 4896.$$

Так что, наше утверждение, что «золото будет у киевских динамовцев, серебро – у араратцев, а бронзовые – у московских спартаковцев» – будет справедливым лишь одной комбинации из 4896-ти возможных (!).

При этом легко убедиться в том, что если $k=n$, то A_n^n – это число таких РАЗМЕЩЕНИЙ, которые отличаются только порядком расположения элементов, но не самими элементами. Такие размещения называются ПЕРЕСТАНОВКАМИ. Тогда их число по формуле (14) можно вычислить как

$$A_n^n = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1 = n! = P_n. \quad (15)$$

Для справки: число n может принимать не только натуральные значения, но может быть и равным нулю. Однако, пустое множество (выборка) является подмножеством любого множества, и естественно считать, что оно может быть упорядочено только одним способом. Также принято считать, что $0! = 1$.

На практике же не всегда важен именно ПОРЯДОК РАСПОЛОЖЕНИИ в выборках, что и продемонстрируем – в задаче 2.

Задача 2. В полуфинал первенства РФ по шахматам участвуют 20 шахматистов, а в финал из них попадут только трое [7. – С. 27]. Поэтому участнику в принципе безразлично, какие из первых трёх мест занять (вполне возможен случай, когда занявший в полуфинале 3-е место в финале был первым). Если требуется установить, сколькими способами может образоваться финальная тройка, то надо посчитать только те выборки из 20 элементов по 3, которые одна от другой отличается ХОТЯ БЫ ОДНИМ

элементом. Другими словами, требуется найти число СОЧЕТАНИЙ из $n=20$ элементов по $k=3$:

$$C_n^k = \frac{A_n^k}{P_k} = \frac{n!}{k!(n-k)!}. \quad (16)$$

Применяя формулу (16) для решения задачи 2 о шахматистах, получим искомое нами число возможных финальных троек:

$$C_{20}^3 = \frac{n!}{k!(n-k)!} = \frac{20 \cdot 19 \cdot 18 \cdot (17!)}{3!(20-3)!} = \frac{20 \cdot 19 \cdot 18}{1 \cdot 2 \cdot 3} = 1140.$$

В результате применения формулы (16) получим 1140 возможных «троек».

Да, на первый взгляд кажется, что, «неужели столько много?». Однако, это действительно так. На данном фрагменте напоминания о комбинаторике пока завершим. Запомним только содержание таких понятий, как «размещение», «сочетание», «перестановки». Заметим только, что из всех перечисленных здесь трёх понятий комбинаторики наибольший последующий интерес для нас будет иметь именно понятие «сочетание». Далее рассмотрим дискретный по своей природе, как отдельную предыдущую формулу для СВ Бернулли, биномиальный закон распределения как один из основных при организации контроля качества производимых изделий.

3.3. Определение биномиального распределения

Будем называть ИСХОД, в котором мы заинтересованы, как *успех* или *выигрыш*. Обозначим p как вероятность успеха в любом одном испытании и через $q = (1 - p)$ - вероятность «проигрыша». Тогда вероятность при бросании кости одного очка $p = 1/6$. Ну и вероятность выпадения «орла» при бросании монеты $p = 1/2$. Тогда, обобщая, вполне можем записать следующее выражение (17):

$$P(k \text{ успехов в } n \text{ испытаниях}) = b_{n,p}(k) = C_n^k p^k q^{n-k}, \quad (17)$$

где, как обычно, p означает вероятность успеха в одном испытании, а вероятность «неуспеха» равна $q = 1 - p$. Ещё и биномиальный коэффициент C_n^k (число возможных сочетаний) вычисляется, как и в случае предыдущего распределения Бернулли, повторяя выражения (15) и (16):

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (18)$$

где C_n^k - число различных комбинаций, в которых получается k успехов в n испытаниях. Формула, что можно заметить без труда, весьма и весьма похожа на формулу (3а)? – Действительно похожа, как формула расчёта числа возможных сочетаний, столь необходимой как для теории, так и для практики применения закона вычисления отдельных вероятностей по известной нам уже формуле Бернулли. Да и вычисления, обозначенные в формулах (9) – (13), весьма похожи на формулы системы соответствующих уравнений (4а). Ведь так?

В заключение данного подраздела апробируем полученные формулы (17) и (18), например, на уже знакомой нам ситуации (12), где вероятность выпадения 1-ой ожидаемой нами «единички» ($k = 1$) из 3-х бросаний, то есть $k = 1$ при $n = 3$ для обобщённых формул (17) и (18). Тогда, согласно выражению (18):

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{3!}{1!(3-1)!} = \frac{1 \cdot 2 \cdot 3}{1 \cdot 1 \cdot 2} = 3.$$

– тоже «3», как и в формуле (12).

Или по выражению (17):

$$P(1 \text{ успех в } 3 - \text{х опытах}) = 3 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^{3-1} = 0,374 = 34,7\%.$$

вероятность получена та же самая, что и в выражении (12).

Чтобы лучше оценить приведённое выше, обратимся к более продуктивному для понимания уже сказанного нами, то есть к более наглядному примеру.

Пример 5. Мы бросаем монету 4 раза («орёл» или «решка»: на ребро монета встанет вряд ли, что вполне понятно). Подсчитываем число «орлов» k . Какова вероятность получения различных возможных значений $k=0$ (ни одного «орла»), 1, 2, 3 и 4 – максимальное число возможного выпадения «орлов».

То есть отметим для себя: $n = 4$ (число наших опытов). А вероятность выпадения монеты «орлом» или «решкой» - вполне понятно, - равна $1/2$. Тогда искомая вероятность – просто биномиальное распределение $b_{n,p}(k)$ с $n=4$ и $p = q = 1/2$. Иначе говоря, по формуле (17):

$$P(k \text{ "орлов" в } n = 4 \text{ опытах}) = b_{n,p}(k) = C_4^k p^k q^{n-k} = C_4^k \left(\frac{1}{2}\right)^4 = \\ = C_4^k \cdot 0,0625$$

Может возникнуть вопрос: а почему сомножитель $p^k q^{n-k} = (1/2)^4$? – вполне очевидно потому, что вероятность выпадения «орла» или «решки» совершенно определённо равна $(1/2)$. Но иногда в подобных случаях надо всё увиденное даже в формулах постараться соотносить полученные результаты и ход их получения с позиции т.н. «здравого смысла». Так, заметим, и в данном случае.

Далее формулу вычислить легко по частям. Итак, последний член равен $0,0625$. Остаётся вычислить биномиальные коэффициенты. Первый - для случаев невыпадения «орла» ($k=0$), тогда коэффициент по (18) будет:

$$C_4^0 = \frac{4!}{0!(4-0)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot (4)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot (1 \cdot 2 \cdot 3 \cdot 4)} = 1. \\ C_4^1 = \frac{4!}{1!(4-1)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot (3)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot (1 \cdot 2 \cdot 3)} = 4. \\ C_4^2 = \frac{4!}{2!(4-2)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot (2)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot (1 \cdot 2)} = 6. \\ C_4^3 = \frac{4!}{3!(4-3)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot (1)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot (1)!} = 4. \\ C_4^4 = \frac{4!}{4!(4-4)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4 \cdot (0)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 1} = 1.$$

Теперь нам остаётся лишь умножить все полученные результаты и (на всякий случай сложить – вдруг единица не получится, то есть имели дело не с полной группой событий, помните?). Проверим: $(1 + 4 + 6 + 4 + 1) \cdot 0,0625 = 16 \cdot 0,0625 = 1,0000$ – всё верно. Можно строить график (рис. 3) в процентах или относительных единицах (см. рис. 3):

$$6,25 + 25,00 + 37,50 + 6,25 = 100,00 (\%)$$

$$1 \cdot 0,0625 + 4 \cdot 0,0625 + 6 \cdot 0,0625 + 4 \cdot 0,0625 + 1 \cdot 0,0625 = 1,0000$$

$b_{n,p}(k), \%$

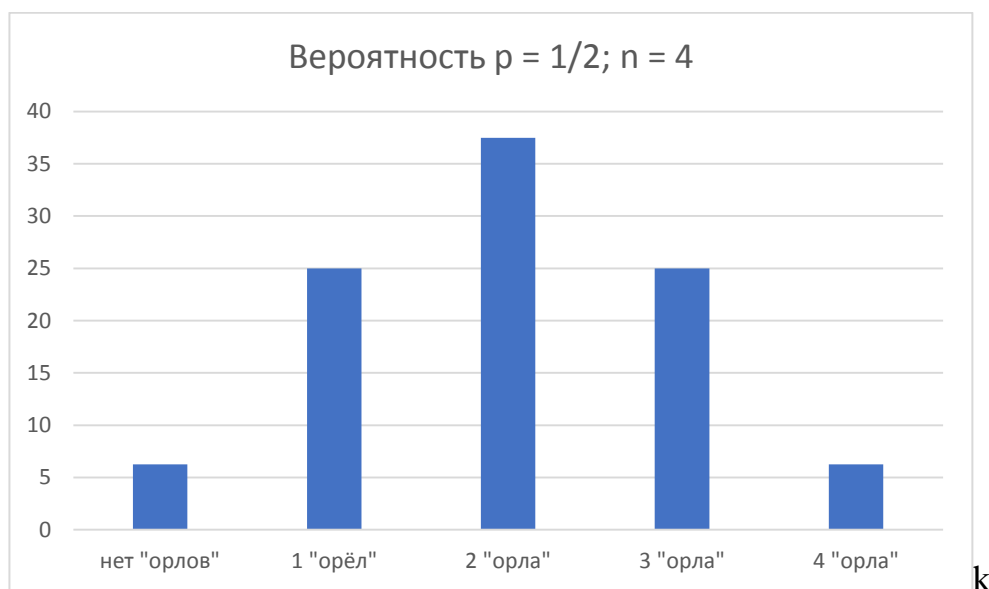


Рис. 3. Биномиальное распределение $b_{n,p}(k)$ с $p = 1/2$, $n = 4$, для вероятности выпадения k «орлов» при подбрасывании 4-х монет.

Из рис. 3: наиболее вероятное выпадение «орла» $k = 2$ (как и следовало ожидать). Оттуда же: вероятность выпадения трёх «орлов» такая же, как и при выпадении лишь одного «орла», а вероятность выпадения аж четырёх «орлов» (!) равна такой же вероятности выпадения ни одного «орла». Но, как мы видим, на рис. 3, такая симметрия существует только при $p = 1/2$. Это так. Ну а теперь вспомним что нам не столько важно само это биномиальное

распределение, сколько свойства, вытекающие из него. И какое, наконец, они имеют отношение к контролю качества продукции (и услуг)?

Отметим также и следующее немаловажное обстоятельство.

Итак, мы ознакомились лишь с общими положениями и особенностями биномиального закона распределения. Однако, для нас представит больший интерес, безусловно, как данный закон распределения случайных величин станет полезным нам при организации контроля качества продукции (и услуг), ведь так? – Попытаемся дать ответ и на этот естественный вопрос. Только для начала, и весьма кратко, остановимся на СВОЙСТВАХ биномиального распределения.

3.4. Свойства биномиального распределения

Начнём, пожалуй, с главного, а именно: биномиальное распределение $b_{n,p}(k)$ определяет вероятность реализации k «успехов» в n испытаниях в случае, когда p есть вероятность успеха в единственном испытании. Пусть наш эксперимент состоит из n испытаний, повторённый много раз, то мог возникнуть вопрос: а каково же СРЕДНЕЕ ЧИСЛО УСПЕХОВ k ? Опуская промежуточные преобразования, отметим, что среднее число успехов [3. – С. 199] можно посчитать по формуле (19):

$$k^{cp} = \sum_{k=0}^n k b_{n,p}(k) \quad (19)$$

и легко вычисляется тоже как среднее число выигрышей при биномиальном распределении

$$k^{cp} = np. \quad (20)$$

Поскольку переход от выражения (19) к выражению (20) не воспринимается как очевидный, тот же автор Тейлор в работе [3] на стр. 211 отсылает интересующихся к задаче с номером 10.8, которая предлагает среднее число выигрышей в формуле (19) привести к формуле (20). Причём отмечается, что «существует много способов доказать это», и один из

лучших состоит в записи биномиального разложения, которое справедливо для любых двух чисел p и q и любого целого $n > 0$:

$$(p + q)^n = p^n + np^{n-1}q + \dots + q^n = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k},$$

где биномиальный коэффициент, обозначающий число сочетаний из n элементов по k согласно [7. – С. 27] можно обозначить как более привычную для нас запись

$$\binom{n}{k} = C_n^k = \frac{n!}{k!(n-k)!}.$$

Поскольку равенство, приведённое выше, верно для любых p и q , то можно продифференцировать это равенство по p . Если затем применить известное равенство $p + q = 1$ и умножить его почленно на p , то в итоге неизбежно получим результат в виде выражения (20).

То есть, когда $p = 1/2$ (как в случае бросания монет), то среднее число успехов k^{cp} , согласно (20), равно просто $n/2$, что вполне соответствует т.н. «здравому смыслу».

Прокомментируем это выражение (20) (и «подозрительно простое» к тому же). Согласно ему, если мы с вами повторим серию n испытаний много-много раз, то среднее число УСПЕХОВ будет равно вероятности успеха в одном испытании p , умноженное на это самое n .

Тогда стандартное отклонение (с.к.о.) для нашего числа успехов:

$$\sigma_n = \sqrt{np(1-p)}. \quad (21)$$

Однако, заметим, что распределение при $p \neq 1/2$ биномиальное распределение **НЕСИММЕТРИЧНО**. В этом плане любопытно было бы сравнить данное биномиальное распределение $b_{n,p}(k)$ с распределением Гаусса $f_{x,\sigma}(x)$ для непрерывной СВ. Ну, во-первых, биномиальное распределение описывает **ДИСКРЕТНЫЕ** значения $k = 0, 1, 2, \dots, n$, тогда как Гаусс описывает **НЕПРЕРЫВНЫЕ** значения измерений величины $x^{cp} = p^{cp}$. То есть биномиальное распределение *симметрично* лишь относительно своего максимума только при $p = 1/2$.

Заметим также, что далее следующие выражения (22) и (23) для вычисления средней величины и с.к.о. от неё вместо ранее приведённых, соответствующих им выражений (20) и (21), в реальной практике применяются достаточно широко при оценке верхней и нижней границы в *p*-контрольных картах Шухарта, широко применяемых в контроле качества продукции, что будет показано далее, см. [4]. Получается что-то вроде подбрасывания монеты, но с вероятностью совсем иного вида и рода (и уж явно с вероятностью $p \neq 1/2$).

В это связи рассмотрим более подробно тот же, уже приведённый нами пример, который не раз встретится в экзаменационных билетах, взятый из [4].

Пример 6. Ещё раз рассмотрим данный пример для более подробной иллюстрации применения биномиального закона распределения. Для контроля качества выпускаемой продукции организовали отбор готовых изделий в 10 партиях по 100 штук изделий. Затем подсчитывалось число несоответствующих единиц продукции. Такие исходные данные для выбора метода и процесса решения приведены в табл. 1:

Таблица 1

Результаты контроля качества в 10 партиях по 100 изделий в каждой

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, шт.	100	100	100	100	100	100	100	100	100	100
Число брак. изд., шт.	1	2	0	3	14	4	1	2	15	3

Здесь в каждой партии *i* число контролируемых элементов n_i , $i = 1, N = 10$, то есть ровно по 100 единиц, и поэтому $n_i = n = const = 100$.

Как нам уже известно, выбор вида того или иного статистического распределения осуществляется на основе свойств данного распределения, характер которого и должен установить сам исследователь.

Вот мы пока подробнее изучили два закона распределения случайных величин (ЗРСВ) – Бернулли и биномиального распределения. Вообще-то нас всегда интересует некоторое среднее значение x^{cp} наблюдаемой

случайной величины в виде множества $X = \{x_i\}$, $i = 1, N$ – общее число таких единиц, и её стандартное отклонение (или с.к.о) σ_x .

Вспомним ЗРСВ Бернулли на одном из примеров. «Например, при некоторых условиях вероятность появления события A (здесь – бракованного изделия) равна вероятности p . Найти вероятность того, что серия из n независимых испытаний даст k появлений и $(n - k)$ не появлений события A » [2. – С. 53]. Но нам вероятность p нам неизвестна - единая, кстати, для ВСЕХ испытаний; а у нас в табл. 2 вероятность p от партии к партии меняется. Зато в третьей строке табл. 1 мы сами зафиксировали число (именно НЕ ВЕРОЯТНОСТЬ, но конкретное дискретное по своей природе ЧИСЛО бракованных изделий (наступление того самого события A).

Так где же тогда та ЕДИНАЯ вероятность наступления (или не наступления) события A ? В общем, закон Бернулли здесь вряд ли можно использовать: вероятности-то наступления события A (брака) от партии к партии – разные (см. табл. 1). Попробуем теперь закон биномиального распределения: годится ли он для последующего описания на основе исходных данных по табл. 1?

А ведь верно, именно со стороны биномиального ЗРСВ с число успехов k , следует, что биномиальное распределение $b_{n,p}(k)$ определяет вероятность k успехов в n испытаниях в случае, когда p есть вероятность успеха в ЕДИНСТВЕННОМ испытании [3. – С. 199]. Тогда уместнее было бы обозначить (эксплицировать, формализовать то есть) ту вероятность как p_i , $i = 1, n$, где n – число изделий в испытываемых партиях. Согласно же исходной табл. 1, берутся партии именно по 100 штук готовых изделий ($n_i = n = 100 = const$).

Так что из двух, известных нам ЗРСВ для исходных данных табл. 1 содержательнее подходит именно *биномиальный* ЗРСВ. Тогда, в соответствии с выбранным нами вида ЗРСВ, и составим новую табл. 2 на основе табл. 1, дополнив её ещё одним столбцом «Сумма» и дополнительной

строкой с вероятностью p с попутной экспликацией (здесь – введением обозначений) уже перечисленных в табл. 1 переменных первого столбца:

Таблица 2

Результаты контроля качества в 10-ти партиях по 100 изделий

№ партии	1	2	3	4	5	6	7	8	9	10	Сум- ма
Объём партии, n_i , шт.	100	100	100	100	100	100	100	100	100	100	1000
Браков. изд., np_i , шт.	1	2	0	3	14	4	1	2	15	3	45
Доля события A , p_i , %	1	2	0	3	14	4	1	2	15	3	-

Примечание: поскольку p есть вероятность наступления события A , то должна вычисляться как $p = (\text{число событий } A) / (\text{общее число событий в данной партии объёмом } n_i) \cdot 100\%$, то реально $p_i = (np_i / n_i) \cdot 100\% = (1/100) \cdot 100\% = 0,01 \cdot 100\% = 1\%$ (например, для первого информационного столбца табл. 2), что и показано в последней строке табл. 2. В общем, вероятность брака для каждой партии – величина уникальная и обозначена нами как p_i . А вот число выбранных для испытаний элементов от партии к партии, напомним, – строго одинаковое: $n_i = n = 100 = const$ в данном случае.

Итак, с видом закона распределения содержательно определились. А уж далее, раз определились с видом ЗРСВ типа события A (не Бернулли, но именно биномиальный по сути своей), воспользуемся его возможностями чисто ФОРМАЛЬНО.

В отличие от формулы общего вида (20), воспользуемся её модификацией в виде формулы (22) и посмотрим на табл. 2, как её можно применить к исходным данным нашего пример 1. Всего исследовано $N = 10$ партий изделий, каждая – по 100 штук. Если всё просуммировать по строкам 2 и 3 табл. 2, то сможем вычислить так называемую «МЕХАНИЧЕСКУЮ СРЕДНЮЮ» (в отличие, скажем, от средней взвешенной, характерной для вариационных рядов [5] $p_{ср}^{cp}$) с именем $p_{ср}$, вычисленную как отношение числа браков в каждой партии к общему числу испытанных деталей продукции во всех N сериях. Искомую среднюю величину $p_{ср}$, согласно источнику [4], обозначается как CLp – средняя доля несоответствующих

единиц продукции, % при $n_i = n = 100 = const$ для табл. 1 и 2. Заметим также ещё раз, что, в отличие от ранее приведённых формул для среднего значения (20) и для с.к.о (21) далее используются их модификации в виде (22) и (23) - соответственно, применяемые в реальном контроле качества при работе, например, с p -контрольными картами:

$$CL_p = p_{cp} = \frac{\sum_{i=1}^N np_i}{\sum_{i=1}^N n_i} \cdot 100\% = \frac{\sum_{i=1}^N np_i}{N} = \frac{45 \cdot 100\%}{1000} = 4,5\%. \quad (22)$$

Теперь необходимо вычислить величину стандартного отклонения σ_A . Если нельзя напрямую воспользоваться формулой (21), которая конкретно относится лишь к отклонению ожидаемых событий k при $p_i = 1/2$ (в примере 1 явно $p_i \neq 1/2$).

Тогда для механической средней p_{cp} - формула (22) и для с.к.о. σ_p необходимо применять формулу (23) из того же источника [4. – С. 19]:

$$\sigma_p = \sqrt{\frac{p_{cp}(100 - p_{cp})}{n}}. \quad (23)$$

Тогда по формуле (21а) получим модифицированную величину с.к.о.:

$$\sigma_p = \sqrt{\frac{4,5(100 - 4,5)}{100}} = 2,07\%.$$

Тогда верхняя (ВКГ) и нижняя контрольные границы (НКГ):

$$UCL_p = p_{cp} + 3 \cdot \sigma_p = 4,5 + 3 \cdot 2,07 \approx 10,71\%. \quad (24)$$

$$LCL_p = p_{cp} - 3 \cdot \sigma_p = 4,5 - 3 \cdot 2,07 \approx -1,71\% \rightarrow 0. \quad (25)$$

Нижняя контрольная граница, как и следовало ожидать, как и в прежнем кратком примере, получилась отрицательной; следовательно. $LCL_p \rightarrow 0\%$ (ось абсцисс графика).

Итак, среднее значение вероятности наступления брака и с.к.о. от него нам известны по формуле (23), а верхнюю и нижнюю контрольные границы мы с вами также вычислили. Теперь можно отобразить результаты наших вычисления графически, то есть на основе табл. 2, выражений (22), (24) и (25) визуализировать их (см. рис. 3).

Из рис. 4 нам и ранее явно видно было, что партии с номерами 5 (14 бракованных изделий) и 9 (15 браков) должны быть из рассмотрения выведены, поскольку превышают верхнюю контрольную границу. А вот ответ на вопрос: однородна ли приведённая выборка, можно получить и другим путём, например, вычислением коэффициента вариации $h_{расч}$ [5. – С. 34] как отношения с.к.о. к её средней взвешенной и последующим сравнением с некоторым граничным значением $h_{гран}$, равным 0,30 (или 30%) в одних источниках и 0,33 (33%) – в других.

P, %

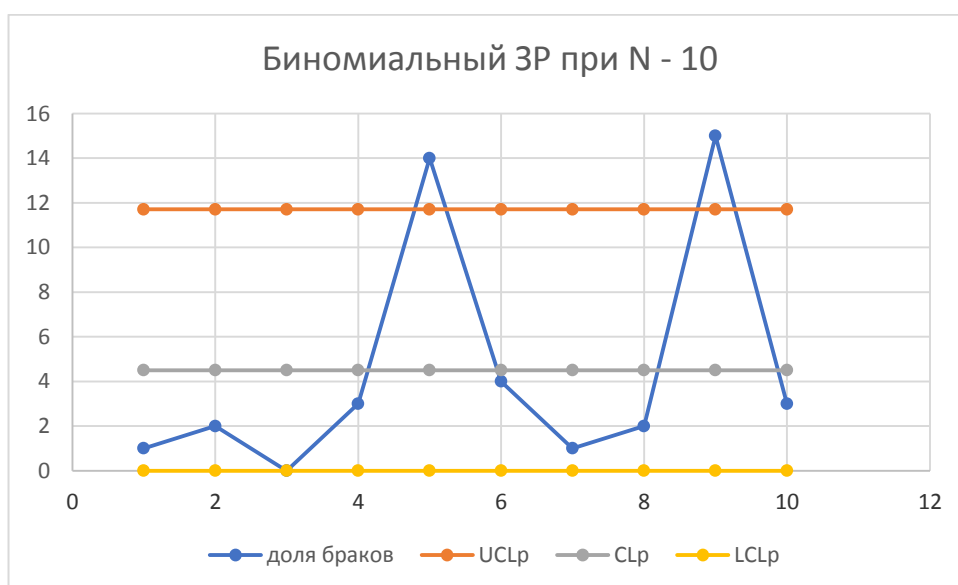


Рис. 4. Результаты выборочной статистики по биномиальному закону распределения при N = 10, n = 100 изделий в партии.

Действительно, вариабельность оказывается повышенной:

$$h_{расч} = \frac{\sigma_p}{p_{ср}} \cdot 100\% = \frac{2,07}{4,5} \cdot 100\% = 0,46 (46\%) > h_{гран} = 0,30.$$

Итак, во-первых, 5 и 9 партии принято считать бракованными. Во-вторых (без прекращения производства! – это важно) создаётся, как уже упоминалось, экспертная группа и выявляются причины брака, что будет подробнее рассмотрено при изучении процесса контроля на предприятии в

целом, например, при изучении согласно РПД Темы 8 «Организация и контроль качества продукции и профилактика брака».

А поскольку результаты 5 и 9 партии из рассмотрения вынесены, оставшиеся восемь партий по 100 контрольных изделий каждая ($n = 100$ шт.) по тем же формулам покажут следующие результаты: $CLp = p_{cp}^{cp} = 2,0\%$; $\sigma_p = 1,4\%$; $UCLp = 6,2\%$; $LCLp = -2,2\% \rightarrow 0\%$, рассчитанных по тем же формулам – (23), (24) и (25). То есть очевидно: оставшиеся *восемь* партий будут признаны качественными относительно верхних и нижних контрольных границ. Построим график и для оставшихся 8-ми партий (кроме 5-й и 9-ой) по той же табл. 2. Рассчитаем новые показатели для восьми, а не 10-ти.

Итак, согласно формуле (14) по вычислению (здесь – взвешенного) среднего:

$$CLp = p_{cp} = \frac{\sum_{i=1}^N np_i}{\sum_{i=1}^N n_i} \cdot 100\% = \frac{(45 - 14 - 15) \cdot 100\%}{1000 - 100 - 100} = 2,0\%.$$

Для расчёта требуемых границ необходимо также определить новое значение с.к.о. по формуле (21а):

$$\sigma_p = \sqrt{\frac{2,0(100 - 2,0)}{100}} = 1,4\%.$$

Тогда верхняя и нижняя границы определяются легко по (24) и (25):

$$UCLp = 2,0\% + 3 \cdot 1,4\% = 2,0 + 3 \cdot 1,4 = 6,2\%.$$

$$LCLp = 2,0\% - 3 \cdot 1,4\% = 2,0 - 3 \cdot 1,4 = -2,2\% \rightarrow 0.$$

Так как $-2,2\% < 0$ (отрицательного брака в природе не существует), то вместо расчётного значения ($-2,2\%$) принимаем $LCLp = 0$, как и в формуле (25).

Построим новый график, аналогичный рис.4 того же процесса, приведённый на рис. 5.

Как видно из рис. 5, текущее значение для событий A как p , % находится уже в пределах верхней и нижней контрольных границ. Поставленная задача в данном примере наконец-то решена.

Вычислим заодно и новое значения коэффициента вариации:

$$h_{\text{расч}} = \frac{\sigma_p}{p_{\text{ср}}} \cdot 100\% = \frac{1,4}{2,0} \cdot 100\% = 0,70 \text{ или } 70\% > h_{\text{гран}} = 0,30.$$

Да, этот коэффициент уже 70% вместо бывшего 46%. Но вычисление такого коэффициента в первом случае достаточно продуктивно, тогда как на втором этапе вычислений с учётом удалённых партий с номерами 5 и 9 он лишь характеризует распределение бракованных изделий уже внутри вновь вычисленных границ и может быть только принят к сведению, что даже полученная картина является статистически неоднородной, что может быть также дополнительным мотивом для совершенствования схемы контроля соответствующей экспертной группы.

P, %

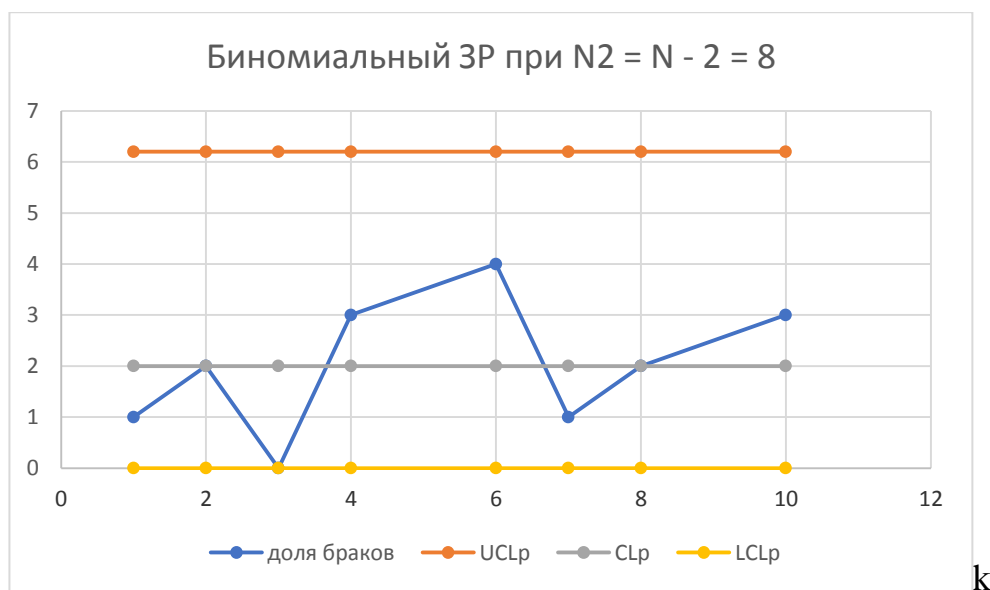


Рис. 4. Результаты выборочной статистики по биномиальному закону распределения при N2 = 8, n = 100 изделий в партии (партии 5 и 9 удалены).

Итак, нами был только что рассмотрен пример выбора схемы контроля, исходя из условий табл. 1, где каждый номер партии предполагал контроль по 100 изделий (n = 100 шт.). Случай, так сказать, классический (для нас – модельный, то есть наглядный и простой). Решение примера 1 завершено.

Однако, рассмотрим и следующий за этим пример, решённый нами на одном из практических занятий, а впоследствии сравним результаты уже решённого примера 1 и следующего за ним примера 2. На таком примере мы хотим показать, что проведённые нами ранее рассуждения и расчёты можно существенно расширить применительно к иным областям деятельности, например, в области оценки качества учебного процесса того или иного вуза.

Пример 7. Итак, в отличие от предыдущего примера, где в каждой партии исследовалось по 100 единиц продукции и фиксировались бракованные изделия по принятой на предприятии измерительной методике. В данном же примере объектом исследования являются не партии по 100 единиц, а *отдельные кафедры* отдельного вуза на предмет оценки аттестованных в середине семестра и не аттестованных студентов. Такие исходные данные по кафедрам приведены в табл. 3[4. – С. 27 - 29].

Таблица 3

Результаты внутрисеместровой аттестации студентов вуза

№ кафедры вуза	1	2	3	4	5	6	7	8	9	10	Сум-ма
Число аттестуемых	42	27	14	26	38	34	45	47	10	30	313
Аттестов. успешно	15	9	2	10	11	13	11	15	4	14	104

Снова исходные данные надо как-то формализовать (эксплицировать). Как и прежде: вторую строку обозначим через n_i , а третью – как np_i . Тогда вероятности p_i как % студентов, прошедших аттестацию: $p_i = (np_i/n_i) \cdot 100\%$: что и представим с изменённой первой графой на той же табл. 3 (табл. 4):

Таблица 4

Экспликация исходных данных и оценка доли аттестованных, %

№ кафедры вуза	1	2	3	4	5	6	7	8	9	10	Сум-ма
Число аттестуемых, n_i	42	27	14	26	38	34	45	47	10	30	313
Аттестов. успешно, np_i	15	9	2	10	11	13	11	15	4	14	104
Доля аттестов-ых, $p_i, \%$	35,7	33,3	14,3	38,5	28,0	38,2	24,4	31,9	40,0	46,7	-

Обратим внимание на то, что в первом примере среднее значение $p_{\text{вс}}^{\text{ср}}$ находилась нами через отношение сумм $\sum np_i$ к $\sum n_i$ по прежним пределам

сумм с последующим умножением на 100% при переходе от относительных единиц $p_i = [0; 1]$ к её процентному выражению $p_i = [0\%; 100\%]$. Таким же образом находится среднее и во примере 7 по формуле (22): $CLp = p_{cp} = (101/313) \cdot 100\% = 33,2\%$. Однако, если в примере 6 были все партии по 100 деталей, то в примере 7 каждая партия уникальна по числу аттестуемых студентов. То есть $p_i \neq p \neq \text{const}$. Поэтому и величины с.к.о. для каждой кафедры – УНИКАЛЬНЫЕ, но находится по той же формуле (21а). Поэтому для каждой кафедры будут уникальными и верхняя и нижняя контрольные границы по формулам (24) и (25) - соответственно. Только для табличного удобства найденные, уникальные для каждой кафедры с.к.о., будем сразу в соответствии с формулами вычисления $UCLp$ и $LCLp$ либо прибавлять слагаемое $3\sigma_p$, либо вычитать его – соответственно. Остальные результаты приведены в строках, добавленных к табл. 4 в табл. 5.

Ну и предварительно выполним вспомогательные вычисления по суммированию числовых показателей во второй и третьей строке табл. 4: $\sum n_i = 313$ (общее число аттестуемых); $\sum np_i = 104$ (общее число аттестованных).

Итак, по формуле (22) вычислим новое значение среднего взвешенного:

$$CLp = p_{cp} = \frac{\sum_{i=1}^N np_i}{\sum_{i=1}^N n_i} \cdot 100\% = \frac{\sum_{i=1}^N np_i}{N} = \frac{104 \cdot 100\%}{313} = 33,2\%.$$

А вычисления по формулам для верхней и нижней границы проделаем по тем же формулам (24) и (25), но уже для каждой кафедры в табл. 5:

Как видно из табл. 5, механическое среднее $p_{cp} = 33,2\%$ не выходит за пределы верхней и нижней контрольных границ UCL и $LCLp$ -соответственно, как, впрочем, и число аттестованных на каждой кафедре студентов, что и представим графически на рис. 4. Согласно табл. 5 и её графической иллюстрации на рис. 6, несмотря на разрозненную картину

Таблица 5

Оценка среднего, контрольных границ и долей аттестованных

№ кафедры вуза	1	2	3	4	5	6	7	8	9	10	Сум-ма
Число аттестуемых, n_i	42	27	14	26	38	34	45	47	10	30	313
Аттестов. успешно, np_i	15	9	2	10	11	13	11	15	4	14	104
Доля аттестов-ых, $p_i, \%$	35,7	33,3	14,3	38,5	28,0	38,2	24,4	31,9	40,0	46,7	-
$3 \cdot \sigma_p$ по формуле (23)	22,2	27,2	28,1	28,6	22,1	25,0	19,2	20,4	46,5	27,3	-
UCL _p (ВКГ) по (25)	55,4	60,4	61,3	61,8	55,3	58,2	52,4	53,6	79,7	60,5	-
P_{cp} по формуле (22)	33,2										-
LCL _p (НКГ) по (24)	11,0	6,0	5,1	4,6	11,1	8,2	14,0	12,8	0	5,9	-

числа аттестованных на ВСЕХ кафедрах, объективно получается, что число аттестованных студентов находится в норме – между НКГ и ВКГ.

$P, \%$

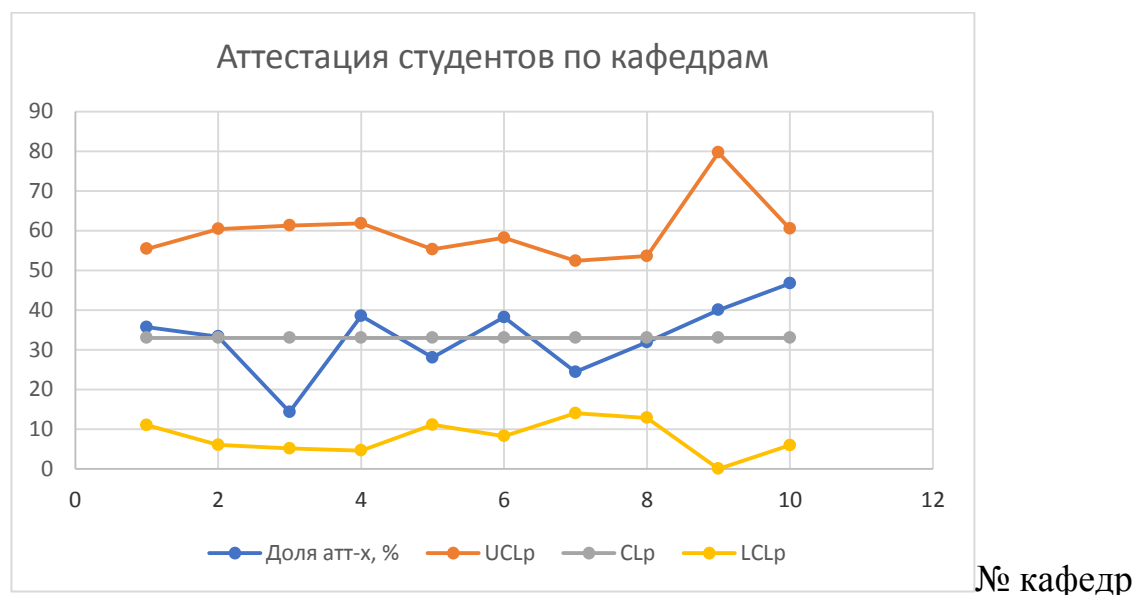


Рис. 6. Визуализация данных табл. 5: число аттестованных находится между верхней и нижней контрольных границ.

Данное обстоятельство ещё раз свидетельствует о том, что предлагаемый здесь расчётный материал изначально по качеству контроля, оказывается не предметно-ориентированной, но методо-ориентированной конструкцией. А это уже МЕТОДОЛОГИЧЕСКИЙ вывод, что для нас здесь весьма важно. То есть ранее рассмотренный на данный предмет материал, если делать это методологически корректно, можно распространять и на иные области деятельности. Так что степень сходства и различия между

примерами 6 и 7 нам теперь известна: для этого надо только сравнить графики, приведённые на рис. 4 и рис. 5.

И всё же остаётся невыясненным следующий вопрос.

Нам по математической статистике всегда было иметь привычное дело с такой разновидностью ЗРСВ, как НОРМАЛЬНЫЙ ЗАКОН РАСПРЕДЕЛЕНИЯ Карла Фридриха Гаусса. Отметим только, что распределение исследуемой СВ по этому закону имеет место, когда действующих на сам процесс причин – множество, часть из них взаимно компенсируют друг друга, и в итоге фиксированных средних значений в распределении получается заметно больше, чем также фиксированных значений по его краям. Поэтому на данном этапе рассмотрения законов распределения случайных величин вполне уместным будет воспользоваться новым для нас, Гауссовым подходом, для тех же исходных данных, приведённых в эксплицированной табл. 2 по табл. 1, в которой представлены исходные данные.

Отметим для себя только то, что закон распределения СВ Гаусса органично присущ именно непрерывной СВ, в отличие от ЗР по схеме Бернулли (т.н. биномиального) и ЗР Пуассона. Однако наши намерения вполне корректны: непрерывную СВ всегда можно отобразить в виде не дискретного, но интервального вариационного ряда (ВР) в виде n вариантов-строк как небольшой рабочей таблицы с дискретными же характеристиками его вариантов-строк по столбцам (графам) «эмпирическая частота f_i » и «средние значения интервалов ряда x_i^{cp} ».

Попробуем это сделать по рекомендациям источника [5. – С. 43 - 49] для проверки распределения элементов в примере 6 (табл. 2) с проверкой рабочей гипотезы распределения по Гауссу. Иначе говоря, при рассмотрении биномиального закона распределения нам представляется полезным естественное его сравнение с применением такого знакомого уже нам закона распределения, как нормального ЗР СВ К.Ф. Гаусса. В этой связи подробнее рассмотрим сам процесс построения ВР.

3.5. Проверка применения нормального закона распределения

Эту проверку мы проведём на исходных данных того же примера 6, табл. 2. Для начала (ведь у нас остались только 8 измерений после исходных 10-ти, после удаления из модели результатов №5 и №9 из общей картины контроля, ведь так? Перечислим их, эти самые восемь значений: 1; 2; 0; 3; 4; 1; 2; 3. То есть вместо исходных 10-ти значений имеем только прошедшие контроль восемь: $N_2 = 8$. Небольшой, конечно, ряд цифр, но попробуем с ним хоть что-нибудь сделать. Пусть это будет эксплицировано нами как некоторая статистическая совокупность $Y = \{y_j\}, j = 1, N_2 = 10 - 2 = 8$. Для проверки на гауссовское распределение необходимо сначала построить вариационный ряд (ВР) для определения опытного значения частот $f_j, j = 1, 8$, которые, в свою очередь, могут быть средством для сравнения с теоретическими частотами по критерию уже χ^2 -Пирсона с целью определения: подходит ли наш исходный ряд элементов как статистической совокупности под нормальный закон распределения Гаусса или не подходит. Итак, сначала надо построить вариационный ряд. А сам процесс его построения можно рассматривать как операцию отображения исходного множества Y с m элементами СВ на более компактное множество X с n вариантами-строками (как отображение m элементов на n):

$$\tau: Y \rightarrow X.$$

То есть построение ВР есть как бы «сжатие исходной информации», более удобной для последующих вычислений, в чём нам легко убедиться. И действительно: во множестве СВ с именем Y было $N_2 = 8 = m$ исходных элементов, а будет новое множество X (в виде таблицы, правда) с $n = 4$ вариантами-строками.

Построение ВР. Для этого надо определиться с его величиной шага $h = R/n$, где под символом R понимается размах выборки: $R = y_{max} - y_{min} = 4 - 0 = 4$ (%); n – как уже упоминалось, число интервалов в ВР (вариантов) – см. формулу (26).

Построение ВР существенно усложняет предыдущие расчёты для обязательного нахождения эмпирических (опытных) частот f_i , $i = 1, N2 = 8$, а уж потом найти среднее взвешенное, с.к.о., а также уже дополнительно составить ещё одну таблицу для нахождения т.н. «теоретических частот» $f_i^{\text{теор}}$, для чего надо ещё и войти в специальную таблицу Гаусса в приближении Стьюдента и определить (иногда с округлением до целых значений) теоретические частоты, взятых из той же, построенной нами новой таблицы.

Сравнение тех и других частот по Пирсону мы и узнаем, можно ли исходную статистическую совокупность признать как распределённую по нормальному закону или нет. Если да, то воспользуемся его формулами для вычисления критерия χ^2 -Пирсона с последующим сравнением вычисленного и табличных значений при заданном уровне значимости α .

Число вариантов (строк в ВР) устанавливается по приблизительной формуле Стерджесса [6] при $N2 = 8$:

$$n \approx 1 + 3,322 \lg N2 = 1 + 3,322 \lg 8 = 1 + 3,322 \cdot 0,9 = 3,99 \approx 4 \text{ (инт.)}. \quad (26)$$

Следовательно, искомый шаг для построения ВР будет следующим:

Оценим с учётом формулы (26) величину шага $h = R / n = 4\% / 4 = 1,0\%$, начиная от минимального значения, равного нулю. Теперь можно начинать строить ВР, что есть таблица с n - числом строк (вариантов) в табл. 6. То есть мы тем самым и строим уже упомянутое отображение τ :

$$\tau: \text{Ряд с числом эл-тов } N2 = 8 \rightarrow \text{ВР с числом эл-тов } n = 4. \quad (27)$$

То есть вследствие построения отображения (27) мы «сжали» исходную информацию с 8-ми единиц до 4-х. А если бы исходное множество U содержало бы не 8, а, скажем, 100 единиц? – Тогда по формуле Стерджесса (26) нам для их отображения τ потребовалось бы только примерно 6 – 7 вариантов-строк ВР. Если бы множество U состояло из 1000 элементов, то вследствие того, что $\lg 1000 = 3$, нам потребовалось бы по формуле (26) 9 -10 строк всего-то. Итак, ВР есть *средство сжатия* информации (см. табл. 6):

Примечание: знак (+) в заглавии строк означает, что, согласно [6], интервалы по правой стороне считаются закрытыми, а слева - открытыми.

Таблица 6

Вариационный ряд исходной совокупности $N_2 = 8$ единиц, где $n = nvr = 4$ – число вариантов

Варианты, i	Границы инт. $x_{нач} \div x_{кон} (+)$	Эмпирические частоты f_i		Накопленная частота q_i	Средние значения x_i^{cp} интервалов
		счёт из N_2	значения		
1	2	3	4	5	6
1	0,0 – 1,0	///	3	3	0,5
2	1,0 – 2,0	//	2	5	1,5
3	2,0 – 3,0	//	2	7	2,5
$nvr = 4$	3,0 – 4,0	/	1	8	3,5
Всего:	-	8	8	-	-

Здесь для нас цель построения ВР (назначение которого, как известно, значительно шире) – получение эмпирических частот f_i и получение середин интервалов x_i^{cp} . Для продолжения необходимо получить не только эмпирические частоты f_i , но и их т.н. «теоретические значения» $f_i^{теор}$ для последующего применения критерия согласия χ^2 -Пирсона.

Но для этого сначала необходимо найти среднее взвешенное значение $p_{вз}^{cp}$ и на его основе - стандартное с.к.о. σ_p от взвешенного среднего $x_{cp}^{вз}$. Рассчитаем это по методике, представленной в [5]. Но среднее взвешенное в терминах ВР, в отличие от прежней формулы (22), да и с.к.о. – также ВЗВЕШЕННЫЕ (тоже в терминах ВР: и то, данные в столбце 6 табл. 6 встречаются неодинаковое число раз, а сколько в каждом интервале – см. столбец 3 той же таблицы):

$$x_{cp}^{вз} = \frac{\sum_{i=1}^n x_i^{cp} f_i}{\sum_{i=1}^n f_i} = \frac{0,5 \cdot 3 + 1,5 \cdot 2 + 2,5 \cdot 2 + 3,5 \cdot 1}{3 + 2 + 2 + 1} \approx 1,62. \quad (28)$$

$$\sigma_p = \sqrt{\frac{\sum_{i=1}^n (x_i^{cp} - x_{cp}^{вз})^2 \cdot f_i}{\sum_{i=1}^n f_i}} = \sqrt{\frac{8,875}{8}} = 1,05. \quad (29)$$

Итак, нами получены взвешенные - среднее значение и с.к.о. σ_p вокруг него – по три σ_p от центра $x_{cp}^{вз} = p_{вз}^{cp}$, где непрерывная функции Гаусса в виде колпака будет в этой точке иметь максимум. Всё мы это делали для того, чтобы произвести так называемое «выравнивание вариационных рядов», то

есть сделали предпосылки к т.н. «построению теоретического распределения» [5. – С. 45].

Эмпирические частоты нами получены ещё в ВР, приведённым в табл. 6: это f_i , которые были определены нами для всех четырёх интервалов ВР (табл. 6) как 3, 2, 2 и 1 (всего восемь = $N/2$, ничего не «потеряно» то есть).

Ну а далее следует отметить, что кривая Гаусса изначально была равна $(2\pi)^{1/2}$ (что поясним несколько ниже). Ну а чтобы её привести к единице, на это число всё и делят. Известный интеграл Гаусса здесь не приводим, но по предмету исследования сообщаем, что «при выравнивании вариационного ряда (ВР) по кривой нормального распределения» теоретические частоты ряда $f_i^{теор}$ определяются по известной формуле [5. – С. 128]:

Далее сначала вычисляются вспомогательные параметры - нормированные отклонения

$$t_i = \frac{(x_i^{cp} - x_{cp}^{вз})}{\sigma_p} \text{ и}$$

$$\text{const} = \frac{Nh}{\sigma_p} = \frac{8 \cdot 1}{1,05} \approx 7,62 (\%)$$

для последующего вычисления вероятности наступления каждого события (здесь x_i^{cp} , когда i меняется от 1 до $n = 4$) по следующей формуле, которая приводится в приложениях к учебникам по статистике, например, в [2. - С. 128]:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \quad (30)$$

Поэтому будущую расчётную табл. 7 на основе табл. 6, видимо, придётся существенно расширить.

Построение продолжения табл. 6 для ВР в виде табл. 7. Итак, таблица в виде ВР (табл. 6) нами уже построена. Остаётся лишь вычислить значения t_i по числу вариантов-строк ($n = 4$) и войти в статистическую таблицу, табулирующую зависимость $\varphi(t)$, например, в [5. – С. 128] следующим образом: например, в графе 4 табл. 7 в первой строке для первого

интервала записана величина « $t_1 = -1,07$ » (здесь и далее – округлять до сотых долей единицы). Входим в таблицу [5. – С. 128] по абсолютной величине $t_1 = |-1,07| = 1,07$ и находим в таблице значение «2251», которую следует записать в графу 5, строку 1 как 0,2251 для $\varphi(t_1) = 0,2251$ (см. табл. 7, столбец 5, первая строка).

А так как величина $const = 7,62$ нами уже вычислена, то найденное число $\varphi(t_1)$ необходимо на него умножить и получить искомое значение $f_1^{meop} = 7,62 \cdot 0,2251 = 1,72$ (штуки) ≈ 2 штуки (попаданий числа исходных значений СВ y_j в первый интервал ВР $(0,0 \div 1,0)$ – см. табл. 6, графа 2. Хотя в данный интервал попало 3 значения исходной СВ y_j , то есть значение эмпирической (взятой из опыта) частоты $f_1 = 3$. А теоретическая частота $f_1^{meop} = 1,72$ или 2, если округлять до целых (в разных задачах рекомендуется по-разному). На наш взгляд, надо проверять и для одного варианта, и для другого. В таблице 7, например, расчёты велись с округлёнными значениями теоретических частот.

А далее следует только воспользоваться либо критериями согласия [5. – С. 45 – 49] Пирсона χ^2 , либо Романовского Rom или Колмогорова λ . То есть для этого необходимо достроить табл. 6 до табл. 7 с получением искомых значений частот, но уже теоретических f_i^{meop} ; $i = 1, nvr = n = 4$. А уже после построения табл. 7 непосредственно приступаем к расчёту т.н. «критериев согласия».

3.6. Расчёты критериев согласия

Сначала мы выдвигаем гипотезу (предположение) о том, что элементы СВ y распределены по закону Гаусса, которая может быть подтверждена последующими расчётами или опровергнута. Но прежде надо найти теоретические частоты f_i^{meop} для каждого интервала построенного нами ВР,

поскольку эмпирические частоты f_i для тех же интервалов-вариантов мы получили ещё на этапе построения ВР.

Для этого на основе табл. 6 строим дополнительную расчётную табл. 7 для получения теоретических значений частот распределения вариантов в обязательном ВР $f_i^{теор}$. Из табл. 6 позаимствуем лишь столбец 6 со средними значениями i -вариантов ($i = 1, n = 4$) для вычисления $(x_i^{cp} - x_{вз}^{cp})$ – здесь для большей ясности $p_{вз}^{cp} = x_{вз}^{cp}$, что понятно и без дополнительных комментариев.

Таблица 7

Вычисление теоретических частот $f_i^{теор}$ для ранее построенного ВР

Варианты, I	Средние значения x_i^{cp} интервалов	$(x_i^{cp} - x_{вз}^{cp})$	$t = (x_i^{cp} - x_{вз}^{cp})/\sigma_p$ (округлять до сотых; знаки + и – не учитывать)	$\varphi(t)$ из Приложения 1 [5. –С. 128]	$f_i^{теор} = (Nh/\sigma_p) \cdot \varphi(t)$ с округлением до целых
1	2	3	4	5	6
1	0,5	- 1,12	- 1,07	0,2251	1,72 \approx 2
2	1,5	- 0,12	- 0,11	0,3965	3,02 \approx 3
3	2,5	0,88	0,84	0,2803	2,14 \approx 2
n = 4	3,5	1,88	1,79	0,0804	0,61 \approx 1
Всего теоретических частот:					7,49 \approx 8

В итоге расчётов в табл. 7 получили два значения теоретических частот: округлённые их значения до целых чисел и их же неокруглённые, вещественные значения.

Итак, всего после первого контроля (см. табл. 2 с иллюстрацией на рис. 3) осталось из $N = 10$ испытаний всего $N_2 = 8$ (штук) испытаний (5-е и 9-е с числом бракованных изделий числом 14 и 15 из каждой сотни деталей $n = 100$ были из рассмотрения нами как экспертами выведены). Так что оставшиеся 8 испытаний дали нам основание для построения ВР с определением числа появления того или иного варианта – см. табл. 6, то есть эмпирических частот f_i общим их числом также 8. Хотя и сумма теоретических частот из округлённых частных значений по вариантам ВР получилась также равной 8-ми, сумма неокруглённых значений составила всего $7,49 < 8$, что, вообще говоря, настраивает нас делать расчёты по критерию согласия Пирсона для

обоих вариантов – для округлённых значений теоретических частот и для неокруглённых.

Итак, согласно табл. 6, – *воплощения построенного нами ВР*, - мы насчитали в сумме восемь появлений СВ и четыре опытных (**эмпирических**) частот f_i - по числу полученных интервалов для ВР. Применяв необходимые преобразования в табл. 7, также получили в сумме те же восемь СВ, но уже **теоретических частот $f_i^{теор}$** для тех же интервалов ВР. Это и даёт нам наконец-то основания для применения любого доступного нам *критерия согласия*. Начнём с критерия согласия Пирсона χ^2 [5. – С. 48]:

$$\chi_{расч}^2 = \sum_{i=1}^n \frac{(f_i - f_i^{теор})^2}{f_i^{теор}}, \quad (31)$$

Для этого выражения (31) для удобства вычислений требуется построить дополнительную расчётную таблицу. Пусть ей будет табл. 8 следующего вида (в данной таблице мы принимаем во внимание лишь округлённые значения теоретических частот $f_i^{теор}$).

Расчётное значение $\chi_{расч}^2$ критерия согласия χ^2 -Пирсона нами в табл. 7 получено: $\chi_{расч}^2 = 0,83$. Для проверки справедливости выдвинутой нами ранее гипотезой о распределении исследуемой СВ – множества с именем У, - по нормальному закону распределения (НЗР) необходимо и достаточно сравнить полученное нами расчётное значение параметра Пирсона с его табличным значением, которое, в свою очередь, находится из статистической таблицы «Значение χ^2 -критерия Пирсона при уровне значимости 0,10; 0,05; 0,01» в зависимости от числа степеней свободы df .

Таблица 8

Расчёт значения критерия Пирсона $\chi_{расч}^2$

Варианты, I	Эмпирические частоты f_i	Теоретические частоты $f_i^{теор}$	$(f_i - f_i^{теор})$	$(f_i - f_i^{теор})^2$	$(f_i - f_i^{теор})^2 / f_i^{теор}$
1	2	3	4	5	6
1	3	2	1	1	0,50
2	2	3	- 1	1	0,33
3	2	2	0	0	0

n = 4	1	1	0	0	0
Всего:	8	8	-	-	$\Sigma = \chi^2_{\text{расч}} = 0,83.$

Найденное табличное значение критерия $\chi^2_{\text{табл}}(df, \alpha)$, где α – уровень значимости, сравнивается с его расчётным значением. Если $\chi^2_{\text{табл}}(df, \alpha) > \chi^2_{\text{расч}}$, то расхождения между эмпирическими и теоретическими частотами можно считать случайными, и выдвинутая нами гипотеза о близости эмпирического распределения к нормальному не опровергается[5. – С. 46].

Для такой проверки остаётся уточнить лишь значение степеней свободы $df = n - r - 1$, где r – число параметров эмпирического распределения, использованных для нахождения теоретических частот. А для этого мы использовали в табл. 6 и 7 среднее взвешенное ВР и его стандартное отклонение: всего в сумме два параметра, поэтому в случае проверки гипотезы о нормальном распределении $r = 2$. Тогда $df = n - r - 1 = 4 - 2 - 1 = 1$, и при принятым нами (что и рекомендуется рядом авторов) $\alpha = 0,05$ (ошибка не превышает 5%, то есть достоверность 95%) табличное значение $\chi^2_{\text{табл}}(df = 1, \alpha = 0,05) = 3,84$. Тогда по графе 6 табл. 7 для округлённых и точных значений $f_i^{\text{расч}}$ будет таким:

$$\chi^2_{\text{расч}} = 0,83 < \chi^2_{\text{табл}}(df = 1; \alpha = 0,05) = 3,84 - \text{для целых } f_i^{\text{расч}};$$

$$\chi^2_{\text{расч}} = 1,56 < \chi^2_{\text{табл}}(df = 1; \alpha = 0,05) = 3,84 - \text{для дробных } f_i^{\text{расч}}.$$

Таким образом, расхождения между эмпирическими частотами f_i и теоретическими частотами $f_i^{\text{теор}}$ по всем $n = 4$ вариантам ВР можно с достоверностью 95% считать случайными, и выдвинутую ранее гипотезу о близости эмпирического распределения к нормальному закону распределения не опровергается. В итоге выдвинутая нами «нулевая гипотеза» о нормальном ЗР исследуемой СВ *подтверждается*.

Подтверждается даже несмотря на видимую нами разницу между данными столбцов 2 и 3 табл. 8. То есть в данном случае видимая нами разница между этими столбцами является т.н. «статистически кажущейся».

Ну и далее в качестве альтернативного критерия согласия воспользуемся хотя бы критерием Романовского (впервые опубликован автором в 1938 году на научной конференции в Средней Азии). Для этого вычисляется сам критерий $Rom_{расч}$ и сравнивается с граничным $Rom_{гран} = 3$. Если $Rom_{расч} < Rom_{гран}$, то гипотеза о нормальном распределении не отвергается. Итак, из [5. – С. 48 – 49] по известной формуле (32):

$$Rom_{расч} = \frac{|\chi^2 - df|}{\sqrt{2df}} = \frac{|0,83 - 1|}{\sqrt{2}} = 0,12 < Rom_{расч} = 3. \quad (32)$$

Гипотеза о нормальном распределении исходной статистической совокупности, отражающей процент бракованных изделий (появление события A в примере 1) в исследуемых партиях по 100 штук деталей (готовой продукции) $p_i, \%$ (см. табл. 2 примера 1) и после применения критерия согласия Романовского также подтверждается.

Заметим только, что критерий Романовского, как показывает практика, несколько даже «мягче» критерия Пирсона. Ну а критерий Колмогорова рекомендуется использовать, когда исследуемая совокупность близка, скажем, к 50-ти единицам, что к нашему примеру 6 отнести при испытании лишь 10 партий, а теперь и из 8-ми партий (хоть и по 100 единиц в каждой) также достаточно сложно.

Однако, на исходных данных примера 6 после получения среднего значения и величин с.к.о. теперь уже в рамках нормального закона распределения (а не биномиального, как с самого начала решения примера 6) можно оценить среднее значение, а также верхнюю и нижнюю контрольные границы:

Раньше при иллюстрации биномиального закона распределения случайной величины при решении пример 6 мы исходили из рассуждений (гипотетических, концептуальных положений, по содержанию то есть, без каких-либо расчётов) о наступлении достаточно редких событий (см. табл. 1 и 2), пришли к выводу, что исследуемая нами величина наступления события A (появление в партии изделий брака p_i , $i = 1, N = 10$ партий), - когда сами

«партии», состояли из целых 100 (!) изделий, - пришли к выводу о целесообразности использования именно БИНОМИАЛЬНОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ. Далее, в рамках принятого нами ЗРСВ, мы оценили как 4,5% по формуле (20а) или (23), рассчитали с.к.о. $\sigma_p = 2,07\%$ как вспомогательный параметр по формуле (21а), а с помощью него по соответствующим формулам и обе контрольные границы: верхнюю $UCL_p = 10,71\%$ и нижнюю $LCPL = - 1,71\% \rightarrow$ т.е. 0,00%. Ещё для наглядности представили полученные результаты графически на рис. 3.

Вот тогда-то и получили все основания партии №5 ($p_5 = 14$ браков $> UCL_p = 10,71\%$) и №9 ($p_9 = 15$ браков, также больше той же UCL_p) из потока поступления продукции потребителю **ИСКЛЮЧИТЬ** как **бракованные** изделия. Напомним только, что «проценты, %» и штуки у нас совпадали (см. табл. 2), поскольку партии состояли из 100 изделий.

Словом, напоминая содержание примера 6, вывели партии с номерами №5 и №9 из последующего рассмотрения и снова посчитали необходимые показатели уже в новом, уменьшенном составе годных изделий при $N_2 = 8$ партиям: среднее число бракованных изделий (они же и в процентах в примере 6 те же), а также значение с.к.о.- по формуле (21а); новое значение $\sigma_p = 1,05\%$. Затем, применим закон «трёх сигм» (от $- 3\sigma$ до $+ 3\sigma$). То есть под «колоколом» или «колоколом Гаусса» располагаются значения 99,7% всех элементов нормального распределения. Далее – как и прежде: прибавили по 3 штуки с.к.о. к среднему по формуле (24) или отнимая их же по формуле (25), получим новые верхнюю и нижнюю границы, но уже для оставшихся $N_2 = N - 2 = 10 - 2 = 8$ партий.

В случае же, если мы этак РЕШИЛИ ВДРУГ выдвинуть т.н. СВОЮ новую «нулевую гипотезу»: не биномиального закона распределения. но **НОРМАЛЬНОГО ЗАКОНА РАСПРЕДЕЛЕНИЯ** исследуемой случайной величины – браков в каждой партии. Или, то же самое, появление/непоявления исследуемого события A . Тогда в терминах повторённого нами рис. 4, среднее значение будет вычислено по формуле (22) и равно $p_{\text{вз}}^{cp} =$

1,62%; $\sigma_p = 1,05\%$ по формуле (21a). Тогда $UCLp = p^{cp}_{\epsilon_3} + 3\sigma_p = 1,62 + 3 \cdot 1,05 = 4,77 (\%)$; $LCLp = p^{cp}_{\epsilon_3} - 3\sigma_p = 1,62 - 3 \cdot 1,05 = - 1,53 (\%) \rightarrow 0$. Отообразим результаты на рис. 7.

Сравнивая содержание рис. 7, отражающего БИНОМИАЛЬНОЕ распределение, и нового НОРМАЛЬНОГО распределения на рис. 8 видим, что исходные доли браков A остались прежними (см. «Доля события A , p_i , %»), как и нижние контрольные границы: все они приравнены нами нулю. Зато изменилось положение верхней контрольной границы: в случае нормального распределения она существенно снизилась с 6,20 штук (или процентов – партии-то по 100 штук) до 4,77 штук (или процентов). Однако же, общая картина осталась прежней: все партии, кроме №5 и №9 хоть для рис. 5 (наша гипотеза о биномиальном распределении), хоть для рис. 7 (тоже наша гипотеза, но уже о нормальном распределении) оказались не забракованными (!). И это важно.

Но при сравнении рис. 5 и рис. 7 вполне очевидно, что верхняя граница $UCLp = 4,77\%$ почти приблизилась к максимальному числу бракованных изделий ($y_{max} = y_5 = 4\%$ бракованных изделий), и появилась угроза как бы ещё одного брака для всей партии из 100 деталей. Хотя и по рис. 5 даже партия из 100 изделий, содержащая 4 брака ранее считалась вполне пригодной для производства. В общем, пришла пора вспомнить известную всем *максиму*: есть две принципиальные ошибки – ошибка 1-го рода в пропуске брака и ошибка 2-го рода – забракование годного изделия. Ошибки обоих родов являются недопустимыми в процессе контроля качества произведённой на предприятии продукции.

Итак, оба рисунка - рис. 5 визуализирует нашу рабочую гипотезу о биномиальном законе распределения бракованных изделий в партиях, а рис. 7 – ту же картину, но при гипотезе о нормальном распределении тех же браков как случайных величин. В итоге мы имеем одну и ту же качественную картину, но нормальный закон является более «требовательным» за счёт снижения величины верхней контрольной границы с 6,20% до примерно

4,77% (снизилась приблизительно на 23% или почти на одну пятую, что существенно). Поэтому проведём более подробные сравнения.

Проценты, %

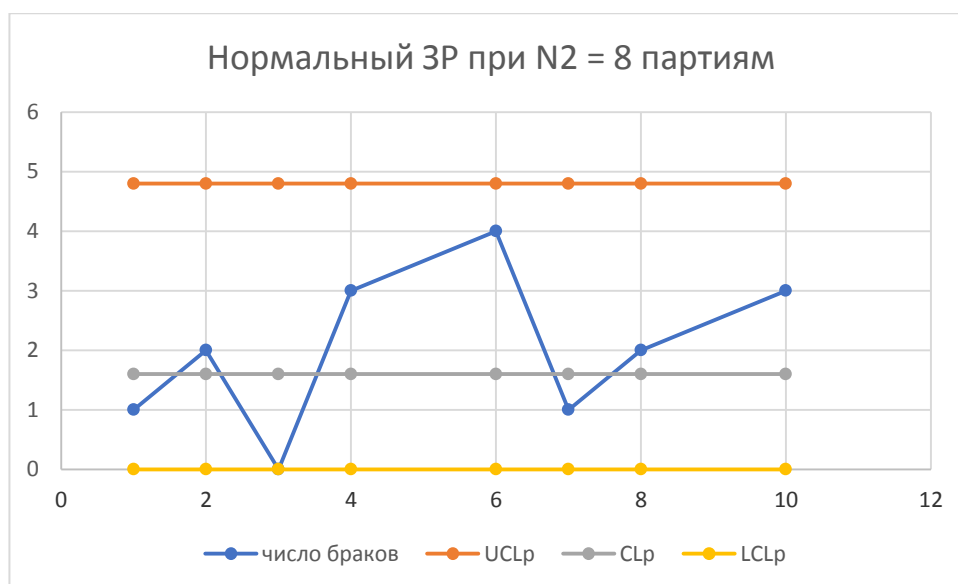


Рис. 7. Результаты выборочной статистики по нормальному закону распределения при том же исключении из анализа партий №5 и №9.

Таким образом, уместность применения биномиального закона распределения и (или) нормального закона распределения нами была рассмотрена на одних и тех же исходных данных, приведённых в примере 6. Остаётся лишь сравнить результаты применения обоих законов распределения с оценкой их как с положительной, так и отрицательной сторон, так или иначе связанных с производством необходимых вычислений. Иногда, как будет показано далее, бывает предпочтительным отчасти поступиться строгостью самого подхода (например, к проверке рабочей гипотезы о ЗР СВ), если ожидаемый результат сможет вполне оправдать себя в производственном плане в целом.

3.7. Результаты сравнения двух законов распределения.

Во-первых, исходные данные в табл. 1 и, будучи нами же эксплицированными, в том же виде, только с новыми обозначениями

приведены и в табл. 2 – дубле табл. 1 по своей природе. И что мы видим? – Да, закон биномиального распределения (классический аналог его, как вы помните, состоит в подбрасывании шестигранной кости) ориентирован ВСЕГДА на дискретные данные, что мы и наблюдаем в табл. 1 (или табл. 2). Тогда как закон распределения тех же случайных величин Гаусса ориентирован именно на НЕПРЕРЫВНЫЙ ряд данных (иногда – чисто условно, как показала, организованная наша же, практика).

Да и с расчётами в первом случае намного проще: надо лишь перевести число браков в каждой партии в проценты. Благо, что в модельном примере б все партии по 100 штук, отчего, как уже упоминалось нами, число бракованных изделий, выраженных в штуках и процентах численно совпадают. Однако в реальности проценты по исходным данным в виде штук надо вычислять. Особенно для возможного случая, когда партии, выбранные для контроля, неодинаковы по своему объёму, что очевидно. В теории множеств такие объёмы называются ещё «мощностью выборки» или МОЩНОСТЬЮ (или числом элементов множества или подмножества) - в данном случае контролируемых подмножеств в партиях с номерами № (см. табл. 1 и 2).

Далее, согласно принятой нами гипотезы о целесообразности применения биномиального закона распределения (чисто *на концептуальной основе*, безо всяких дополнительных вычислений то есть), расчёты существенно упрощаются. Посмотрите только на формулы, если вы концептуально (мысленно, феноменологически – все лишь) смогли, коли основания к тому были (а они по факту **ДЕЙСТВИТЕЛЬНО** были), вам остаётся лишь воспользоваться буквально четырьмя простыми формулами: для вычисления среднего (простого, механического среднего, кстати, в отличие от взвешенного среднего в терминах ВР) по формуле (22), с.к.о. σ (23), из которых легко получаются контрольные границы верхняя (24) и нижняя (25) путём прибавления или отнимания от среднего (опят же – самого

простого, - механического среднего) - либо $(+3\sigma)$, либо (-3σ) . Ну и график наподобие рис. 7 для наглядности построить.

Если нет выходящих за верхнюю границу числа браков в стандартной серии, то операция контроля качества успешно завершена. Если же таковые выходы за верхнюю границу всё же имеются, то их необходимо удалить из последующего рассмотрения (как, например, партии №5 и №9) и все эти четыре действия для оставшихся необходимо повторить сначала.

С гипотезой о распределении числа браков как элементов статистической совокупности по нормальному закону (очередная ваша гипотеза – предположение то есть) – немного сложнее. Надобно прежде доказать (и прежде всему – самим себе), что элементы исследуемой вами статистической совокупности (третья строка в табл. 1 или четвёртая строка в табл. 2) распределены именно по закону К.Ф. Гаусса, то есть по нормальному закону распределения. Хотя, повторим ещё раз: перед расчётами по видам табл. 1 или 2 дискретное к тому же распределение браков всё же ближе по сути своей к биномиальному закону распределения, чем к более привычному нам – непрерывному нормальному закону распределения браков как случайных величин.

Отметим также и последовательность действий исследователя при применении им нормального закона распределения изучаемой случайной величины, поскольку она предполагает больший объём вычислений.

Краткая последовательность при «подозрении» данных табл. 2 на нормальное распределение. Итак, выбор в статистической совокупности как множества с именем $Y = \{y_j\}, j = 1, N_2 = 8$ элементов, нахождение размаха анализируемой выборки $R = y^{max} - y^{min}$, вычисление десятичного логарифма lgN_2 , вычисления по формуле Стерджесса числа вариантов $nvr = n$ для построения будущего вариационного ряда (меньше $nvr = 4$ просто нельзя, иначе для входа в таблицу Пирсона требуется число степеней свободы не менее $df = nvr - 3$ (при числе степеней свободы $df = 0$ таблиц ни Стьюдента, ни Пирсона, ни подобных им других – просто в природе не существуют;

значит, число nvr вариантов вариационного ряда должно быть не менее четырёх, что весьма важно).

Расчёт числа вариантов строимого нами ВР и на его основе шага $h = R/nvr$ следует непосредственное построение собственно вариационного ряда с попутным определением фактических (эмпирических частот) f_i , $i = 1, nvr \geq 4$ (штук). Построение ВР заканчивается таблицей в виде табл. 6.

Далее необходимо строить его продолжение для оценки значений теоретических частот $f_i^{теор}$ – по числу вариантов, т.е. строк ВР (ибо вариационный ряд – это, повторим ещё раз, – всегда таблица). А теоретические частоты мы определяем в табл. 7, да ещё входя при этом для определения необходимых значений $\varphi(t)$ в соответствующую таблицу табулированных значений функции нормального распределения, предварительно вычислив константу, равную $(N2 \cdot h/\sigma_p)$ и округлив полученные значения теоретических частот до целых чисел (см. последнюю графу табл. 7).

Итак, все значения эмпирических частот (ещё в таблице ВР были определены), и ряд значений теоретических частот найдены. Остаётся лишь сравнить эмпирические частоты (полученные из нашего статистического опыта) f_i по числу вариантов когда-то нами построенного ВР и теоретические частоты в том же числе $f_i^{теор}$ – см. рабочую табл. 8, что отражает процесс вычисления расчётного значения критерия Пирсона.

Если $\chi^2_{расч} < \chi^2_{табл}$ при уровне доверия α (рекомендуется величина уровня доверия $\alpha = 0,05$), то наши сомнения относительно справедливости выдвижения нами «нулевой гипотезы» о нормальном распределении можно отвергнуть: исследуемые элементы статистической совокупности (СС) принадлежат именно такой совокупности, элементы которой распределены (и нами это было доказано) именно по нормальному закону.

Как видно, с нормальным законом всё сложнее. А чтобы быть уверенным, что перед вами дискретная картина подчинена СОДЕРЖАТЕЛЬНО биномиальному закону распределения – достаточно

лишь понимать суть его распределения, которая приведена здесь в районе начала данного раздела.

После рассмотрения двух законов распределения случайных величин – Бернулли и биномиального с последующей апробацией на тех же данных гипотезы о нормальном распределении той же случайной величины, рассмотрим теперь подробнее ещё один закон распределения, а именно, распределение Пуассона.

4. ЗАКОН РАСПРЕДЕЛЕНИЯ ПУАССОНА

Этот закон достаточно подробно и наглядно изложен, например, в [3. – С. 213 – 222] с некоторыми попутными обобщениями по другим законам распределения случайных величин.

Поэтому, в дополнение к предыдущему распределению необходимо получить возможно большее представление также об ещё одном дискретном законе распределения - ЗР Пуассона. Напомним только, что предыдущие ЗР Бернулли и биномиальный ЗР также оперируют с *дискретными* значениями случайной величины. Данное дополнение позволит не только наиболее полно охватить дискретные ЗР, применяемые исследователями при контроле качества, но и так же сравнить его со знакомым уже нам нормальным законом распределения Гаусса, который, напомним ещё раз, оперирует уже с *непрерывными* СВ.

При этом не станем забывать и того, что при известных условиях даже дискретные законы распределения иногда сходятся к закону Гаусса.

4.1. Возможные варианты творческого участия студентов-молодых специалистов

Оговоримся сразу о том, что мы с Вами постараемся избежать сложных ветвящихся схем-рисунков. И действительно, студенты как будущие

дипломированные специалисты могут столкнуться лишь с тремя видами или способами проявления *вашей творческой активности*. Итак, в начале трудовой деятельности выпускников может возникнуть три ситуации (заметьте попутно, что в теории вероятностей (ТВ) такая ситуация соответствует «полной группе событий»).

1. Процесс производства фирмой или заводом серийных изделий налажено. Важно обратить внимание, на каком ЗАКОНЕ РАСПРЕДЕЛЕНИЯ (об обобщении их и как ими пользоваться мы обязательно с вами, как мы планируем, познакомимся ещё подробнее). А набор этих законов распределения случайных величин (здесь – ожидаемое событие есть появление БРАКА) в общем немаленький: закон Бернулли, биномиальный закон, закон Пуассона, нормальный закон распределения Гаусса). Какой из них применить при уже действующих на предприятии или его участке? – Здесь наша с вами задача по предварительной оценке СОДЕРЖАНИЯ процесса и организации КОНТРОЛЯ КАЧЕСТВА продукции (услуг) сводится к следующему: насколько используемый закон распределения, отражённый, скажем, в p -контрольных картах (или pn -картах), будет для данного состояния производства наиболее УМЕСТЕН? А может, если квалифицировать появление брака по какому-либо иному закону распределения даст лучший результат? – А уж это зависит от вас, молодых специалистов.

Ну и возможные варианты: 1) выпуск какого-то изделия по объёму существенно возрос, а контрольные p -карты с прежними законами распределения случайной величины – брака то есть, - остались прежними; (а, может, их пора поменять или не надо?); 2) выпуск изделий уменьшился вследствие их усложнения: годятся ли прежние контрольные карты? 3) перешли на выпуск новых изделий: 4) смешанный вариант: часть изделий – из старого ассортимента, часть – из нового и т.д.

2. Но вот предприятие приступило только к выпуску НОВОЙ ПРОДУКЦИИ. Опять вопрос: а как контролировать её качество? Всегда ведь

существует дилемма: не пропустить бы брак (потребитель вашей продукции будет недоволен) – всегда квалифицировалась как «ошибка 1-го рода», и как бы случайно не забраковать годное изделие (прямые убытки предприятию) – «ошибка 2-го рода». Хорошо и привычно пользоваться средними (как правило – взвешенными) и дисперсиями (корень из них и есть среднее квадратическое отклонение – с.к.о.). Но, если вы вспомните (предварительно просмотрев, конечно же, их вычисления), скажем, по нормальному закону распределения Гаусса (НЗР), а сам характер производственного процесса и, соответственно, перечисленные выше величины вычисляются совсем по другим формулам, ибо закон распределения «ожидаемых событий – брака то есть, - подчиняется совсем другому закону распределения? – Так что сначала вид ЗР СВ необходимо установить.

Отметим сразу: только при проверке на НЗР алгоритмически вычисляются теоретическое распределение ожидаемых событий, а уж затем используются критерии согласия Пирсона, Романовского и Колмогорова. И только ОСТАЛЬНЫЕ законы распределения – основные из них Бернулли, биномиальный и Пуассона определяются как бы «на глаз», только *по сути* производственного процесса и появления ожидаемых случаев – бракованных изделий.

3. Ну и потом п. 1 и п.2 на разных стадиях могут носить более или менее смешанный характер, как уже упоминалось в конце п. 1.

Поэтому наша задача – научиться квалифицировать сам процесс и ожидаемые события брака как средства контроля качества для законов распределения Бернулли, биномиального и Пуассона – как бы «на глаз». Тогда как принадлежность (или непринадлежность к НЗР, а также, кстати, иногда и для ЗР СВ Пуассона, *напоминаем*, именно РАССЧИТЫВАЕТСЯ. То есть рассчитываются определённые параметры и сравниваются с табличными значениями. А по результатам таких сравнений исследователь уже приходит к умозаключению: распределены элементы СВ по НЗР (ЗР Пуассона) или нет. Однако, если ошибка в принятии гипотезу о нормальном распределении

невелика, она может быть учтена в последующих рассуждениях, расчётах и окончательных выводах отдельно.

Понятно, что рассчитать по исходным данным что-то методически, может быть, и проще, что отнюдь не значит «полезнее для дела». Ведь законы распределения случайных величин, т.е. бракованных изделий, для действительных их законов распределения (Бернулли, биномиального, да и Пуассона в отдельных случаях) надо действительно неформализовано оценивать именно по существу исследуемого процесса («на глаз»). А для этого надо хорошо представлять себе их сущность. Ну а поскольку два первых нами уже рассмотрены ранее, рассмотрим далее подробнее закон распределения СВ Пуассона, столь же часто применяемого при контроле качества выпускаемой предприятием (фирмой, заводом) продукции (услуг) для того или иного сегмента рынка.

4.2. Распределение случайной величины согласно закону Пуассона

Существует множество источников описания содержания этого закона распределения СВ. Ну, например, обратимся к уже упомянутому источнику [3. – С. 212 - 232]. Это уже будет третий дискретный закон предельного распределения (ЗР): Бернулли (или биномиальный) и Пуассона. А более подробно суть непрерывного нормального ЗР Гаусса рассмотрим несколько позже.

Заметим также о сходствах и различиях понятий «сущность» и «содержание» как философских категорий, так как студенты нередко воспринимают их как синонимы, что, конечно же, далеко не так, и вот почему. Философская категория «сущность» предполагает оценку исследователем в объекте изучения (ОИ) внутренних, необходимых и повторяющихся связей по мере углубления в свойства ОИ и объективно выражается в «явлениях» - первого рода, второго рода и так далее. Поэтому данные категории изучают как парные «сущность-явления». С другой

стороны, философская категория «содержание» предполагает оценку исследователем в ОИ внутренних, повторяющихся, необходимых, а также СЛУЧАЙНЫХ связей. Содержанию всегда соответствует адекватная ему категория «формы», поэтому и эти категории изучают как парные «содержание-форма».

Однако, наличие вот этих вот случайных связей (а необходимые это связи в ОИ или случайные, покажет только другая философская категория – «практика» (как 1) источник познания; 2) предмет познания и уже хорошо известный нам как 3) критерий истины, что всем известно). А поскольку на момент исследования часто методически трудно установить характер связей – необходимые ли они или случайные, - форма, таким образом, обладает свойством обратного диалектического влияния на содержание (понятно, что изначально содержание первично, а форма – вторична). Хорошей иллюстрацией к сказанному может послужить известная морская краткая притча: «Потому и корабли в море тонут, что у матроса в тумбочке беспорядок». Трудно с этим поспорить, ведь так?

Поэтому при оценке процесса производства продукции в виде изделий и услуг необходимо наиболее адекватно учитывать особенности процесса как непосредственного производства, так и проверки его на соответствие требуемому уровню качества. И это всё – кроме обязательного учёта в производстве и контроле требований действующих стандартов и других подзаконных актов. Поэтому именно в таких ситуациях от исследователя как субъекта познания потребуются все его знания, умения и навыки, чтобы преимущественно выявить в его процессе именно внутренние, повторяющиеся и необходимые связи, иначе говоря, именно «сущность» объекта познания. Для этого и предназначено данное учебное пособие, которое хотя и не является достаточно кратким, но может отражать лишь незначительную часть реальной практики, которая всегда богаче и разнообразнее, чем кем-то в своё время описанные её отдельные стороны. А данное обстоятельство имеет самое прямое отношение и к оценке характера

распределения исследуемых случайных величин, которые для нас являются средствами выявления возможных случаев брака.

Далее подробнее рассмотрим свойства распределения случайных величин по закону Пуассона. И, прежде всего, о том, когда же всё-таки скорее всего применяется именно этот дискретный ЗР СВ Пуассона.

Когда применяется ЗР Пуассона. Это распределение, описывает результаты экспериментов, в которых считают события, происходящие *случайно*, но в определённом *среднем темпе*. И содержание этих двух строчек необходимо принять во внимание, а лучше - запомнить.

В качестве классического примера ЗР Пуассона обычно используют работу счётчика Гейгера, считающего k штук электронов, испущенных в РАДИОАКТИВНЫХ РАСПАДАХ за одну минуту. Опустим подробности процесса (каждое ядро характеризуется определённой вероятностью распада за минуту и др.). Кстати, если бы мы знали эту вероятность (очень маленькую по своему значению) и число ядер (весьма большое их число), то могли бы весьма просто подсчитать *ожидаемое среднее число распадов в минуту*. То есть, если нам известно было бы число ядер « n » и вероятность распада каждого p , то вероятность k распадов – это просто вероятность k -успехов в n испытаниях, или (запись ЗР Пуассона) $b_{n,p}(k)$.

Однако число испытаний n (т.е. число ядер) – просто огромно ($n \sim 10^{20}$), и вероятность «успеха» (т.е. распада) для любого одного ядра – просто ничтожна ($p \sim 10^{-20}$), Однако вполне можно показать, что биномиальное распределение НЕОТЛИЧИМО от более простой функции, называемой распределением Пуассона, т.е. что

$$P(k \text{ отсчётов за любой определённый интервал}) = p_{\mu}(k). \quad (33)$$

А в формуле (33), содержащее распределение Пуассона, $p_{\mu}(k)$ определяется как:

$$p_{\mu}(k) = e^{-\mu} \frac{\mu^k}{k!}. \quad (34)$$

В формуле (34) μ – просто пока просто положительный параметр ($\mu > 0$) для k . Нам же для полной ясности остаётся лишь выяснить, что такое параметр μ . Для раскрытия его содержания придётся кое-что из математического анализа просто вспомнить или же – просто поверить следующим выкладкам. Для начала вычислим среднее число ожидаемых отсчётов k в многочисленных испытаниях:

$$k_{\text{cp}} = \sum_{k=0}^{\infty} k \cdot p_{\mu}(k) = \sum_{k=0}^{\infty} k \cdot e^{-\mu} \frac{\mu^k}{k!}. \quad (35)$$

Понятно, что первый член при $k=0$ может быть опущен (он равен нулю), а вот $(k/k!)$ можно для упрощения (35) заменить на $[1 / (k-1)!]$. Далее вынесем за знак суммы, оказавшийся общим множителем $\mu e^{-\mu}$, и получим

$$k_{\text{cp}} = \mu e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!}. \quad (36)$$

Хотя и формула (36) кажется не вполне удобной, но это только на первый взгляд. И действительно, в дроби под знаком суммы просматривается вот такой, знакомый ещё из математического анализа, так называемый «второй замечательный предел» (это когда изучали бесконечно малые и операции с ними. Вспомнили?). Так что остающаяся бесконечная сумма:

$$1 + \frac{\mu^1}{1!} + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \dots = e^{\mu}. \quad (37)$$

С учётом известного всем давно преобразования типа формулы (37), в формуле (36) эти самые e^{μ} и $e^{-\mu} = (1 / e^{\mu})$ просто сокращаются, и из выражения (36) приходим к совсем простой формуле (38):

$$k_{\text{cp}} = \mu. \quad (38)$$

Таким образом, параметр μ , характеризующий распределение Пуассона $p_{\mu}(k)$, это просто **среднее число отсчётов, которое мы ожидаем в случае многократного повторения счётного эксперимента.**

Однако, не станем останавливаться только на формуле (38), и ещё рассмотрим другие свойства изучаемого нами сейчас распределения Пуассона. Начнём с самого простого – с доступной нам всегда визуализации

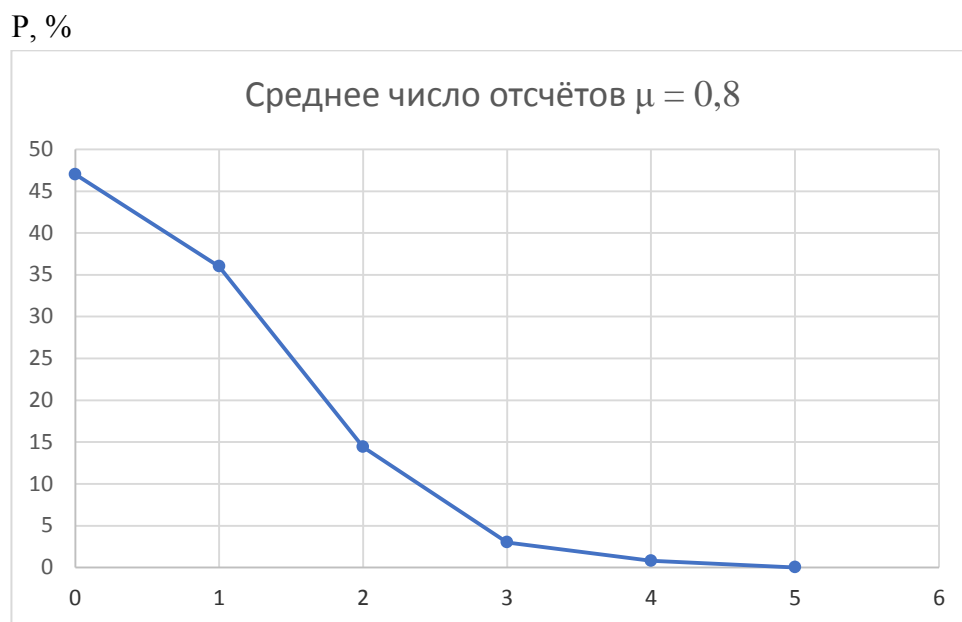
информации, представленной на графиках рис. 9 и рис. 10, так как всегда было известно, что до 90% информации усваивается именно через зрительный канал, то есть увиденное, о чём и свидетельствует известная поговорка: «Лучше один раз увидеть, чем сто раз услышать». А чтобы обещанные нами рисунки стали бы ещё более наглядными, уточним, что на рис. 9 отобразим картину распределения вероятностей для случая, когда $\mu = 0,8$, а $n = 4$, а на рис. 10 – ту же картину для $\mu = 3$ и $n = 9$. Для расчёта конкретных значений вероятностей для каждого случая благоприятного исхода k воспользуемся формулой (34). Тогда вероятности $P(k)$, в %, событий для разного числа k можно представить на рис. 9 и 10.

Тогда по содержанию рис. 9 из [3. – С. 215 – 216] видно, что наиболее вероятные числа отсчётов равны для $k = 0$, $k = 1$ (44,9% и 36,0% - соответственно), тогда как для $k = 2$ и $k = 3$ вероятность их появления также является достаточно видимой (22,4% и 22,4%), что на указанных рисунках является вполне наглядным. То есть с ростом величины μ максимальные значения смещаются вправо по оси абсцисс для k . Впрочем, на обоих рисунках 9 и 10 приведённые на них распределения являются явно асимметричными.

Для варианта, когда $\mu = 0,8$, а $n = 4$, вероятности (%): $P(0) = 44,9$; $P(1) = 36,0$; $P(2) = 14,4$; $P(3) = 3,8$; $P(4) = 0,8$; $P(5) = 0,01$ (см. рис. 9).

Кстати, обратим внимание при сравнении графиков на рис. 9 и 10 на тот предмет, что они при изменении величин μ не только меняют свою форму, но и наивероятнейшие их значения также смещаются.

Для варианта, когда $\mu = 3$ и $n = 9$, вероятности (%): $P(0) = 5,0$; $P(1) = 14,9$; $P(2) = 22,4$; $P(3) = 22,4$; $P(4) = 16,8$; $P(5) = 10,1$; $P(6) = 5,0$; $P(7) = 2,2$; $P(8) = 0,8$; $P(9) = 0,3$ (см. табл. 10).



число k

Рис. 9. Распределение Пуассона со средним числом отсчётов $\mu = 0,8$.

Здесь также будет уместным привести ещё один вариант исходных данных как продолжения сравнения рис. 9 и 10 на простой визуальной основе: при увеличении «среднего числа отсчётов μ , при котором мы ожидаем в случае многократного повторения счётного эксперимента» [3. –С. 215] график на рис. 10 напоминает асимметричного распределения Гаусса.

Действительно, при увеличении числа μ до ∞ распределение Пуассона [там же] «становится всё более симметричным и стремится к распределению Гаусса с тем же средним и стандартным отклонением». Для иллюстрации сказанного приведём ещё один график того же ЗР СВ Пуассона, скажем, для $\mu = 9$. Хотя информация на рисунках по своему характеру является дискретной, однако по своей форме напоминает огибающую кривую НЗР Гаусса, в чём и убедимся на содержании рис. 11.

P, %



число k

Рис. 10. Распределение Пуассона со средним числом отсчетов $\mu = 3$.

Из рис. 11 также видно, что огибающая – та же кривая К.Ф. Гаусса (НЗР), с тем же центром (средним значением), если проверять по соответствующим таблицам ЗР Гаусса, и с тем же стандартным отклонением (закон трёх сигм) [10], [11].

Даже на рис. 11 при $\mu = 9$ наблюдается некоторая асимметрия, в случае больших (точнее, «немаленьких» значений) μ ЗР Пуассона весьма близок к нормальному ЗР Гаусса. Отметим для себя также что число $\mu = 9$ является несколько особенным (см. ниже пример Э. Деминга из [15]), что нам весьма понадобится далее при проверке вычислительных суждений отдельных авторов (или отдельного автора – как получится). Поскольку для авторов данного пособия является главным не только популяризация обязательной проверки так называемых «нулевых гипотез» при оценке того или иного закона распределения случайной величины (в нашем случае – доли браков в производстве), но и правомерность некоторого отступления от общих правил в пользу *оперативности* проводимого контроля качества.

P, %



число k

Рис. 11. Распределение Пуассона со средним числом $\mu = 9$

Кстати, ещё одно интересное свойство ЗР Пуассона обнаруживается, если вычислить его стандартное отклонение σ_k (или с.к.о. для НЗР)

$$\sigma_k^2 = (k - k_{cp})^2. \quad (39)$$

Преобразования формулы (39) заключаются в том, что среднее квадратов отклонений преобразуется по правилам статистических вычислений в их среднее от формулы (39) и, используя уже знакомое выражение (38), заменяют σ_k^2 из формулы (39) на среднюю разность квадрата k и квадрат среднего с последующим получением формулы для с.к.о.

$$\sigma_k = \sqrt{\mu}. \quad (40)$$

То есть распределение Пуассона со средним числом отсчетов $\mu = k_{cp}$ согласно (38) имеет стандартное отклонение (40) $\sigma_k = (\mu)^{1/2}$.

Сопутствующие комментарии. Мы привыкли к тому, что раньше строго соблюдалась размерность. Например, при $x_{cp} = 8,50$ штук, дисперсия $\sigma_x^2 = 2,25$ (штук²). И с.к.о. $\sigma_x = (\sigma_x^2)^{1/2} = [(2,25 \text{ штук}^2)]^{1/2} = 1,5$ штуки. Тогда можно записать $x_{cp} \pm \sigma_x$ или $8,5 \text{ шт.} \pm 1,5 \text{ шт.}$, то есть $x_{cp} =$ от 7 до 10 или $7 \div 10$, или в терминах закрытого интервала [7; 10], и всё в тех же единицах измерений – в «штуках». А здесь, в выражении (8), если $\mu = 8,5$ штук, то

(ОПЯТЬ ЖЕ ПО РАЗМЕРНОСТИ ТОЛЬКО) $(\mu)^{1/2} = (\text{штук})^{1/2}$. Корень из «штук»? – Ответ: хоть и размерность нарушается, но формулы (38) и (40) являются вполне математически корректными.

Вот таков по своей природе ЗР случайных величин Пуассона (см. рис. 9 – 11). И ещё раз подчеркнём близость законов распределения по своей природе как Пуассона, так и Гаусса (важно только, чтобы согласно формуле (38) $k_{cp} = \mu$ было бы не слишком маленьким (см. рис. 9). Тем более, что у нас по контролю качества – примерно та же картина: берём наугад образец выпускаемой продукции – и подвергаем его контролю качества. Ожидаемое событие «брак», наверное, при налаженном деле (производственном процессе) уж явно не будет ни дробным, что очевидно, ни ощутимо большим – порядка единиц из партии – вряд ли больше. А когда $k_{cp} = \mu$ будет достаточно ощутимым, *можно вследствие простоты ЗР Пуассона пользоваться именно им, а не ЗР Гаусса.*

То есть в оперативно-техническом отношении нам ЗР Пуассона вполне подходит, чем НЗР Гаусса. Тогда в качестве вполне уместного здесь обобщения необходимо сравнить эти оба ЗРСВ.

4.3. Процесс сравнения законов распределения биномиального закона, законов Пуассона и Гаусса

Во-первых, распределение Гаусса $f_{x,\sigma}(x)$ является *непрерывным*, поскольку случайная величина x - непрерывная переменная, а распределение Пуассона $P_\mu(k)$ той же величины – статистическая совокупности случайных элементов с именем X – *дискретно* по природе своей, да и подобно биномиальному ЗР, поскольку значения $k = 0, 1, 2, 3, \dots$ представляют собой именно дискретный ряд. Во-вторых, гауссово распределение $f_{x,\sigma}(x)$ определяется *двумя* параметрами: средним значением x и шириной отклонения (с.к.о.) σ_x . Тогда как распределение Пуассона чуть ли не

автоматически определяется *одним* параметром - средним μ , тогда как с.к.о. $\sigma = (\mu)^{1/2}$.

Однако, если среднее отсчётов μ велико, то дискретная природа k становится МЕНЕЕ СУЩЕСТВЕННОЙ, и распределение Пуассона (подобно биномиальному) хорошо аппроксимируется функцией Гаусса С ТЕМИ ЖЕ средним и шириной распределения, то есть

$$p_{\mu}(k) \approx f_{x,\sigma}(x), \text{ когда } \mu \text{ достаточно велико, где} \quad (41)$$

$$x_{\text{ср}} = \mu; \sigma = \sqrt{\mu}.$$

Приближение (41) называется «гауссовой аппроксимацией» распределения Пуассона. Оно аналогично соответствующей аппроксимации биномиального распределения, рассмотренного нами ранее, особенно когда параметр распределения μ достаточно велик (ну, например, когда $\mu \geq 9$ и более – см. рис. 10 и 11).

Пример 8. Рассчитать распределение Пуассона для $\mu = 64$. Вероятность, например, $k = 72$ отсчётов при исследовании радиации за некоторое время:

$p(72 \text{ отсчёта}) = p_{64}(72)$ – далее по знакомой формуле (34):

$$p_{64}(72) = e^{-64} \frac{64^{72}}{72!} = \text{трудоёмкие вычисления} = 2,9\%.$$

С другой стороны, по Гауссу σ – одна и та же:

$$\sigma = \sqrt{64} = 8.$$

Тогда вероятность $P(72 \text{ отсчёта}) \approx f_{x,\sigma}(x) \approx P_{\mu}(k) = f_{64,8}(72) =$ далее - простые вычисления = 3,0%, что считается весьма хорошим приближением, так как более точное значение $2,9\% \approx 3,0\%$.

Ещё примеры распределения по Пуассону. Классическими примерами по содержанию случайных процессов следует признать следующие:

- 1) откладывание домашними птицами на птицеферме за час число яиц;
- 2) число рождений в день в родильном доме.

Понятно, что в обоих случаях следует учитывать возможные сезонные колебания. Для дальнейших расчётов надо считать события каждый день, то есть много раз. Затем представить всё на рисунке, и проверить на ЗР Пуассона. Если желательно измерить (или применить количественный тест), то можно применить и критерий χ^2 , например.

Так примерно обстоит и в контроле качества производимых изделий: брак встречается редко, но при отлаженном производстве (без сбоев) **ДЕЙСТВИТЕЛЬНО** либо из каждой контрольной партии (например, в 100 единиц продукции, как в табл. 1 или 2), встречаются буквально лишь **НЕСКОЛЬКО** бракованных изделий из каждой-то контролируемой партии. Под нормальный закон распределения числа браков как случайных величин мы уже, помнится, не однажды пробовали. А давайте попробуем то же самое представить и с позиции законом Пуассона?

Исходная информация весьма проста [5. – С. 44 – 45]: 1) ВР представляет собой распределение по дискретному признаку, где по МЕРЕ увеличения значений признака наблюдаемого нами x частоты **ВДРУГ** резко уменьшаются и где средняя арифметическая ряда (взвешенная, конечно же, которое ещё надобно как-то подсчитать) **РАВНА** или **БЛИЗКА** по значению к *дисперсии*. Иначе говоря, согласно последней формуле, например, формуле под номером (41):

$$x_{cp} = \mu; \sigma = \sqrt{\mu}.$$

То есть $x_{cp} = \sigma^2$. Значит, такой ряд фиксируемых наблюдений можно **ВЫРОВНЯТЬ** (т.е. найти адекватную теоретическую зависимость) по кривой Пуассона, аналитическое выражение которой нами было ещё получено в формуле (34):

$$p_{\mu}(k) = e^{-\mu} \frac{\mu^k}{k!} = p_x = e^{-\mu} \frac{\mu^x}{x!},$$

где просматривается та же запись, в которой p_x – вероятность наступления отдельного события в виде x ; $x_{cp} = k_{cp} = \mu$.

Для иллюстрации приведённого здесь возьмём тот же пример 6 (табл. 2) под именем табл. 9, в которой партии (№5 и №9) с 14 и 15 браками выводятся из рассмотрения. Поэтому далее (см. табл. 9) новые номера партий обозначены через знак «/».

Как мы помним, из примера 6. Мы с вами ещё ранее установили, что раз мы удалили из той табл. 2 (или этой табл. 1) два столбца со старыми номерами 5 и 9 (по 14 и 15 браков – соответственно, потому что верхнюю границу существенно превосходили, как мы помним), отчего впоследствии изменились и средние значения браков по всем партиям (механические средние).

Таблица 9

Учёт числа бракованных изделий после выведения 5 и 9 наблюдений

№ партии / Новые №	1/1	2/2	3/3	4/4	6/5	7/6	8/7	10/8	Сум-ма
Объём партии, n_i , шт.	100	100	100	100	100	100	100	100	1000
Браков. изд., pr_i , шт.	1	2	0	3	4	1	2	3	45/15
Доля события A, p_i , %	1	2	0	3	4	1	2	3	-

Итак, сравним результаты расчётов по биномиальному закону распределения (рис. 5) и нормальному закону распределения (рис. 8):

Уровни на рисунках-графиках	Из рис. 5	Из рис. 7
Верхняя граница UCLp	6,2%	4,8%
Среднее CLp	2,0%	1,6%
Нижняя граница LCLp	- 1,7% → 0,0%	- 2,2% → 0,0%

Таким образом, из сравнения двух законов распределения рассматриваемой нами случайной величины в данном случае следует, что закон распределения Гаусса несколько «строже» биномиального закона распределения. А ведь при этом может случиться так, что возрастёт опасность 2-го рода: можно забраковать годное изделие. Попробуем всё же расширить данный пример с учётом, если предварительно примем гипотезу о

том, что исследуемая случайная величина (число браков) согласно тем же исходным данным (см. табл. 2) распределена по закону Пуассона. Для этого лишь надо воспользоваться формулами для вычисления среднего значения

$$x_{\text{ср}} = \mu \text{ и к. о. } \sigma = \sqrt{\mu},$$

после чего рассчитать верхнюю и нижнюю контрольные границы по тем же формулам (24) и (25) – соответственно.

Если мы не намерены ориентироваться на нормальный закон распределения браков, необходимость в построении вариационного ряда отпадает. Тогда в качестве μ можно взять как рассчитанную по механическому среднему (как после принятой гипотезы о биномиальном законе распределения – см. рис. 5). Тогда по ЗР Пуассона $\mu = 2,0\%$: $\sigma = \sqrt{2,0} \approx 1,4$. По формулам (24) и (25) получим: $ULCp = 2,0 + 3 \cdot 1,4 = 6,2\%$; $LCLp = 2,0 - 3 \cdot 1,4 = - 2,2\% \rightarrow 0\%$. Откуда видно, что верхняя и нижняя контрольные границы вполне совпадают, если изначально нами была принята гипотеза о биномиальном законе распределения (см. табл. 5).

С другой стороны, если ориентироваться для сравнения на гипотезу о нормальном распределении браков, то по закону Пуассона $\mu = 1,6\%$: $\sigma = \sqrt{1,6} \approx 1,26\%$. По формулам (24) и (25) получим: $ULCp = 1,6 + 3 \cdot 1,26 = 5,4\%$; $LCLp = 1,6 - 3 \cdot 1,26 = - 2,2\% \rightarrow 0\%$.

Таким образом, при сравнении использования предварительной гипотезы о распределении Пуассона по отношению к уже скорректированному ряду из восьми случаев брака практически совпадает контрольными границами для рис. 5 и немного «демократичнее» по отношению к рис. 8 (при гипотезе о нормальном законе распределения числа бракованных изделий из серии по 100 штук деталей продукции): при этом верхняя контрольная граница 4,8% возрастает до 5,4%. То есть она, эта верхняя контрольная граница $ULCp$, занимает промежуточное место между результатами применения нами гипотезы о биномиальном распределении исследуемой случайной величины (случаев брака) и нашей же гипотезой о

нормальном распределении забракованных изделий. Что и следовало ожидать в конечном итоге. И что ещё раз к тому же свидетельствует о том, что на практике действительно непросто справиться с решением извечной дилеммы: какому же конкретному закону распределения следует изучаемая нами случайная величина?

Да, формально отличия «имеют место быть», но, как мы убедились, они не носят уж очень заметный характер. Следовательно, выбор и применение на практике нами той или иной гипотезы распределения случайной величины (здесь – забракованных из партии число изделий) в самом начале исследования может быть всегда нами же и уточнён путём проведения проверочных (верифицирующих) действий при осмыслении (формализации) результатов практического контроля. Что является весьма важным в производственной деятельности *наших будущих выпускников*.

Итак, распределение браков как СВ мы рассматривали через призму биномиального ЗР (рис. 5) и нормального закона распределения (рис. 8) при тех же исходных данных, откуда и выражение для среднего значения браков $p_{cp}, \%$ и стандартного отклонения σ_p вычислялись при биномиальном ЗР, в котором $N = 10, n_i = n = 100 = const$:

$$p_{cp} = \frac{\sum_{i=1}^N np_i}{\sum_{i=1}^N n_i} \cdot 100\% = \frac{45 - 14 - 15}{1000 - 2 \cdot 100} \cdot 100\% = 2,0\%.$$

$$\sigma_p = \sqrt{\frac{p_{cp}(100 - p_{cp})}{n}} = \sqrt{\frac{2(100 - 2)}{100}} = 1,40 = 1,4\%.$$

$$UCLp = p_{cp} + 3 \cdot \sigma_p = 2,0 + 3 \cdot 1,4 = 6,2\%.$$

$$LCLp = p_{cp} - 3 \cdot \sigma_p = 2,0 - 3 \cdot 1,4 = -2,2\% \rightarrow 0\%.$$

Теперь же обратимся, как картина, представленная своими, исходными для нас данными, в табл. 2 теперь уже соответствует ЗР Пуассона [3].

Итак, мы располагаем исходными данными табл. 2, где встречающееся число ожидаемых событий - «бракованное изделие». Измеряется оно, понятно, в целых единицах k , хотя и среднее $\mu = k_{cp}$ может быть представлено

и вещественным числом, что также понятно. К тому же показанное несколько выше $p_{cp} = k_{cp} = \mu = 2,0\%$, если в терминах ЗР Пуассона. Ну, давайте проверим и его. Для этого необходимо и достаточно лишь воспользоваться формулой (34), задавая для неё различные значения случаев брака k (см. табл. 1). Так переменная k изменяется от 0 до 4 (т.е. $k=0, 1, 2, 3, 4$, возьмём ещё и 5, чтобы «с запасом»). Итак, при $\mu = p_{cp} = k_{cp} = 2\%$. Теперь проведём расчёты по формуле (34) вероятности появления события k :

$$\text{при } k = 0 \quad p_2(0) = e^{-2} \frac{\mu^0}{0!} = e^{-2} \frac{1}{1} = 0,140 \text{ или } 14\%;$$

$$\text{при } k = 1 \quad p_2(1) = e^{-2} \frac{\mu^1}{1!} = e^{-2} \frac{2}{1} = 0,270 \text{ или } 27\%;$$

$$\text{при } k = 2 \quad p_2(2) = e^{-2} \frac{\mu^2}{2!} = e^{-2} \frac{4}{2} = 0,270 \text{ или } 27\%;$$

$$\text{при } k = 3 \quad p_2(3) = e^{-2} \frac{\mu^3}{3!} = e^{-2} \frac{8}{6} = 0,180 \text{ или } 18\%;$$

$$\text{при } k = 4 \quad p_2(4) = e^{-2} \frac{\mu^4}{4!} = e^{-2} \frac{16}{24} = 0,090 \text{ или } 9\%;$$

$$\text{при } k = 5 \quad p_2(5) = e^{-2} \frac{\mu^5}{5!} = e^{-2} \frac{32}{120} = 0,040 \text{ или } 4\%.$$

Прежде чем строить график, вспомним, что в ЗР Пуассона

$$\sigma_p = \sqrt{\mu} = \sqrt{2} = 1,4\%,$$

как у нас прежде и было при биномиальном ЗР (см. чуть выше). Всё пока сходится. А теперь график по полученным результатам начертим на рис. 12:

И что же мы видим на графике рис. 12? – Да, прежде всего то, что при $k=1$ и $k=2$ вероятности одни и те же, и равны они 27% (или 0,270 в относительных единицах).

Ну а теперь – как и было изложено выше, когда речь шла только о ЗНР, - сделаем то же самое. Итак, практический (эмпирический) вид кривой Пуассона мы с вами имеем. Ну и не худо бы при этом ещё и эмпирические частоты f_i , $i = 1, N = 8$ знать. Известно лишь одно средство – всё тот же вариационный ряд (ВР): 1) оценить размах выборки R ; 2) рассчитать

примерное число интервалов n (или nvr) ВР по той же формуле Стерджесса [6];3) найти шаг $h = R/n$ изменения исследуемой величины – здесь X – СВ в виде статистической совокупности, элементы которой – отдельные случаи брака;4) построить собственно ВР – таблицу со строками-вариантами; 5) сосчитать эмпирические частоты f_i в каждой строке таблицы (то есть для каждого варианта вариационного ряда).



Рис. 12. Графическое изображение поведённых расчётов при $\mu = 2\%$.

Затем опять наступит этап применения критериев согласия Пирсона, Романовского, а может быть, и Колмогорова. Ну а уж потом – последовательное применение критериев согласия, – реализация уже известных этапов. А пока, по-прежнему, опять начнём строго по порядку - с построения ВР.

Построение очередного вариационного ряда. Мы уж тут не напоминаем, для чего именно этот ВР для получения эмпирических частот f_i мы строим: для нормального ли ЗР или ЗР Пуассона ли. Важно начать именно с построения ВР. Размах выборки (см. последнюю строку табл. 1) составляет от 0% до 4% случаев брака, то есть $R = 4 - 0 = 4\%$. Следующий этап – определение примерного числа интервалов по формуле Стерджесса:
 $n = nvr \approx 1 + 3,322 \cdot \lg N = 1 + 3,322 \cdot \lg 8 = 1 + 3,322 \cdot 0,9 = 1 + 2,99 = 3,99 \approx 4$.

Полученное число вариантов ВР по формуле Стерджесса $nvr = 4$ (то есть *не менее 4-х* интервалов: далее примем $nvr = 5$) нас вполне устраивает: для входа в таблицу Пирсона для НЗР необходимы число степеней свободы $df = n - 3 = 4 - 3 = 1 > 0$, для ЗР Пуассона $df = n - 2 = 2$.

И действительно: число бракованных изделий от партии к партии измеряется от 0 до 4-х. Тогда мы строим не *интервальный*, но *дискретный* ВР (см. табл. 10). А в роли вариантов тогда и будут фиксированные значения k . Тогда пусть $n = nvr = 5$ и шаг h для интервального ВР $h = 4\%/5 = 0,8\%$, то есть $0,8\%$ приходится на один интервал будущего дискретного ВР с $h = const$. Но и число свободы для будущего входа в критерий Пирсона будет уже $df = n - 3 = 5 - 3 = 2$, для которого $\chi^2_{табл} = 5,99$ для уровня значимости $\alpha = 0,05$.

Определение теоретических частот $f_i^{теор}$ значительно проще, чем определения их же для проверки распределения на ЗР Гаусса (нормальный закон распределения). Для определения вида закона выравнивания, согласно [5. – С. 45], для которого $f_i^{теор}$ вычисляется по формуле:

$$f_i^{теор} = N \cdot p_x, \quad (42)$$

где N , в отличие от НЗР, представляет собой *все* исходные значения исследуемой СВ, тогда как при проверке гипотезы о нормальном их распределении тот же символ в подобном же случае означает лишь *малое* число вариантов построенного нами же ВР по исходной СС с именем $Y = \{1, 2, 0, 3, 4, 1, 2, 3\}$; $|Y| = 8$ после удаления партий с номерами 5 и 9.

Таблица 10

ВР для получения эмпирических частот

№ варианта i	$x_i^{нач} \div x_i^{кон}$ интервала; (+)	Эмпирические частоты f_i	
		счёт	всего штук
1	2	3	4
1	0,0 – 0,8	/	1
2	0,8 - 1,6	//	2
3	1,6 - 2,4	//	2
4	2,4 - 3,2	//	2
5 = n	3,2 - 4,0	/	1
Всего:	-	8	8

Именно в этом как сходство этих символов N , так и их различие при проверке гипотез о ЗР СВ Пуассона и ЗР СВ Гаусса, о чём необходимо при соответствующих расчётах теоретических частот не забывать - как при проверке гипотезы о распределении СВ по закону Пуассона, так и при проверке гипотезы о распределении СВ по закону Гаусса.

Проверка по критерию согласия Пирсона о ЗР Пуассона. Найдём их по такой простой формуле (42) и, как и в предыдущем случае НОРМАЛЬНОГО ЗАКОНА. При этом, как и прежде, полученные результаты округлим до целых: $f_1^{теор} = 8 \cdot 0,14 = 1,12 \approx 1$; $f_2^{теор} = 8 \cdot 0,27 = 2,16 \approx 2$; $f_3^{теор} = 8 \cdot 0,27 = 2,16 \approx 2$; $f_4^{теор} = 8 \cdot 0,18 = 1,44 \approx 1$; $f_5^{теор} = 8 \cdot 0,09 = 0,72 \approx 1$, после чего внесём в столбец 4 в табл. 11 – сразу для как бы «попутной» проверки критерия согласия χ^2 Пирсона, уже знакомой нам из предыдущего материала для НЗР.

Тогда $\chi^2_{расч} = 1,00 < \chi^2_{табл}(df = 2; \alpha = 0,05) = 5,99$, что, в свою очередь, означает, что расхождения между эмпирическими и теоретическими частотами можно считать случайными, и выдвинутая нами (хоть и негласно) гипотеза о близости эмпирического распределения к распределению Пуассона *не опровергается*. Приводимый здесь ход приёма проверки работоспособности предварительно выдвинутой «нулевой гипотезы» о ЗР СВ является универсальным.

Таблица 11

Расчёт значения критерия Пирсона $\chi^2_{расч}$

Варианты, I	Эмпирические частоты f_i	Теоретические частоты $f_i^{теор}$	$(f_i - f_i^{теор})$	$(f_i - f_i^{теор})^2$	$(f_i - f_i^{теор})^2 / f_i^{теор}$
1	2	3	4	5	6
1	1	1	0	0	0
2	2	2	0	0	0
3	2	2	0	0	0
4	2	1	1	1	1
n = 5	1	1	0	0	0
Всего:	8 шт.	7 шт. (?)	-	-	$\Sigma = \chi^2_{расч} = 1.$

То, что число теоретических частот не равно числу эмпирических, при подобных вычислениях, особенно для сверхмалых выборок ($N < 25 - 30$), нередко при вычислениях всё же имеет место (здесь $8 \neq 7$). Обычно в таких случаях *уменьшают* число интервалов ВР.

Итак, $\chi^2_{расч} = 1 < 5,99$, о чём и свидетельствует содержание табл. 11.

Проверка по критерию согласия Романовского о ЗР Пуассона. Как альтернатива предыдущей проверке применим к нашей первой проверке по критерию согласия Пирсона ещё один – критерий Романовского. Итак, из [5. – С. 48 – 49] по формуле (43), в которой, в отличие от (32), число степеней свободы выражено через величину k (встречается и такое их обозначение):

$$Rom_{расч} = \frac{|\chi^2 - k|}{\sqrt{2k}} = \frac{|1,00 - 1|}{\sqrt{2 \cdot 1}} = 0,71 < Rom_{гран} = 3. \quad (43)$$

Гипотеза о ЗР Пуассона статистической совокупности, отражающей процент бракованных изделий (т.е. появление события A) в исследуемых партиях по 100 штук деталей (готовой продукции) $p_i, \%$ (см. табл. 1, 2) и после применения критерия согласия Романовского также подтверждается.

Критерий Колмогорова исследует так называемые «накопленные частоты (или частоты)», что и приведено в табл. 12 (здесь общее число исследуемых единиц СВ $N = 8$). Тогда по критерию согласия Колмогорова:

$$\lambda = \frac{D}{\sqrt{N}} = \frac{1}{\sqrt{8}} = \frac{1}{2,828} = 0,35.$$

По таблице $P(\lambda)$ [5. – С. 133, Приложение 6] получим для $\lambda = 0,35$ вероятность $P = 0,9997$ (или 99,97%). Следовательно, и по данному критерию согласия можно полагать, что расхождения между эмпирическими частотами f_i и теоретическими частотами $f_i^{теор}$ носят СЛУЧАЙНЫЙ ХАРАКТЕР, и выявленными расхождениями можно далее пренебречь также и в рамках верификации гипотезы о критерии согласия Колмогорова. Хотя следует отметить, что данный критерий является, по мнению ряда авторов, «менее мощным» критерия согласия Пирсона, а применение критерия Колмогорова рекомендуется при выборочной совокупности в 50 единиц и более.

Таблица 12

Проверка критерия согласия Колмогорова на ЗР Пуассона

Эмпирические частоты f_i , шт.	Теоретические частоты $f_i^{\text{теор}}$, шт.	Накопленные частоты		$ q_i - q_i^{\text{теор}} $
		Эмпирич. q_i	Теоретич. $q_i^{\text{теор}}$	
1	1	1	1	0
2	2	3	3	0
2	2	5	5	0
2	1	7	6	$\max 1 = D$
1	1	8	7	1

На этом верификация гипотезы о ЗР СВ Пуассона завершается.

Сравнение наших исследований по ЗР нормальному и ЗР Пуассона.

Исходные данные для обоих ЗР мы использовали одни и те же: табл. 1 и табл. 2. И у нас в результате расчётов получилось следующее.

Если сравнить предыдущие результаты, то придётся признать, что все они (критерии согласия Пирсона, Романовского) имеют примерно одинаковую долю достоверности. Правда, не всегда была осуществлена проверка на критерий согласия Колмогорова. Но если её провести, то результат будет также успешным, что ещё раз свидетельствует о том, что при приближении исследуемых значений до известного числа (до 9-ти, например), что НЗР Гаусса и ЗР СВ Пуассона также, как и закон Бернулли, биномиальный (особенно последний) так или иначе сходятся к параметрам нормального закона распределения при увеличении числа испытаний, в чём нетрудно убедиться по результатам расчётов по соответствующим критериям согласия.

Ещё раз отметим только, что НЗ распределения Гаусса оперирует с непрерывными величинами, а остальные, рассмотренные нами законы, базируются на именно дискретных исходных данных. И то, что закон биномиального распределения как бы «похож» на закон Пуассона – ничего удивительного нет (при определённых условиях, конечно же). Как и нормальный закон распределения Гаусса также «смыкается» с показателями Пуассона, когда численность исследуемых «ожидаемых» случаев, например,

приближается, скажем к 9-ти или, ещё больше – 15-ти ожидаемым случаям (данном случае – случаев «брака» из испытываемой партии готовых изделий).

Тогда закон распределения Пуассона вообще непросто отличить от закона нормального распределения Гаусса: средние ожидаемые величины у них совпадают, и вычисленные средние отклонения у них также одинаковы, если число ожидаемых случаев удачных наблюдений (здесь – случаи брака) где-то порядка 9 – 15-ти.

Поэтому проверка на тот или иной закон распределения предполагает прохождение соответствующего теста на его соответствие, сопровождаемого как правило с дополнительными вычислениями.

4.4. Общие рекомендации применению законов распределения случайных величин

Ранее мы рассмотрели четыре возможных закона распределения элементов случайных величин: Бернулли, биномиальный, Пуассона и нормальный Гаусса. Первые три – ЗР распределения дискретных СВ, закон Гаусса – для непрерывных СВ. Ну и наиболее характерные признаки, когда надобно применять тот или иной ЗР, что и показано на конкретных примерах в источниках информации [2] и [3].

То есть уже по самому характеру примеров можно вполне принять контроля исследуемой случайной величины (здесь - брака). То есть ЗАПОМИНАНИЕ самого содержания приводимых здесь примеров уже может дать ориентацию их пригодности, либо непригодности при организации схемы (алгоритма) планируемого контроля качества выпуска продукции (товаров, услуг). Причём как бы «вплотную» это относится к первым трём схемам, тогда как четвёртую (нормальный ЗР) обычно ПРЕДВАРИТЕЛЬНО доказывают на формулах. Тогда как решение применять (или не применять) первые три закона распределения решается

именно на интуитивном уровне при условии ЗНАНИЯ сущности исследуемого процесса контроля.

Примеры применения законов распределения случайных величин.

Далее будут просто приводиться примеры. После из подробного восприятия, читатели смогут в какой-то мере соотнести их с конкретным исследуемым производственным процессом.

Примечание: каждый приводимый здесь пример мы попытаемся приспособить непосредственно под контроль качества продукции. В роли «продукции» здесь может выступать как собственно изделие на производстве, так и показатели учебного процесса, что доступно любому студенту (здесь табл. 2 приводится как табл.13; табл. 4 как табл. 14). Здесь табл. 13 отражает итоги измерений в партиях готовых изделий:

Таблица 13

Результаты контроля качества в 10-ти партиях по 100 изделий

№ партии	1	2	3	4	5	6	7	8	9	10	Сум-ма
Объём партии, n_i , шт.	100	100	100	100	100	100	100	100	100	100	1000
Браков. изд., pr_i , шт.	1	2	0	3	14	4	1	2	15	3	45
Доля события A, p_i , %	1	2	0	3	14	4	1	2	15	3	-

Напротив, табл. 14 отражает итоги аттестации студентов кафедр и ориентирована как на собственно оценку результатов аттестации, так и на работу кафедр и преподавателей, иначе говоря, на возможности управления учебно-воспитательным процессом.

Таблица 14

Оценка по кафедрам вуза доли аттестованных студентов, %

№ кафедры вуза	1	2	3	4	5	6	7	8	9	10	Сум-ма
Число аттестуемых, n_i	42	27	14	26	38	34	45	47	10	30	313
Аттестов-ны успешно, pr_i	15	9	2	10	11	13	11	15	4	14	104
Доля аттестов-ых, p_i , %	35,7	33,3	14,3	38,5	28,0	38,2	24,4	31,9	40,0	46,7	-

Предметные области последние две таблицы отражают разные. Однако методика работы с табл. 13 и табл. 14 является одинаковой.

Пример 9 применения ЗР Бернулли [2. - С. 49 – 50].

Волокна хлопка определённого сорта в среднем на 75% имеют длину, меньшую 45 мм, и на 25% - длину большую или равную 45 мм. Найти вероятность того, что среди 3-х НАУДАЧУ взятых волокон 2 будут короче, а одно – длиннее 45 мм. (Ответ: вероятность события C как $P(C) = 27/64 \approx 0,42$ или 42%.

Тогда, в терминах уже знакомого примера 9 для табл. 5: если найденное нами среднее значение вероятности $p_{cp} = 4,5\%$, найти вероятность того, что среди 3-х наудачу взятых изделий 2 будут меньше p_{cp} , а одно – превосходить его. Для табл. 6: среднее значение аттестованных $p_{cp} = 32,2\%$. Найти вероятность того, что среди 3-х наудачу взятых кафедр на 2-х из них среднее число аттестованных будет меньше 32,3%, а одна покажет больший результат.

Пример 10 применения ЗР Бернулли. Известно, что в среднем на 515 мальчиков рождается 485 девочек (известен в мире демографии как коэффициент 1,05 – 1,06. В одной семье ШЕСТЕРО детей. Найти вероятность того, что среди них не больше двух девочек,

(Ответ: $P \approx 0,37$ или 37%).

Пример 11 применения биномиального ЗР [3. – С. 199]. Биномиальное распределение $b_{n,p}(k)$ определяет вероятность реализации k успехов в n испытаниях в случаях, когда p есть вероятность успеха в ЕДИНСТВЕННОМ испытании. Оно всегда применяется, например, в случае бросания n костей, монет. Успех испытания p , а $q = (1 - p)$ обозначим как неуспех. Так, имея k успехов а n испытаниях, вероятность p находится с помощью *биномиального распределения*.

И хотя это распределение не имеет большого значения для, скажем, элементарной физики, однако из-за простоты является привлекательным (да

и прямо ведёт к распределению Гаусса). Примеры? – см. непосредственно работу с табл. 5 и табл. 6, что изложены в предлагаемом материале.

Также вспомним и то, что биномиальный ЗР [3], [7] $b_{n,p}(k)$ предполагает исследование эксперимента с *дискретными* значениями ожидаемых исходов $k=0, 1, 2, \dots, n$, тогда как НЗР Гаусса $f_{x,\sigma}(x)$ описывает *непрерывные* значения. Впрочем, эти два ЗР мы с вами уже взаимно рассматривали.

Пример 12 распределения Пуассона [3. – С. 213]. Имеем дело с образцом радиоактивного материала плюс счётчик Гейгера: каждую минуту считаем k электронов. Повторяем эксперимент много раз и получаем разное число k , и имеющаяся *вариация* отражает, скорее всего, СВОЙСТВА процесса радиоактивного распада. Вот так и у нас – см. наши примеры в табл. 5 и 6: условия применения ЗР Пуассона – весьма похоже. Мы ведь тоже не знаем, сколько бракованных изделий окажется в очередной партии из, например, 100 деталей (или число аттестованных студентов по какой-то кафедре). Но – считать легко: среднее равно μ , а с.к.о. равно $(\mu)^{1/2}$. Да ещё, если μ – несколько единиц – уже к распределению Гаусса сходится. Так что наши таблицы 13 и 14 годятся и для этого распределения.

Пример 13 нормального распределения Гаусса. Мы им уже пользовались, и не раз. Это когда наши с вами измерения ПОДВЕРЖЕНЫ ДЕЙСТВИЮ множества небольших и случайных ошибок (факторов). Ну да: средних значений много, а слишком малых и слишком больших – заметно меньше, чем средних (хотя и данный ЗР может быть асимметричным, особенно при малых числе измерений n). Применять его для табл. 13 мы уже с вами пробовали в предыдущем, чисто техническом, фрагменте.

Краткие выводы. Теперь немного, хоть и частично, повторим ещё раз. Итак, вы приступили к работе по специальности. При этом прямо при вас меняется и номенклатура выпускаемой продукции, и даже прежняя – как-то модифицируется, изменяется, не говоря уж о единичном сначала выпуске новой продукции. Во всех случаях уже СУЩЕСТВУЕТ некая, принятая на

предприятию, система контроля качества в соответствии с действующими стандартами и требованиями

Вот здесь и могут оказаться полезными приведённые и придуманные нами же **Примеры 9 – 13**. А для этого надо не только хорошо представлять, как квалифицировать появление случайной величины (брака качества), но и по КАКОМУ ЗАКОНУ он распределён, а затем сделать соответствующие расчёты. И главный акцент здесь – «по какому закону распределения появление случайной величины («брака») следует ожидать? И это надо сделать именно на умоглядном уровне, при условии, конечно же, полученного вами приведённого здесь теоретического материала.

Ну а само его применение не должно составить для читающих данное учебное пособие заметного труда, если вспомнить знаменитую мысль, услышанную мной ещё в Школе юных математиков, от проф. Смышляева: «Что учёный человек, что неучёный – одинаковы в том, что ни то, ни другой НИЧЕГО о чём-то конкретном не знают. Но учёный всегда знает, где об этом можно прочитать...».

Если будут иметь место какие-либо затруднения, сделайте запрос на своём компьютере в любом поисковике и смотрите на появившийся текст. Если что-то вам нечитабельным показалось (бывает и такое) – обратитесь к сайту рядом с первоначально увиденным. Иногда просто поискать надо. В конечном итоге – найдёте всё, что необходимо для решения вашей задачи. Проверено!

Ну а на следующей части попробуем вам показать, что «нормальный закон распределения Гаусса» - это также легко и просто, как и прост переход к нему от дискретных ЗР: Бернулли, биномиального и Пуассона. А ещё один дискретный ЗР Бернулли – просто незаменим там, где он по содержанию своему окажется единственно уместным. Особенно в теории массового обслуживания, что, однако, несколько выходит за пределы поставленных нами целей. Теперь же рассмотрим подробнее наиболее нам привычный – закон нормального распределения СВ Гаусса.

5. ЗАКОН НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ГАУССА

Этот закон распределения случайных величин занимает особое положение среди иных законов распределения, поэтому рассмотрим его подробнее. Заметим только, что все ЗР СВ, рассмотренные ранее, характерны именно для дискретных переменных, тогда как НЗР – для непрерывных.

5.1. Общие положения

Как правило, математическая статистика (далее – просто «статистика») оперирует с исходными данными, основу которых являет собой исследуемая (или исследуемые) статистическая совокупность (СС) с символическим именем, скажем, $Y = \{y_i\}$, $i = 1, N$, где y_i – элементы СС, i - текущий индекс или «счётчик» элементов переменной с именем X , а число её элементов понимается нами как её «мощность» $|X| = N$ или число её элементов. А символом Y мы обозначаем имя этой исходной статистической совокупности потому, что нам будет необходимым воспользоваться математическим аппаратом вариационных рядов (ВР), представляющих собой отображение исходной совокупности с именем Y на более компактную и продуктивную (да и более привычное для нас имя X для новой статистической совокупности), но существующей уже в виде ВР путём построения отображения одной совокупности на другую:

$$\tau: Y \rightarrow X, \text{ в которой } N \rightarrow n \text{ (} n < N \text{ или } n \ll N \text{)}. \quad (44)$$

Так, выражение (44) предоставляет нам возможность вместо анализа большой по объёму совокупности Y представить в виде компактной совокупности с именем X , которая представляет собой имена строк-вариантов (а ВР – это компактная таблица) числом n , что весьма окажется полезным, как ещё раз убедимся в том далее.

Обычно статистика или её практически обязательный отдел под названием «Общая теория статистики» (ОТС) во многих статистических учебниках изложена с разной степенью полноты: где более подробно, где менее подробно. Но в большинстве случаев (а каждая книга, в том числе и учебник по статистике всегда, так или иначе, ограничена своим объёмом) остается т.н. «полу-знание» или, в лучшем случае, недоумение: а в чём же всё же заключается СУТЬ этого ЗНР? – Хотя и весьма понятны ситуации, когда его применение кажется (и действительно является) уместным. Ну это когда как бы «действуют множество незначительных и не учитываемых нами на исследуемый процесс причин (факторов), которые, к тому же, могут легко компенсировать действия друг друга», что также не создаёт дополнительных недопониманий.

Ну и ещё: глядя на график дискретных частот при построении, скажем, непрерывного вариационного ряда (ВР) в виде гистограммы или в случае дискретного ВР – в виде полигона распределения, - действительно подчас наблюдаем, что частот в середине в пределах общего размаха исследуемой выборки много, а по краям нашего предварительного, опытного (эмпирического) распределения – частот мало. И в целом, если мысленно провести огибающую кривую по вершинам полигона распределения или по серединам столбиков гистограмм, то часто мысленно наблюдаем именно «гауссову шапку», накрывающую и ограничивающую нарисованную нами картинку сверху, под которой средних значений много, а слишком малых и слишком больших - мало.

При проведении же более подробных и корректных статистических расчётов весьма часто бывает необходимым убедиться, а действительно ли исследуемая нами случайная величина, получившая при проводимой нами сначала экспликации (формализации) имя Y , распределена (точнее, её элементы y_i) в соответствии с НЗР? Или по каким-либо другим законам распределения? Бернулли, скажем, или по биномиальному закону распределения, или по закону распределения Пуассона? И список подобных

законов распределения приведённым здесь списком далеко ведь не исчерпывается, с чем, по всей видимости, так или иначе знакомы те, кто в своё время «проходил» такую учебную дисциплину, как «Математическая статистика и теория вероятностей».

Кстати, все перечисленные здесь законы имеют свои номинальные обобщения в виде записей, позволяющих отразить специфику каждого из приведённых здесь ЗР. Так, НЗР Гаусса обычно обозначается в литературе как $f_{X,\sigma}(x)$, где X – некоторый центр распределения (среднее), а σ – среднее квадратическое отклонение (с.к.о.) или стандартное отклонение, или стандартная ошибка. ЗР Бернулли записывается как $P_n(k)$, так на его графиках всегда показывается фиксированная величина – вероятность $p = const$ для каждого отдельного опыта, n – число опытов; k – появление ожидаемого события (в штуках) [2. – С. 52 – 59]. ЗР биномиальный обозначается как $b_{n,p}(k)$, где n – число испытаний; p – вероятность реализации k успехов [1. – С. 198]. ЗР Пуассона $p_\mu(k)$, где $\mu = k_{cp}$ – среднее число удачных исходов, k – сами удачные исходы из n испытаний (например, при $\mu \approx 9$ ЗР Пуассона трудно отличить от НЗР с теми же средним и дисперсией [3. – С. 215 – 216].

Остаётся только напомнить ещё раз, что законы Бернулли, биномиальный и Пуассона отражают *дискретные* изменения СВ, тогда как НЗР Гаусса отражает свойство *непрерывности* исследуемой СВ, что тоже, в известном смысле, является на практике абстракцией: мы ведь всегда работаем с конечным набором случайных чисел, ведь так?

Таким образом, здесь были приведены отдельные (и основные для наших исследований) законы распределения как ПРЕДЕЛЬНЫЕ распределения – то есть теоретические идеализации, к которой никогда практически нельзя абсолютно точно приблизиться на эксперименте [3. – С. 113]. Но без них в статистике хоть как-то обойтись отчего-то не получается.

Так что необходимо, пожалуй, остановиться более подробно на каждом из упомянутых здесь ЗР, и продолжим это с именно НЗР К.Ф. Гаусса.

5.2. Закон нормального распределения Гаусса

Обычно закон нормального распределения (ЗНР) Гаусса обозначается как просто экспоненциальная функция вида:

$$y = f_{x,\sigma}(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (45)$$

В формуле (45) видно, что когда $x = 0$, то это отдельное значение функция Гаусса (45) строго равна единице. Однако, при увеличении x от значения 0 в обе стороны, функция будет тем быстрее уменьшаться, чем меньше значение σ – «параметра ширины», как ещё его назовём (т.е. с.к.о.). Да и просто здравый смысл подсказывает: чем меньше значение с.к.о., тем быстрее убывает функция (или «шапка») Гаусса, тем острее её пик (или больше «эксцесс» как момент четвёртого порядка. В этой связи, видимо, необходимо напомнить, что момент первого порядка означает математическое ожидание (её оценка – средняя величина исследуемой переменной); момент второго порядка означает дисперсию (квадратный корень из неё – есть с.к.о., как мы помним); момент третьего порядка измеряет степень асимметрии (правой или левой), тогда как четвёртый момент (эксцесс) означает, как уже упоминалось, степень как бы «заострённости» кривой Гаусса.

Отметим также сразу, что в данном случае для нас представляет интерес прежде всего НОРМИРОВАННЫЕ значения функции Гаусса, то есть такие значения, которые находятся под «шапкой» знаменитой кривой Гаусса, которые в пределе в сумме все вместе взятые, исчисляются **единицей**. То есть вся площадь под нормированной кривой равна строго единице. Тогда выражение для такой функции плотности распределения СВ Гаусса будет выглядеть, в отличие от табличной функции (30) для $\varphi(t)$, следующим образом:

$$f_{x,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x - X)^2}{2\sigma^2}\right]. \quad (46)$$

В формуле (44) x – это некоторое текущее значение исследуемой переменной, X – центр её распределения. А вот знаменатель первой дроби $\sigma\sqrt{2\pi}$ в выражении (46) как-то настораживает: откуда он появился, и откуда именно в таком виде? Однако, для начала давайте-ка нормированный вид функции Карла Фридриха Гаусса, описывающей НЗР вида (46), сначала нарисуем (ведь по свидетельствам психологов, как уже упоминали, до 90% всего объёма информации поступает к нам всё же через зрительный канал).

Визуализация нормированного ЗНР. Кривую НЗР Гаусса представим графиком на рис. 13, используя возможности табличного процессора Microsoft-Excel. Но для этого необходимо сначала составить рабочую таблицу, которую предварительно необходимо рассчитать, прежде чем отобразить на предполагаемом нами рис. 13. Для этого лишь надо задаться произвольными исходными данными по оси абсцисс графика на рис. 13 для расчёта функции вида (46).

Для наглядности и последующего удобства примем в качестве центра распределения, например, $X = 5$ единиц измерения. Пусть также значения ранее рассчитанной нами «сигмы» (с.к.о.) будут равны, скажем, 1,5 единицы. Хотя бы для графического удобства.

Итак, $X = 5$; $\sigma = 1,5$. Ну а после этого попытаемся построить «плавный» график, для чего с желательным мелким шагом рассчитаем для того исходную таблицу (см. табл. 15). Понятно, что не у всех обучаемых смартфоны располагают одинаковыми вычислительными возможностями: у одних таких возможностей больше, у других меньше. Поэтому автор воспользуется возможностями своей СТАНДАРТНОЙ системой управления базами данных (СУБД) с именем Fox-Pro 2.5 – под операционную систему MCDOS.

Итак, приняли сами: центр распределения $X = 5$ единиц; $\sigma = 1,5$ единиц; сами также задали для расчётов шаг = 0,5 единиц, чтобы детальнее показать, что такое « $\pm 3\sigma$ ». Расчёты по формуле (46) сведём в табл. 15, причём его множитель обозначим через « $a1$ », а множитель за « $a2$ »

Таблица 15

Результаты расчёта по формуле (46)

Формулы	1	2	3	4	5	6	7	8	9	10
x(i)	1	2	3	4	5	6	7	8	9	10
a1	0,266	0,266	0,266	0,266	0,266	0,266	0,266	0,266	0,266	0,266
a2	0,03	0,14	0,41	0,80	1,00	0,80	0,41	0,14	0,03	0,01
f = a1·a2	0,008	0,036	0,109	0,213	0,266	0,213	0,109	0,036	0,008	0,001

Из табл. 15 видно, что первый сомножитель уже нормированной формулы (46) является для принятых центра распределения $X = 5$ (выделено шрифтом) и величины среднего квадратического отклонения (с.к.о.) $\sigma = 1,5$. Напомним, что вся площадь под кривой Гаусса строго равна ЕДИНИЦЕ.

И ещё видно в строке для «a2» - экспоненциального множителя формулы (46), что в степени «0» экспонента равна строго единице, как и любая степенная функция. Теперь – о том, откуда взялся первый множитель «a1». Так, в [3. – С. 119 - 120] показано, что функция Гаусса (47) представляет собой колоколообразную кривую с центром в $x = 0$ (если без X):

$$\exp \left[-\frac{(x - X)^2}{2\sigma^2} \right] \quad (47)$$

и достигает максимума при $x = X$, а также спадает симметрично в обе стороны от центра распределения X (см. рис. 13).

С.к.о. σ также отмечены по оси абсцисс рис. 13 и располагаются графически на ней в районе знаков « \Rightarrow ». Ну и, наверное, будет нелишним напоминать, что максимуму нормированной плотности распределения будет соответствовать максимум, равный

$$\frac{1}{\sigma\sqrt{2\pi}}$$

для середины распределения, то есть для $x = 5$. На рис. 13 видно, что при увеличении чисел σ по обе стороны от середины распределения график асимптотически приближается к оси абсцисс.

Плотность вероятности $f(x)$

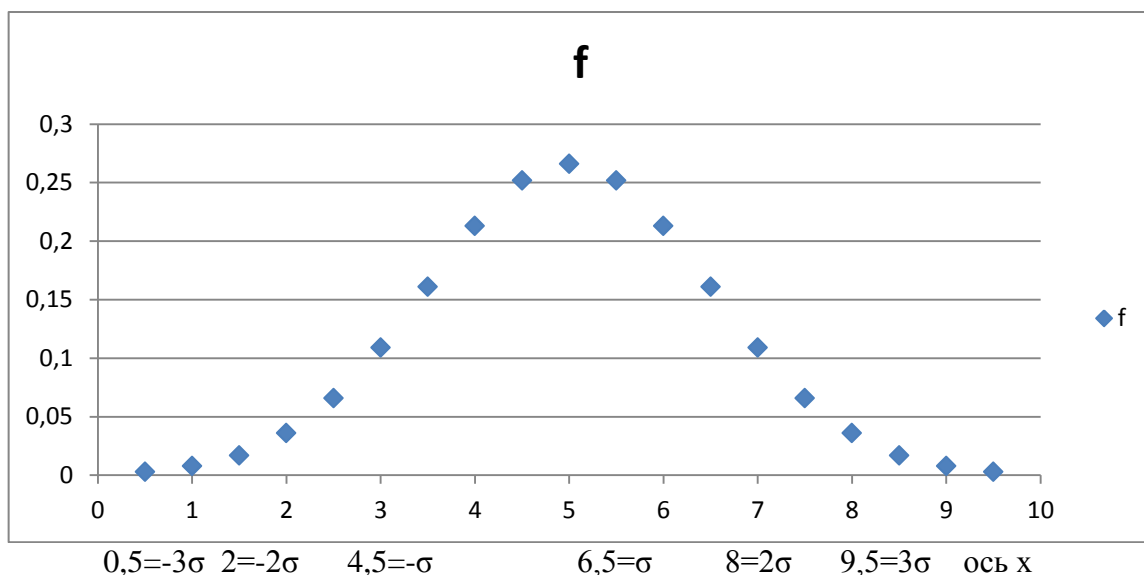


Рис. 13. Закон нормального распределения Гаусса; центр = 5; $\sigma = 1,5$

Теоретическая сторона вопроса. Однако, выражение (47) является не окончательным, так как оно ещё не описывает ПРЕДЕЛЬНОЕ РАСПРЕДЕЛЕНИЕ, поскольку любое распределение должно быть **нормировано**, иначе говоря, приведено к единице, то есть должно удовлетворять условию:

$$\int_{-\infty}^{\infty} f(x)dx = 1. \quad (48)$$

Ну и чтобы выполнить это условие (48), мы (именно мы сами!) введём такое обозначение $f(x)$ с учётом формулы (47):

$$f(x) = N \exp \left[-\frac{(x - X)^2}{2\sigma^2} \right]. \quad (49)$$

В формуле (49) умножение на N не изменяет ни формы, ни положения максимума при $x = X$, как уже отмечалось ранее, что косвенно и представлено на рис. 13. Теперь же нам надо выбрать «нормировочный множитель» N таким образом, чтобы функция $f(x)$ в формуле (49) была бы **НОРМИРОВАНА** в соответствии с требованием (48). Итак, подставив формулу (49) в формулу (48), будем иметь:

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} N \left[-\frac{(x - X)^2}{2\sigma^2} \right] dx. \quad (50)$$

Заметим только, что при оценке и решении интегралов такого вида ВСЕГДА нужно произвести ЗАМЕНУ переменных, чтобы упростить подынтегральное выражение формулы (50). Для этого обозначим $(x - X) = y$ (в этом случае, очевидно, $dx = dy$) и получим:

$$\int_{-\infty}^{\infty} N \exp \left[-\frac{(x - X)^2}{2\sigma^2} \right] dx = N \int_{-\infty}^{\infty} \exp \left[-\frac{y^2}{2\sigma^2} \right] dy. \quad (51)$$

Затем ещё раз обозначим в (51) и снова заменим $(y/\sigma) = z$ (в этом случае опять же $dy = \sigma dz$, что не менее очевидно, как и в прошлом случае), после чего формула (51) приобретёт вид (если с.к.о. то же, то есть постоянная величина, которую, как мы помним, вполне можно вынести за знак интеграла) формулы (52):

$$\text{Формула (51)} = N\sigma \int_{-\infty}^{\infty} \exp \left(-\frac{z^2}{2} \right) dz. \quad (52)$$

Полученный интеграл в выражении (52) – один из стандартных интегралов в математической статистике, его можно вычислить уже известными нам методами, что здесь, понятно, опускаем, и приводим лишь результат конечных вычислений:

$$\int_{-\infty}^{\infty} \exp \left(-\frac{z^2}{2} \right) dz = \sqrt{2\pi}. \quad (53)$$

А теперь возвратимся к выражениям (50) и (52), и запишем, что

$$\int_{-\infty}^{\infty} f(x) dx = N\sigma\sqrt{2\pi}. \quad (54)$$

Так как интеграл (54) ДОЛЖЕН БЫТЬ равен единице, то есть НОРМИРОВАН, как мы с вами ранее и упоминали, то и соответствующий ему нормировочный множитель N следует выбрать как

$$N = \frac{1}{\sigma\sqrt{2\pi}}. \quad (55)$$

Кратко подведём итог наших преобразований (48) – (54): мы, авторы данного учебного пособия, даже не имеем и тени сомнений в том, что приведённые здесь преобразования были прозрачны и понятны всем и каждому. Разве не так? – Да, дифференциальное и интегральное исчисление традиционно принадлежит к области знаний по высшей математике, но и использованные здесь подходы преобразований от вида (48) до вида (55) включительно не являются для наших будущих специалистов в области управления качеством чем-то принципиально новым. Согласны?

Однако, для пользы общего дела ещё раз вернёмся не сколько назад на анализ вида выражения (46), например. Итак, если снова посмотреть на выражение (46), то увидим, что это и есть тот самый нормировочный множитель как множимое в данном выражении, и которое для табл. 15 мы обозначили как $a1$. Вот и говорят, что РЕЗУЛЬТАТЫ ИЗМЕРЕНИЙ (НАБЛЮДЕНИЙ) РАСПРЕДЕЛЕНЫ НОРМАЛЬНО, то есть имеет место гауссовский НЗР вида $f_{X, \sigma}(x)$, если их предельное распределение описывается функцией вида (46). Ну, это мы опять «о форме выражения НЗР К.Ф. Гаусса».

Тогда, рассуждая далее, по «закону трёх сигм» [10] и получается, что в пределах (см. рис. 13) $\pm 1\sigma$ от центра распределения X (у нас в табл. 1 и на рис. 1 $X = 5$), то туда как бы «попадают» 68,3% всех измерений (наблюдений), если они распределены нормально; в пределах $\pm 2\sigma$ – уже «попадут» 95,4% всех случаев, а в пределы $\pm 3\sigma$ – даже до 99,7% всех случаев проявления ИССЛЕДУЕМОЙ случайной величины. Далее кривая Гаусса на рис. 13, что видно весьма наглядно, асимптотически приближается к оси абсцисс. И, естественно ожидаемые нами, 100% всех случаев будут далеко выходить за пределы $\pm 3\sigma$ (хотя в технических условиях для выполнения отдельных работ для заинтересованных ведомств назначается диапазон исследований, например, в $\pm 5\sigma$ и даже выше (по сведениям авторов – для

отдельных оборонных исследований). Так что «попадание» (практическое или модельное – здесь разницы нет никакой) в интервал $\pm 3\sigma$ от середины распределения СВ будет при бесконечном повторении практических опытов всегда 99,7%, в чём при достаточном усердии в повторении практических опытов методически легко убедиться.

Просто напомним к уже сказанному, что согласно [11], при *доказанном* НЗР вероятности, о чём на занятиях слышали не раз, распределяются по числу с.к.о. с традиционным именем σ следующим образом:

$$P(-1\sigma_x < x_{cp}^{B3} < +1\sigma_x) = \Phi\left(\frac{1}{\sqrt{2}}\right) = 0,683 \text{ или } 68,3\% ; \quad (56)$$

$$P(-2\sigma_x < x_{cp}^{B3} < +2\sigma_x) = \Phi(\sqrt{2}) = 0,954 \text{ или } 95,4\% ;$$

$$P(-3\sigma_x < x_{cp}^{B3} < +3\sigma_x) = \Phi\left(\frac{3}{\sqrt{2}}\right) = 0,997 \text{ или } 99,7\%.$$

А символы Φ (табулированные функции) в системе неравенств (56) – это, по сути своей, есть численные табличные значения, которые находятся в Приложениях к практически любому учебнику статистики. Попутно также отметим, что число «сигм» заказчиком может быть существенно увеличено, например, до 5 – 6, чем и будет, как уже упоминалось нами, обеспечена достоверность более 99,7% (см. асимптоту Гаусса на рис. 13, которая НИКОГДА не пересечёт ось абсцисс – по определению асимптоты).

5.3. Порядок применения закона нормального распределения

При работе с реальными дискретными вещественными (т.е. выраженными десятичными) числами или показателями предварительно надо построить ИНТЕРВАЛЬНЫЙ вариационный ряд (ВР) с вычисленными средним взвешенным x_{cp}^{63} как центром распределения X и шириной размаха – с.к.о. σ_x (мы это с вами уже умеем, не так ли?), а также **шагом h** (для чего надобно заранее определённый размах выборки разделить на число интервалов по приближительной и уже известной формуле Стерджесса). А

построить ВР – значит создать компактную таблицу с i вариантами-строками и посчитать в каждой из них, сколько раз туда попадают наши исходные вещественные (или дискретные) показатели. Этим мы и определяем ряд эмпирических (из опыта) частот f_i .

Подготовка к решению (повторению решения) уже знакомого примера. И тогда уже непосредственно воспользоваться не формулой вида (46), а более простой для применения закона Гаусса [5. – С. 43 - 45]:

$$y = f_{x,\sigma}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), \quad (57)$$

где y – ордината кривой нормального распределения;

$$t = \frac{(x - x_{\text{ср}}^{\text{вз}})}{\sigma_x} - \text{нормированные отклонения.}$$

Тогда при ВЫРАВНИВАНИИ ВР по кривой нормального распределения ТЕОРЕТИЧЕСКИЕ ЧАСТОТЫ $f_i^{\text{теор}}$ вычисляются по формуле (58):

$$f_i^{\text{теор}} = \frac{Nh}{\sigma_x} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) = \frac{Nh}{\sigma_x} \cdot \varphi(t). \quad (58)$$

В формуле (58) число $N = \sum f_i$ – означает *сумму всех опытных частот*, что необходимо помнить по отношению к проверке критерия согласия по ЗР Пуассона, что уже нами отмечено. Потому что для НЗР число N – есть число уже по интервалам построенного ВР, что важно всегда помнить.

Понятно, что в выражении (58) первый множитель $(Nh/\sigma_x) = \text{const}$ для каждого отдельного примера расчёта ВР, а второй и третий множители формулы (58) табулирована как $\varphi(t)$, и её легко получить, например, из Приложения 1 [5. – С. 128 - 129], что уже нами рассматривалось не раз.

Итак, эмпирические частоты f_i мы получили ещё на этапе предварительного построения ВР, а теоретические частоты $f_i^{\text{теор}}$ нашли по формуле (58) с учётом нормированного отклонения t . Остаётся только воспользоваться знакомыми нам уже критериями СОГЛАСИЯ Пирсона, Романовского, Колмогорова и сделать вывод: распределены ли элементы исследуемой нами СВ по нормальному ЗР или нет.

Ну и, как мы видим, заметного противоречия между формулами (46) – (55) и реальными расчётными формулами (57) и (58) никаких формальных противоречий не существует. Приведённые здесь формулы (46) – (55) лишь имеют цель показать, что именно представляют из себя и откуда получаются рекомендуемые расчётные формулы (57) и (58) с учётом предварительных расчётов нормированных отклонений t .

А для практических расчётов воспользуемся уже рассмотренным и решённым нами примером ещё раз. Да и в самом-то деле: новом пример с новыми цифровыми показателями действительно может представить для читателя лишь некоторый интерес, который при прежних, известных исходных данных уже решённого каким-то иным способом данного примера позволит увидеть процесс решения с несколько иной стороны, что, в свою очередь, объективно способствует комплексному видению самого процесса в целом.

Пример для конкретного решения. Пример исходных данных мы выбрали тот же, что и при прежних исходных данных того же результата выборочного контроля, что приведены нами были ещё в самом начале повествования в табл. 1 и табл. 2 (см. здесь как табл. 16).

Понятно, что мы из предварительных расчётов партии с номерами 5 и 9 забраковали (выделены шрифтом), и у нас в табл. 2 осталось только восемь партий из первоначальных десяти. Но в табл. 16 приведём весь первоначальный набор исходных данных при начале контроля качества.

Таблица 16

Результаты контроля качества в 10-ти партиях по 100 изделий

№ партии	1	2	3	4	5	6	7	8	9	10	Сум- ма
Объём партии, n_i , шт.	100	100	100	100	100	100	100	100	100	100	1000
Браков. изд., np_i , шт.	1	2	0	3	14	4	1	2	15	3	45
Доля события A, p_i , %	1	2	0	3	14	4	1	2	15	3	-

Понятно, что объём партии при этом сократился с 1000 до 800 изделий, а число бракованных изделий – до 16-ти: прежние браки за вычетом

оставшихся $(45 - 14 - 15) = 16$. Ну а далее – как всегда: построение ВР из элементов СВ (четвёртая строка табл. 2) с подсчётом эмпирических частот f_i , а затем – расчёт теоретических частот $f_i^{теор}$, по формуле (58) с использованием табличных значений функции $\varphi(t)$. А когда имеем f_i и $f_i^{теор}$, то последовательно используем критерии согласия Пирсона, Романовского, а также Колмогорова, если сочтём необходимым. Также кстати, отметим для себя, что по ряду соображений критерий Колмогорова не является «мощнее» двух предыдущих критериев по свидетельству многих авторов.

Однако, его применение даёт возможность оценки степени достоверности полученных результатов в виде значения их вероятности в относительных единицах или процентах. Так что пренебрежение в использовании данного критерия согласия в ряде случаев не всегда является оправданным.

Построение ВР. Для этого (уже, видимо, стало привычным для читателей, ибо максимум «повторение – мать учения» ещё никто, помнится, не отменял!) необходимо определиться с его величиной шага $h = R / n$, где R – это размах выборки: $R = y_{max} - y_{min} = 4 - 0 = 4$ (%), n – число интервалов в ВР (вариантов):

$$n \approx 1 + 3,322 \lg N = 1 + 3,322 \lg 8 = 1 + 3,322 \cdot 0,9 = 3,99 \approx 4 \text{ (инт.)} = nvr.$$

Тогда величина шага $h = R/n = 4\% / 4 = 1,0\%$, начиная от минимального значения, равного нулю. Тогда начинаем строить ВР в табл. 17 (бывшая табл. 6). $N = \sum f_i = 8$ штук. Так что **создание ВР** – это и есть построение отображения (44).

Повторное здесь примечание: знак (+) в заголовке строк означает, как уже упоминалось, что, согласно [6], интервалы по правой стороне считаются закрытыми, а слева - открытыми. Ну и ещё отдельные замечания. Число интервалов-вариантов ВР табл. 17 равно $nvr = n = 4$; с.к.о. σ - есть корень квадратный от дисперсии $D = \sigma^2$ – иногда записывают и так.

Таблица 17

Вариационный ряд исходной совокупности $N_2 = 8$ единиц, где $n = nvr = 4$ – число вариантов

Варианты, i	Границы инт. $x_{нач} \div x_{кон} (+)$	Эмпирические частоты f_i		Накопленная частота q_i	Средние значения x_i^{cp} интервалов
		счёт из N_2	значения		
1	2	3	4	5	6
1	0,0 – 1,0	///	3	3	0,5
2	1,0 – 2,0	//	2	5	1,5
3	2,0 – 3,0	//	2	7	2,5
$nvr = 4$	3,0 – 4,0	/	1	8	3,5
Всего:	-	8	8	-	-

Тогда с.к.о. вычисляется как корень квадратный из дисперсии.

$$x_{cp}^{B3} = \frac{\sum_{i=1}^n x_i^{cp} f_i}{\sum_{i=1}^n f_i} = \frac{0,5 \cdot 3 + 1,5 \cdot 2 + 2,5 \cdot 2 + 3,5 \cdot 1}{3 + 2 + 2 + 1} = 1,625 \approx 1,62.$$

$$D = \frac{\sum_{i=1}^n (x_i^{cp} - x_{B3}^{cp})^2 f_i}{\sum_{i=1}^n f_i} = \frac{(0,5 - 1,62)^2 3 + (1,5 - 1,62)^2 2 + (2,5 - 1,62)^2 2 + (3,5 - 1,62)^2 1}{3 + 2 + 2 + 1} = \frac{8,875}{8} = 1,1094.$$

$$\sigma_x = \sqrt{D} = \sqrt{\sigma_x^2} = \sqrt{1,1094} = 1,05.$$

Итак, нами получены взвешенные значения – собственно среднее значение исследуемой переменной с именем X и её с.к.о. σ_x вокруг него – по три σ_x от центра x_{B3}^{cp} , где непрерывная функции Гаусса (в виде колпака – см. рис. 13) будет в этой точке иметь максимум. Не в точке 5, как это было на рис. 13 для данного примера, но в точке $x_{B3}^{cp} = 1,62$ – вся разница с рис. 13 – на нём представлена кривая Гаусса при среднем значении, равном пяти, как помнится нам.

Напомним также, что всё мы это делали для того, чтобы произвести так называемое «выравнивание вариационного ряда», то есть сделали предпосылки к т.н. «построению теоретического распределения» [5. – С. 45].

Эмпирические частоты, полученные ещё в ВР, приведённым в табл. 17, мы уже с вами получили – это f_i , которые были определены нами для всех четырёх интервалов ВР (см. также табл. 3, столбец 4 – пример ведь тот же) как 3, 2, 2 и 1 (всего восемь: $N = 8$ шт.).

Построение продолжения табл. 3 для ВР в виде табл. 4. Величина $(Nh/\sigma_p) = (8 \cdot 1) / 1,05 = 7,62 = const$ для нашего конкретного исследования. Необходимо достроить ещё ранее приведённые табл. 6 до табл. 7 с получением искомых значений частот, но уже теоретических f_i^{meop} , $i = 1, nvr = 4$. А уже после построения табл. 17 как продолжения табл. 2 и непосредственно приступаем к расчёту т.н. «критериев согласия».

Итак, $const = (Nh/\sigma_p) = 7,62$. Надо рассчитать показатель t из выражения (30), войти в Приложение 1 [5], найти табличные значения $\varphi(t)$, умножить их на константу для конкретного ВР, а затем округлить полученные значения до ЦЕЛЫХ ЗНАЧЕНИЙ. Это и будут искомые значения f_i^{meop} . А частоты f_i мы уже определили в табл. 17, столбец 4.

Для построения табл. 18 для определения вариантов f_i^{meop} лишь воспользуемся столбцом 6 из табл. 4 со средними значениями i -вариантов ($i = 1, nvr = 4$) для вычисления $(x_i^{cp} - x_{вз}^{cp})$ – здесь для большей ясности $p_{вз}^{cp} = y_{вз}^{cp}$. Однако, для большей цельности восприятия, всё-таки в бывшей табл. 7 (здесь – табл. 18) вторым после номеров вариантов столбцом назначим 6-й столбец из также бывшей табл. 6 (теперь – табл. 17) – т.е. просто повторим его.

Таблица 18

Вычисление теоретических частот f_i^{meop} для ранее построенного ВР

Варианты, I	Средние значения x_i^{cp} интервалов	$(x_i^{cp} - x_{вз}^{cp})$	$t = (x_i^{cp} - x_{вз}^{cp})/\sigma_p$ (округлять до сотых; знаки + и – не учитывать)	$\varphi(t)$ из Приложения 1 [5. –С. 128]	$f_i^{теор} = (Nh/\sigma_p) \cdot \varphi(t)$ с округлением до целых
1	2	3	4	5	6
1	0,5	- 1,12	- 1,07	0,2251	1,72 \approx 2
2	1,5	- 0,12	- 0,11	0,3965	3,02 \approx 3
3	2,5	0,88	0,84	0,2803	2,14 \approx 2
n = 4	3,5	1,88	1,79	0,0804	0,61 \approx 1
Всего теоретических частот:					7,49 \approx 8

Итак, теоретические частоты $f_i^{теор}$ ещё раз найдены (см. столбец 6)

Расчёты по критериям согласия. Согласно в табл. 3 (в ВР) мы насчитали также восемь опытных (эмпирических) частот f_i . Применяв необходимые преобразования в табл. 4, также получили те же восемь, но уже теоретических частот $f_i^{теор}$ (см. табл. 19). Это и даёт нам основания для применения нами уже знакомого *критерия согласия* Пирсона χ^2 [5. – С. 48] согласно формуле (31), расчёты по которой проведём в рабочей табл. 19 (ранее – табл. 8):

Таблица 19

Расчёт значения критерия Пирсона χ^2 по формуле (59)

Варианты, I	Эмпирические частоты f_i	Теоретические частоты $f_i^{теор}$	$(f_i - f_i^{теор})$	$(f_i - f_i^{теор})^2$	$(f_i - f_i^{теор})^2 / f_i^{теор}$
1	2	3	4	5	6
1	3	2	1	1	0,50
2	2	3	- 1	1	0,33
3	2	2	0	0	0
n = 4	1	1	0	0	0
Всего:	8	8	-	-	$\Sigma = \chi^2_{расч} = 0,83.$

Прежде всего, число степеней свободы df при выравнивании по *кривой нормального распределения* Гаусса К.Ф.: $df = n - r - 1$, где число r означает число параметров, которые применялись для расчёта теоретических частот, которых было ровно два: среднее взвешенное и с.к.о.; поэтому $r = 2$, и число степеней свободы $df = 4 - 2 - 1 = 1$ (меньше df для таблицы Пирсона просто невозможно). Вспомним также, что в процессе нахождения теоретических частот при проверке гипотезы на распределение СВ по закону Пуассона мы пользовались только одним параметром – средним значением, поэтому для того случая $r = 1$.

Далее примем наиболее часто используемый уровень значимости $\alpha = 0,05$ и обратимся к таблице «Значения χ^2 – критерия Пирсона при уровне значимости 0,10; 0,05 и 0,01» (то есть при ошибках в 10%, 5% и 1% - соответственно) – см. Приложение 4 [5. – С. 132].

Итак, по таблице значений χ^2 – критерия Пирсона для степеней свободы $df = 1$ (первая строка таблицы) и уровня значимости $\alpha = 0,05$ определяем, что $\chi^2_{\text{табл}} (df = 1; \alpha = 0,05) = 3,84$. Так как полученное в результате заполнения рабочей таблицы 19 расчётное (фактическое) значение $\chi^2_{\text{расч}} = 0,83 < \chi^2_{\text{табл}} = 3,84$, то, следовательно, можно считать РАСХОЖДЕНИЯ между эмпирическими (фактическими) и теоретическими значениями частот по вариантам **случайными**, а выдвинутую ранее гипотезу (предположение) об БЛИЗОСТИ эмпирического (фактического) к нормальному закону распределения НЕ ОПРОВЕРГАЕТСЯ.

Ну и далее в качестве альтернативного критерия согласия воспользуемся хотя бы критерием Романовского. Для этого вычисляется сам критерий $Rom_{\text{расч}}$ и сравнивается с граничным $Rom_{\text{гран}} = 3$. Если $Rom_{\text{расч}} < Rom_{\text{гран}}$, то гипотеза о нормальном распределении не отвергается. Итак, [5. – С. 48 – 49] по известной нам уже формуле (32):

$$Rom_{\text{расч}} = \frac{|\chi^2 - df|}{\sqrt{2 \cdot df}} = \frac{|0,83 - 1|}{\sqrt{2 \cdot 1}} = \frac{0,17}{1,414} = 0,12 < Rom_{\text{расч}} = 3.$$

Гипотеза о нормальном распределении исходной статистической совокупности, отражающей процент бракованных изделий (появление события A в примере 1) в исследуемых партиях по 100 штук деталей (готовой продукции) $p_i, \%$ (см. табл. 2 примера) и после применения критерия согласия Романовского также подтверждается.

Хотя критерий Колмогорова и рекомендуется применять при числе элементов СВ более 50-ти (у нас только их всего $8 = N$), он всё же по отношению к двум предыдущим критериям (31) и (32) является, как уже упоминалось, как бы «менее мощным» по отношению к критерию Пирсона.

А в заключение данного учебного пособия авторы сделали попытку проверки справедливости автора [15] (и не без оснований) об априорной его попытки использования соотношений из ЗР СВ по Пуассону без предварительной проверки «нулевой гипотезы» на одном из примеров,

приведённых автором в [15. – С. 111 - 113], для чего сочли необходимым посвятить этим расчётам ещё одну главу.

6. ПРОВЕРКА ОЦЕНКИ Э. ДЕМИНГОМ КВАЛИФИКАЦИИ СПЕЦИАЛИСТОВ ПО ДОПУЩЕННЫМ ИМИ ОШИБКАМ

Необходимость проверки оценки Э. Демингом специалистов. При знакомстве с 10-м изданием монографии Э. Деминга [15] авторами работы было обращено внимание на использование автором книги формул, характерных для закона распределения случайных величин по закону Пуассона с получением результатов по оценке верхней и нижней предельных границ в отдельных примерах: ВП и НП – соответственно. В качестве исходных данных для характерного примера реализации принципа «встроенного качества» использованы зафиксированные наблюдения за числом ошибок, совершённых 9-ю специалистами, которые расположились в пределах от 4-х до 23-х [15. - С. 112 – 113] за некоторый промежуток времени. В итоге автором [15] были оценены предельные границы нижняя (НП) и верхняя (ВП): $НП = 2$ ошибкам; $ВП = 23$ ошибкам и сделан вывод о том, что все контролируемые специалисты допустили ошибки в так называемом «системном качестве», отчего все они достойны повышения денежного содержания.

Однако, исходная гипотеза Э. Деминга о распределении числа ошибок каждым специалистом, по мнению авторов, нуждалась скорее в выдвижении гипотезу о нормальном их распределении по закону Гаусса, проверка которой, правда, на том же критерии согласия Пирсона сопряжена со значительно большими вычислениями. При этом конечные результаты оказались практически прежними: $НП = 1$ ошибка, $ВП = 24$ ошибки. Однако, как показали результаты вычислений взвешенного среднего и дисперсии в рамках неполной проверки на нормальное распределение, являются вполне убедительными и для гипотезы о распределении Пуассона. Следовательно, в

изложенном материале главным, видимо, является не столько проверка гипотезы о характере распределения исследуемой величины (здесь – ошибок специалистов), но сам метод управления качеством, сопряжённый к тому же с минимальными вычислениями. А в качестве ключевых слов будем по-прежнему использовать такие, как случайная величина, закон распределения, среднее взвешенное, дисперсия, стандартное отклонение, гипотеза, критерий согласия, контрольные границы.

Постановка задачи. Поскольку далеко не все приведённые примеры из книги [15] кажутся вычисленными корректно, возникла настоятельная необходимость в их проверке. Вместе с тем в книге [15. - С.8] заявлено, что в ней приведена «новая концептуальная, научно-методологическая база менеджмента качества, а также иное миропонимание, другая ценностная ориентация», поэтому со стороны читателей требуется более внимательное и к её общему содержанию, и к отдельным примерам в ней.

В качестве исходных данных воспользуемся данными контроля девяти специалистов в течение определённого промежутка времени [15. - С. 112 – 113] по числу допущенных ими ошибок как случайных величин (СВ), которые представим в виде случайного множества с именем $Y = \{y_j\}, j = 1, m = 9$ (специалистов) или $Y = \{10, 15, 11, 4, 17, 23, 11, 12, 10\}$. Общее же число ошибок $N = \sum_{j=1}^m y_j = 113$.

Далее Э. Деминг ставит вполне справедливый вопрос: а есть ли персонально виноватые в их совершении или все девять сотрудников находятся по числу своих своим ошибок в «одном качестве», а причины руководителям предприятия следует искать не в конкретных специалистах, но в самой «системе», сложившейся на предприятии в силу многих причин?

Для построения аргументированного ответа на этот вопрос автор [15] поступает следующим образом. Во-первых, он постулирует распределение случайной величины «ошибка» в переменной Y по закону распределения Пуассона, а во-вторых пользуется его известным свойством: его стандартное

отклонение от среднего $x_{cp} = \mu$. Тогда стандартное отклонение или среднее квадратическое отклонение (с.к.о.) σ , как мы привыкли его обозначать, есть квадратный корень из него же [3. - С. 215 – 217]: $\sigma = \sqrt{\mu}$.

И уж далее, пользуясь известным законом «трёх сигм» [22], которые, как и в нормальном ЗР Гаусса, охватывают 99,7% диапазона изменения СВ: \pm три «сигма». Далее автором [15] приводится формулу для определения нижней и верхней допустимых пороговых границ (НП и ВП – соответственно):

$$\text{ВП, НП} = \mu \pm 3 \cdot \sqrt{\mu}. \quad (59)$$

Необходимо только вычислить это среднее значение μ . Самым точным его значением будет всегда механическое среднее, которое найдём для вычисления взвешенного среднего если $f_j = f = \text{const} = 1$ для всех событий y_j по формуле (2) – фактически вычисляем механическое среднее:

$$y_{cp}^{вз} = \mu = \frac{\sum_{j=1}^m y_j \cdot f_j}{\sum_{j=1}^m f_j}. \quad (60)$$

$$\mu = \frac{1}{9} (10 + 15 + 11 + 4 + 17 + 23 + 11 + 12 + 10) = 12,555. \quad (61)$$

Тогда формула (59) примет вид:

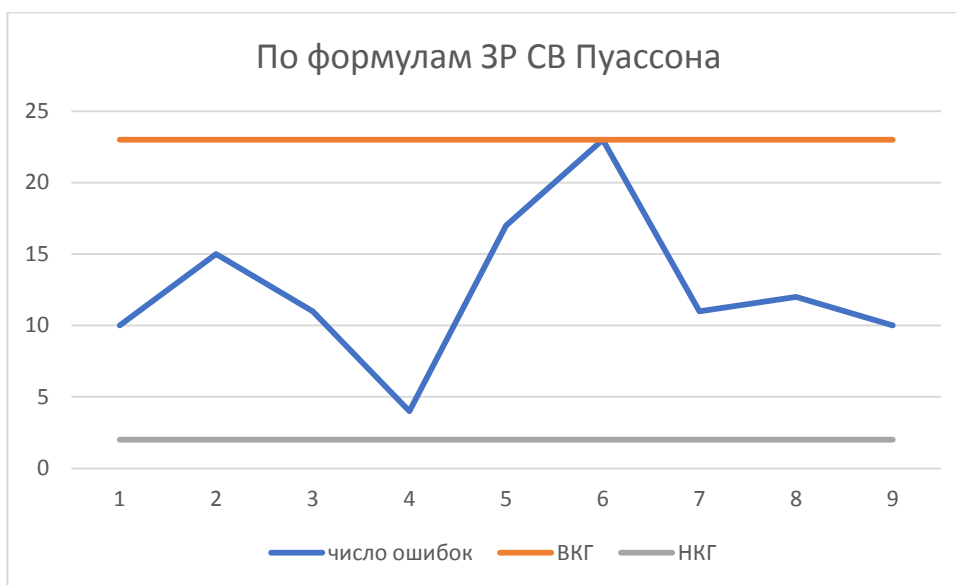
$$\text{ВП; НП} = 12,5555 \pm 3 \cdot \sqrt{12,5555} = 23,18 \approx 23 = \text{ВП}; 1,92 \approx 2 = \text{НП}.$$

Автором [15] получены верхняя (23 ошибки) и нижняя (2 ошибки) контрольные границы, что, наряду с исходными данными для большей наглядности представим на отдельном рисунке (см. рис. 14). Из рис. 14 видно, что все ошибки расположены между контрольными границами, вычисленными автором [15] при распределения их как СВ по закону Пуассона.

Иначе говоря, для выдвинутой гипотезы о ЗР СВ Пуассона дисперсия равна оценке математического ожидания: $\sigma^2 = \mu$ [3. - С. 216], и $\sigma = 3,5434$ ошибки. То есть контрольные границы ВКГ или ВП = 23 ошибки; НКГ или НП = 2 ошибки. Очевидно, что все ошибки располагаются в интервале от 4-х до 23-х, те есть в вычисленных пределах.

Из этого Э. Деминг делает следующий вывод: «Все допущенные 9-ю специалистами ошибки можно отнести за счёт «системы», а не за счёт отдельных специалистов. Следовательно, все специалисты достойны повышения оплаты труда». Да, хоть и воспринимается несколько парадоксально, но справедливо ведь по Э. Демингу.

Число ошибок



условные № специалистов

Рис. 14. Взаимное положение границ и числа ошибок по Демингу[15].

Но мы сделаем попытку присоединиться к нему лишь убедившись, что исследуемая СВ с именем $У$ действительно распределена по ЗР Пуассона, что и проверим по известным правилам.

Проверка рабочей гипотезы Э. Деминга о ЗР СВ по закону Пуассона. Стандартная методика проверки рабочей гипотезы о принадлежности СВ с именем $У$ к тому или иному ЗР, необходимо произвести выравнивание соответствующего вариационного ряда (ВР), то есть заменить его близким к нему по характеру теоретическим (вероятностным) распределением, имеющим определённое аналитическое выражение [5. - С. 43]. То есть сначала необходимо построить ВР как отображение исходной совокупности СВ $У$ на множество новой переменной $Х$, которая в сущности представляет собой варианты-строки компактной по

отношению к множеству Y таблицей с серединами строк x_i^{cp} и эмпирическими частотами f_i , с которыми эти x_i^{cp} проявляются:

$$\tau: Y \rightarrow X. \quad (62)$$

Полезность отображения (62) состоит в том, что оно осуществляет, как уже говорилось не раз, информационное «сжатие»: вместо числа элементов исходного ряда числом $m = 9$ строится конструкция в виде небольшой таблицы с n строками-вариантами, число которых традиционно оценивается по приближённой формуле Стерджесса [25. - С. 95]: $n \approx 1 + 3,322 \cdot \lg m = 1 + 3,322 \cdot 0,95424 = 4,17 \approx 4$ (интервала: всё же лучше округлять в меньшую сторону, как мы с вами уже убедились). Иначе говоря, мы, строя ВР, заменяем $m = 9$ исходных значений СВ на $n = 4$ значения некоторых середин интервалов x_i^c , содержание (характер) которых мы и намереваемся далее определить. Если минимальная и максимальная величины исходного ряда СВ Y : $y_{min} = 4$ ошибки; $y_{max} = 23$ ошибки, откуда размах выборки $R = y_{max} - y_{min} = 23 - 4 = 19$ (ошибок). Тогда искомый шаг для построения ВР в виде компактной таблицы, начиная от минимального значения анализируемой совокупности Y как $y_{min} = 4$ может быть построен с шагом h :

$$h = \frac{R}{n} = \frac{23 - 4}{4} = 4,75 \text{ (ошибки)}. \quad (63)$$

Если найден шаг h и начальное значение СВ $y_{min} = x_{min} = 4$, то в силу (63) можно строить ВР, который и представлен в табл. 1 (графы-столбцы с номерами 1 – 5).

Итак, эмпирические частоты f_i при $i = 1, n = 4$ найдены: свою основную задачу построенный ВР выполнил. Ну что ж, посчитаем число и таких параметров в графах 6 и 7 той же табл. 1.

Сначала для графы 6 табл. 1 надо оценить значения $P(x_i^{cp})$ – вероятности наступления события x_i^{cp} по формуле (64) из [3. - С, 214]:

Однако, необходимо сделать замечания по заполнению графов табл. 2 с номерами 4 и 5, где $P(x_i^{cp})$ – вероятность наступления события x_i^{cp} :

$$P(x_i) = \frac{\mu^{x_i^{cp}}}{x_i^{cp}!} e^{-\mu}, \quad (64)$$

Суть такого замечания относится к знаменателю функции (64) и заключается в следующем.

Таблица 20

Интервальный вариационный ряд множества X с проверкой на ЗР Пуассона и критерием согласия χ^2 Пирсона

I	$x_i - x_{i+1}$	f_i		x_i^{cp}	$P(x_i^{cp})$ по (6)	$f_i^{теор}$ по (8)	$(f_i - f_i^{теор})$	$(f_i - f_i^{теор})^2$	$(f_i - f_i^{теор})^2 / f_i^{теор}$
1	2	3	4	5	6	7	8	9	10
1	4,00 – 8,75	/	1	6,375	0,0225	2,544	-1,544	2,383	0,937
2	8,75 – 13,50	////	5	11,125	0,1071	12,102	-7,102	50,438	4,168
3	13,50 – 18,25	//	2	15,875	0,0691	7,808	-5,808	33,733	4,320
4	18,25 – 23,00	/	1	20,625	0,0110	1,243	-0,243	0,059	0,047
Σ	-	9	9	-	-				9,472

Мы прежде имели дело с вычислением факториалов «!» для ряда *натуральных* (целых) чисел, однако ряд в графе 2 табл. 1 представлен четырьмя *вещественными* числами. Поиск по Интернету позволил оценить методику вычисления факториалов от вещественных (здесь – десятичных) чисел как с помощью логарифмов с последующим потенцированием, так и, более точное вычисление, с применением гамма-функции. Однако в Интернет-источнике [23] приведена ещё одна методика с предварительным преобразованием вещественного числа в некоторую неправильную дробь – см. систему выражений (65), – с последующим вычислением уже собственно факториала как бы предварительно найденной неправильной дроби, в чём, видимо, и состоит главное назначение приводимого в Интернете алгоритма:

$$6,375! = \left(\frac{51}{8}\right)! = 1468,71097; \quad (65)$$

$$11,125! = \left(\frac{89}{8}\right)! = 54206806,14846;$$

$$15,875! = \left(\frac{127}{8}\right)! = 14744562390571,6;$$

$$20,625! = \left(\frac{165}{8}\right)! = 16221580982654100000.$$

Тогда, располагая значениями знаменателей формулы (65), легко заполнить графу 6 табл. 1.

Теперь же на основе результатов по формуле (65) теоретическую частоту $f_i^{\text{теор}}$ легко найти по формуле (66) с учётом формулы (67):

$$f_i^{\text{теор}} = N \cdot P(x_i^{\text{cp}}), \quad (66)$$

где N есть общее число единиц исходного статистического ряда U :

$$N = \sum_{j=1}^m y_j, \quad (67)$$

Тогда результаты вычислений по формуле (8) могут быть внесены в графу (7) в дополнение к уже построенному нами ВР (1 – 5 графы). Значения теоретических частот $f_i^{\text{теор}}$ после выравнивания ВР в соответствии с ЗР Пуассона закончено. Далее для того, чтобы убедиться в том, что СВ – элементы исходного множества U распределены по ЗР Пуассона, остаётся воспользоваться таблицей χ^2 -критерия Пирсона при заданном нами же уровнем значимости α . Для этого, в свою очередь, надо определить число степеней свободы $df = n - r - 1$, где n – число вариантов построенного ВР; r – число параметров эмпирического распределения, использованных далее для нахождения теоретических частот $f_i^{\text{теор}}$ [5. - С. 46]. При выравнивании по кривой Пуассона для расчёта теоретических частот ранее применялся лишь один параметр эмпирического распределения $\mu = u_{\text{cp}}^{\text{вз}} = x_{\text{cp}}^{\text{вз}}$ из выражения (59). Следовательно, $r = 1$, и $df = 4 - 1 - 1 = 2$. И при заданном нами уровне значимости $\alpha = 0,05$ табличное значение $\chi^2_{\text{табл}}(df=2; \alpha=0,05) = 5,99$.

Остаётся рассчитать $\chi^2_{\text{расч}}$ с помощью последовательного заполнения столбцов 8, 9 и 10 табл. 1 как средства вычисления по формуле критерия Пирсона (31)[5. – С. 48]:

$$\chi^2_{\text{расч}} = \sum_{i=1}^n \frac{(f_i - f_i^{\text{теор}})^2}{f_i^{\text{теор}}},$$

Последняя строка графы 8 табл. 1 содержит число $\chi^2_{расч} = 9,47$.

В итоге нами получен результат для сравнения, который мы и представим в выражении (68):

$$\chi^2_{расч} = 9,47 > \chi^2_{табл}(df = 2; \alpha = 0,05) = 5,99. \quad (68)$$

Так как полученное в расчётах значение параметра Пирсона $\chi^2_{табл} = 9,47$ меньше значения χ^2 -критерия Пирсона для принятых $df = 2$ и $\alpha = 0,05$ по выражению (68), то, следовательно, можно считать расхождения между эмпирическими и теоретическими частотами *неслучайным*, поэтому выдвинутая ранее Э. Демингом гипотеза о близости эмпирического распределения СВ по ЗР Пуассона нами *опровергается*, то есть она неверна [5. - С. 48]. Уж так получилось у нас.

Но ведь действительно Э. Деминг пользовался именно этим законом распределения СВ, который на поверку оказывается несостоятельным. Хотя, данная несостоятельность расходится с табличной на 37% или на 58% (зависит от того, какой показатель принять за 100%: $\chi^2_{расч} = 9,47$ или $\chi^2_{табл}(2; 0,05) = 5,99$). Однако же расходятся показатели выражения (68) далеко «не в разы».

Проверка рабочей гипотезы о ЗР СВ по закону Гаусса (авторская гипотеза). Теперь мы выдвигаем иную гипотезу о ЗР СВ: предположим, что СВ распределена по нормальному закону распределения (НЗР) Гаусса. Однако снова строить ВР из исходной совокупности СВ с именем U на требуется (он уже построен нами ранее в графах 1 – 5 табл. 1; размер шага h и эмпирические частоты f_i уже есть; но стандартное отклонение σ надо будет посчитать). Далее остаётся только подсчитать значения теоретических частот $f_i^{теор}$, а далее – повторить вычисления по столбцам 8, 9 и 10 табл. 1 для той же проверки соответствия принятой гипотезы по критерию Пирсона.

Подсчитаем на основе построенного уже ВР прежде отсутствующий в подобных расчётах $f_i^{теор}$ ещё один параметр, с.к.о. σ . Но для его расчёта

понадобится ещё один параметр из того же ВР – уточнённое взвешенное среднее x_{cp}^{B3} по формуле (1), которое применяется в формуле расчёта σ .

$$x_{cp}^{B3} = \frac{6,375 \cdot 1 + 11,125 \cdot 5 + 15,875 \cdot 2 + 20,625 \cdot 1}{1 + 5 + 2 + 1} = \frac{114,375}{9} = 12,7083.$$

Далее по формуле (69) рассчитаем значение с.к.о.:

$$D = \sigma^2 = \frac{\sum_{i=1}^n (x_i^{cp} - x_{cp}^{B3})^2 \cdot f_i}{\sum_{i=1}^n f_i}. \quad (69)$$

$$D = \frac{(6,375 - 12,708)^2 \cdot 1 + (11,125 - 12,708)^2 \cdot 5 + (15,875 - 12,708)^2 \cdot 2}{9} + \frac{(20,625 - 12,708)^2 \cdot 1}{9} = \frac{135,3750}{9} = 15,0417 \approx 15,04.$$

Тогда искомое с.к.о. σ найдётся как корень из дисперсии:

$$\sigma = \sqrt{D} = \sqrt{\sigma^2} = \sqrt{15,0417} = 3,8784. \quad (70)$$

А также постоянное значение для данного ВР:

$$\text{const} = \frac{N \cdot h}{\sigma} = \frac{9 \cdot 4,75}{3,8784} = 11,0226, \quad (71)$$

где N , кстати, вычисляется уже по-другому, не по Пуассону:

$$N = \sum_{i=1}^n f_i - \text{сумма частот вариантов ВР.}$$

Далее вычисляются вспомогательные параметры - нормированные отклонения

$$t_i = \frac{(x_i^{cp} - x_{cp}^{B3})}{\sigma}. \quad (72)$$

для последующего вычисления вероятности наступления каждого события (здесь x_i^{cp}) по формуле (38), которая приводится в приложениях к учебникам по статистике, например, в [5. - С. 128]:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right).$$

Напомним только, что мы ранее решили сравнить работу критерия согласия Пирсона для округлённых и неокруглённых значений теоретических частот.

Поэтому будущую расчётную табл. 21, видимо, придётся существенно расширять по сравнению с таблицей построения обязательного ВР (графы 1 – 5), дополнив графами 6 - 9:

Таблица 21

Интервальный вариационный ряд множества с проверкой нормального закона распределения Гаусса

i	$x_i - x_{i+1}$	f_i		x_i^{cp}	$x_i^{cp} - x_{cp}^{вз}$	t_i по (14)	$\varphi(t_i)$ Прил.	$f_i^{теор}$ по (10)	
		3	4					5	6
1	4,00 – 8,75	/	1	6,375	- 6,333	- 1,633	0,1057	1,16 \approx 1	
2	8,75 – 13,50	////	5	11,125	- 1,583	- 0,408	0,3668	4,04 \approx 4	
3	13,50 – 18,25	//	2	15,875	3,167	0,817	0,2850	3,14 \approx 3	
4	18,25 - 23,00	/	1	20,625	7,917	2,042	0,0498	0,55 \approx 1	
Σ	-	9	9	-	-	-			

В графе 9 табл. 2 было произведено округление теоретических частот в вещественных числах до целых значений. Поскольку в разных источниках пользуются как вещественными, так и натуральными значениями частот, произведём расчёты и для тех (в графах 3, 5, 7 и 9 – для целых значений теоретических частот), и для других (в графах 4, 6, 8 и 10). А затем сделаем вывод из последующего сравнительного анализа итоговых результатов. Далее следуют расчёты для известного критерия Пирсона по формуле (31) для вычисления значений $\chi^2_{расч}$ (для целых и вещественных значений теоретических частот) в табл. 22:

Таблица 22

Расчёт χ^2 - критерия Пирсона для целых и вещественных $f_i^{теор}$

i	f_i	$f_i^{теор}$		$(f_i - f_i^{теор})$		$(f_i - f_i^{теор})^2$		$(f_i - f_i^{теор})^2 / f_i^{теор}$	
		3	4	5	6	7	8	9	10
1	1	1	1,16	0	- 0,16	0	0,0256	0	0,1772
2	5	4	4,04	1	0,96	1	0,9216	0,250	0,2281
3	2	3	3,14	- 1	1,14	1	1,2990	0,333	0,4137
4	1	1	0,55	0	0,45	0	0,2025	0	0,3682
Суммы:								0,583	1,1872

Осталось лишь ещё раз уточнить число степеней свободы df с тем, чтобы выписать из таблицы Пирсона [5. - С. – 132] граничные значения $\chi^2_{табл}(df, \alpha = 0,05)$. Итак, как уже известно, $df = n - r - 1$. Поскольку при расчёте теоретических частот $f_i^{теор}$ используются два параметра эмпирического распределения ($r = 2$): x_{cp} и σ , то $df = n - r - 1 = n - 3 = 1$. Тогда $\chi^2_{табл}(df = 1; \alpha = 0,05) = 3,84$. Затем воспользуемся тем же неравенством (68) для целых значений теоретических частот в неравенстве (73) и для вещественных значений тех же частот в неравенстве (74):

$$\chi^2_{расч} = 0,583 < \chi^2_{табл}(df = 1; \alpha = 0,05) = 3,84. \quad (73)$$

$$\chi^2_{расч} = 1,187 < \chi^2_{табл}(df = 1; \alpha = 0,05) = 3,84. \quad (74)$$

В обоих случаях – для целых и вещественных значений теоретических частот гипотеза о нормальном распределении подтверждается с достаточной достоверностью $\alpha = 0,05$, хотя (73) свидетельствует в пользу выдвинутой нами гипотезы всё же несколько лучше, чем (74). То есть исследуемое нами распределение случайных величин подчинено не закону распределения Пуассона, но нормальному закону распределения Гаусса. В этом, кстати, легко убедиться визуально, если мысленно повернуть содержание графы 3 табл. 1 на 90 градусов влево: число средних интервалах палочек-слежей окажется больше, в первом и последнем интервалах меньше (больше похоже на «шапку Гаусса», чем на графики ЗР Пуассона).

Отсюда же следует, что формула (59), в отличие от автора [15], должна быть записана как формула (75):

$$ВП, НП = x_{cp}^{вз} \pm 3 \cdot \sqrt{\sigma^2}. \quad (75)$$

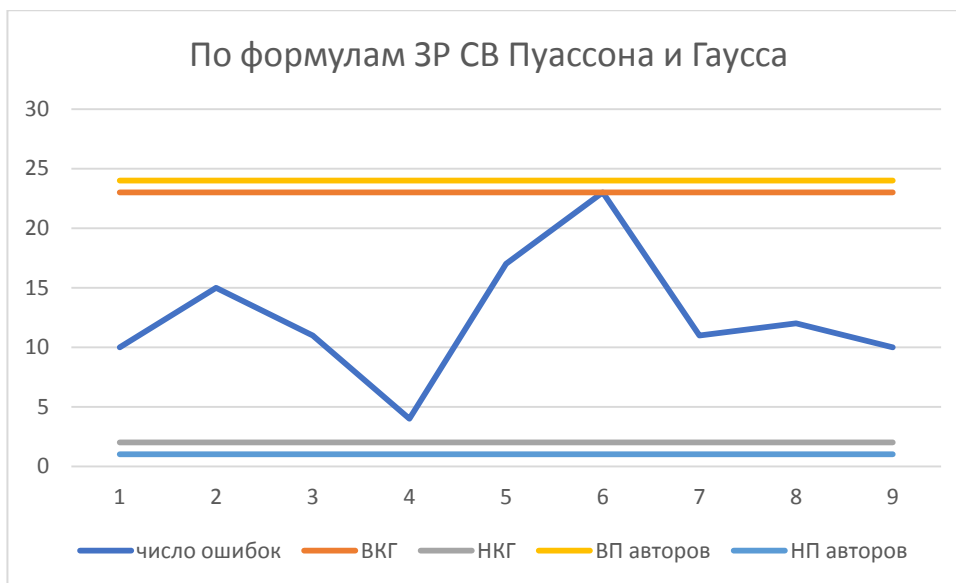
Тогда $ВП = 24,34 \approx 24$ ошибки; $НГ = 1,07 \approx 1$ ошибка (наши расчёты).

$ВП = 23,18 \approx 23$ ошибки; $НГ = 1,92 \approx 2$ ошибки (по Демингу).

Полученные результаты с учётом рис. 14 представим на рис. 15, из которого видно, что в итоге – разница *весьма небольшой*: ВП передвинулся вверх лишь на 1 единицу, НП – также вниз на 1 позицию. То есть мы

установили «коридор для системы» от 1 до 24 ошибок, тогда как из формулы Деминга следует: от 2 до 23 ошибок.

Число ошибок



условные № сотрудников

Рис. 15. Взаимное положение границ и числа ошибок по Демингу[15] и авторским расчётам.

Но содержание рекомендация Э. Деминга для руководителя предприятия остались прежними: если допущенные ошибки каждым специалистом находятся в пределах допустимого, то, во-первых, следует всех поощрить материально, во-вторых, необходимо совершенствовать не квалификацию отдельных специалистов, но существует необходимость совершенствования самой системы производства.

Легко вообразить, как это происходило бы в каком-то ином месте. Местный руководитель, не определяя никаких верхних и нижних контрольных границ, просто, вычислив значение среднего числа ошибок методом подсчёта простого механического среднего (кстати, более точного, чем среднее взвешенное по материалам построенного ВР) и выделил бы тех, кто явно ниже среднего значения по допущенным ошибкам (напомним, эта механическая средняя величина равна $\mu = 12,55$ ошибок). Кто допустил

число ошибок явно меньше значения μ – «идите»; те, кто меньше, но около или равно μ – что называется, «пока идите», а уж кто больше значения μ – «а вас, Штирлиц, я попрошу остаться...».

Однако подход именно Э. Деминга кажется авторам хоть и для них неожиданным, но весьма убедительным для руководителя предприятия.

Какие же в целом можно сделать выводы? Ну, наверное, следующие.

Выводы по разделу. Итак, закон распределения зафиксированных девяти значений ошибок специалистов за какой-то промежуток времени является на проверку нормальным с естественным последующим расчётом обоих допустимых границ по формуле (75), а не по формуле (59), применённой Э. Демингом, характерной для закона распределения той же случайной величины по закону Пуассона. Хотя у Э. Деминга интервал (ВП - НП) = $23 - 2 = 21$ (ошибка), тогда как наш интервал в случае более корректной гипотезы о нормальном распределении ошибок (ВП - НП) = $24 - 1 = 23$ (ошибок). То есть более корректный закон распределения исследуемой случайной величины «раздвинул» границы лишь на 1 единицу вверх для ВП и на 1 единицу вниз для НП. То есть полученной разницей можно вполне практически пренебречь даже в рамках приведённого в [15.- С. 112 – 113] практического примера. Подобные результаты показали проверки и некоторых других примеров из [15].

Хотя, на наш взгляд, следовало бы ограничиться так называемой неполной расчётной проверкой («полупроверкой») прежней гипотезы на основе нашей новой гипотезы, лишь обратившись, например, к [5. - С. 44 - 45], где в частности отмечено следующее: «Если в целом ряде случаев, если ВР представляет собой распределение по дискретному признаку, где по мере увеличения значений признака x частоты резко уменьшаются и где средняя арифметическая ряда равна или близка по значению к дисперсии, то есть $x_{cp} = D = \sigma^2$, то такой ряд можно выровнять по кривой Пуассона...».

И действительно, из формулы (59) для НЗР $x_{cp} = 12,708$; из формулы (69) $\sigma^2 = 15,042$ (разница примерно в 16 – 18%: если брать за 100% - одно или

другое число). То есть, судя по конечным результатам (определение ВП и НП) значения x_{cp} и σ^2 , следует считать достаточно практически близкими. Хотя и кому-то разница на 15 – 18% (в среднем 16,5%) между этими параметрами распределения исследуемой СВ и покажется формально существенной. Так, проведя расчёты по (60) для среднего взвешенного и (69) для дисперсии следовало бы проверить выполнение $x_{cp} \approx \sigma^2$, и в случае их примерной «похожести» - воспользоваться всё же рекомендуемой Э. Демингом формулой (59), характерной именно для ЗР Пуассона: вычислений значительно меньше, а результаты – примерно те же. А методика оценки специалистов и руководства автора [15] действительно впечатляет: поскольку все допустившие ошибки специалисты находятся в существующей «системе» - надлежит всем повысить заработную плату (!).

И, пожалуй, следует упомянуть ещё об одном доводе в пользу априорной гипотезы о выбранном Э. Демингом законе распределения ошибок 9-ти специалистов (ЗР СВ Пуассона, хоть и без выравнивания и проверки по известным критериям согласия). Это, прежде всего, вид графика, приведённого на рис. 19 (стр. 82 настоящего УП) и комментарии к нему при значении $\mu = 9$, где отмечается, что при таком значении μ (и более 9-ти) ЗР СВ Пуассона даже по внешнему виду приближается к нормальному ЗР СВ. Видимо, Э. Демингом поэтому и выбраны именно 9 (!) специалистов. Впрочем, авторы учебного пособия на данном своём умозаключении вовсе и не настаивают...

Да, чисто формально, результат, полученный нами, более *статистически надёжен*. Однако надо учесть и то, что, судя по примечаниям к переводу 10-го издания книги [15], её автор много и плодотворно сотрудничал с широким кругом специалистов, в том числе и носителями знаний по точным наукам. А такие мэтры, как в 50-х годах ХХ века токийский профессор Кауро Исикава с его «рыбьим остовом», столь необходимой моделью для описания как альтернатив, так и критериев в многокритериальных задачах принятия решений, которая использовалась в

полной мере и авторами [24. - С. 20 – 26]. И это при том, что проф. Кауро Исикава всегда считал себя лишь скромным учеником Эдвардса Деминга как в вопросах собственно контроля качества, так и в более широком смысле этого понятия – управления качеством или менеджмента качества.

Ну что ж, будем учиться этому у него и мы.

КРАТКИЕ ВЫВОДЫ

Итак, мы подробнее познакомились с нормальным законом распределения (НЗР) Гаусса, который, в отличие от *дискретных* ЗР Бернулли, биномиального и Пуассона, является *непрерывным*. Однако, дискретные законы можно (а иногда и нужно) приближать к НЗР. Поэтому целесообразно их использовать именно как дискретные законы распределения СВ, в частности биномиальный и Пуассона.

Разумеется, акцент был сделан авторами данного УП на работы в основном отечественных исследователей. Однако, для того, чтобы студенты имели хотя бы общее представление о научных разработках зарубежных учёных, авторы сознательно расширили библиографический список по следующим семи направлениям (в порядке нумерации источников): нормальное распределение [41]–[49]; формы Бернулли [50]–[51]; распределение Пуассона [52]–[55]; биномиальное распределение [56]–[61]; традиционные проблемы по управлению качеством [62]–[63]; современные обзорные статьи по теме УП [64]–[65]; работы по проблемам всеобщего менеджмента качества (TQM) изложены в [66]–[71].

ЗАДАНИЯ НА САМОСТОЯТЕЛЬНУЮ РАБОТУ

Для контроля качества выпускаемой продукции организовали отбор готовых изделий в 10 партиях по 100 штук изделий. Затем подсчитывалось число несоответствующих единиц продукции. Такие исходные данные для

выбора метода и процесса решения приведены (как в табл. 1) по вариантам от 1-го до 20-го. Номера выбирать согласно списку учебной группы. Если же в группе более 20-ти студентов, то следует свой вариант выбирать, например, так: по списку студент значится под № 23: значит, он выбирает №3 и т.д.

Вариант 1 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	2	3	0	3	12	4	1	2	15	3

Вариант 2 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	3	4	3	14	4	1	2	15	3

Вариант 3 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	3	2	3	14	4	1	2	12	3

Вариант 4 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	0	5	13	4	2	2	15	3

Вариант 5 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	3	3	12	4	1	2	11	3

Вариант 6 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	0	3	7	4	3	2	15	3

Вариант 7 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	4	2	2	3	11	4	1	2	15	3

Вариант 8 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	0	3	12	6	7	2	15	3

Вариант 9 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	2	4	2	3	12	4	1	2	15	3

Вариант 10 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	2	4	0	3	13	4	1	2	12	3

Вариант 11 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	5	6	13	4	1	2	15	3

Вариант 12 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	5	63	14	4	1	2	13	3

Вариант 13 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	3	3	12	4	5	2	15	3

Вариант 14 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	3	3	2	3	12	4	1	2	16	3

Вариант 15 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	6	5	3	14	4	1	2	15	3

Вариант 16 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	0	3	14	7	8	2	15	3

Вариант 17 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	2	3	1	3	14	4	1	2	11	3

Вариант 18 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	3	4	4	14	4	1	2	15	3

Вариант 19 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	0	3	12	4	6	3	15	3

Вариант 20 – Результаты контроля качества в 10 партиях по 100 изделий в каждой.

№ партии	1	2	3	4	5	6	7	8	9	10
Объём партии, штук	100	100	100	100	100	100	100	100	100	100
Число брак. изд., штук	1	2	8	3	14	4	6	2	15	3

При расчётах по номерам вариантов необходимо произвести следующие работы:

Все расчёты и графики (можно «от руки») проводить исключительно в РАБОЧИХ ТЕТРАДЯХ, но не на отдельных листах, пусть даже в машинном виде. Далее следует устная краткая защита полученных результатов.

То есть от студента требуется показать свои умения 1) проводить необходимые расчёты, 2) отображать полученные результаты графически и 3) кратко обосновать устно свои умозаключения.

Порядок выполнения домашнего контрольного задания следующий.

1. Выдвинуть предварительную гипотезу о биномиальном распределении случайной величины (случаев забракованных изделий из каждой контрольной партии по 100 штук) и рассчитать среднее значение бракованных изделий, оценить величину с.к.о. и определить верхнюю и нижнюю контрольные границы. Если же какая-либо партия окажется вами забракованной, вывести её из последующего рассмотрения и определить *новое значение* среднего числа бракованных изделий, а также новое положение верхней и нижней контрольных границ.

Если же выведенных из рассмотрения партий вами выявлено не будет, приступить ко второму пункту задания.

2. Проверить новую таблицу на то, что число браков в оставшихся партиях подчинены нормальному закону распределения Гаусса. Для этого необходимо построить вариационный ряд на число вариантов $nvr \geq 4$ с попутным определением эмпирических частот и продолжить его до получения теоретических частот.

Замечание: старайтесь построить вариационный ряд на именно ЧЕТЫРЕ варианта, так как число степеней свободы для входа в таблицу Пирсона $df = nvr - 3$; при этом должно быть соблюдено следующее, общем случае, нестрогое неравенство: $df \geq 1$.

Располагая рядами эмпирических и теоретических частот, применить к ним критерии согласия Пирсона и Романовского (дополнительно, для

желающих – и Колмогорова). Если же обнаружится, что при заданном уровне значимости α (рекомендуется брать табличные значения при $\alpha = 0,05$ или 95%надёжности) $\chi^2_{табл} \leq \chi^2_{расч}$, то ваша гипотеза о нормальном распределении выявленных случаев брака – вполне СПРАВЕДЛИВА. Если же $\chi^2_{табл} > \chi^2_{расч}$, то случаи брака как исследуемая случайная величина не удовлетворяет требованиям НЗР Гаусса. Тогда попытайтесь установить: на сколько именно.

Пример. Пусть $\chi^2_{табл} (df = 1; \alpha = 0,05) = 3,84$, а $\chi^2_{расч} = 4,08$. Тогда превышение над табличным значением легко найти из пропорции:

3,84 - это 100%;

4,08 – это x%.

Откуда $x = (4,08 \cdot 100\%) / 3,84 = 106,25\%$. Или найденное нами значение $\chi^2_{расч} = 4,08$ превышает максимальное табличное значение всего на $106,25 - 100 = 6,25$ (%). А много это или мало, предлагаем пользоваться модифицированной авторами таблицы 23 проф. В.А. Ядова [12. – С. 62]:

Таблица 23

Таблица характера надёжности и ошибки

Характер надёжности	Ошибка, %
максимальный (ошибка – минимальна)	до 3
обыкновенный	3 – 10
приближённый	10 - 20
ориентировочный	20 - 40
прикидочный	более 40

3. Рассчитать среднее значение, верхнюю и нижнюю контрольные границы и сделать выводы с позиций биномиального закона распределения.

4. Выдвинуть нулевые гипотезы о распределении вариантных значений случайных величин последовательно по законам распределения Пуассона и Гаусса.

5. Сделать выводы о результатах применения к предложенным в вариантах случайных значений долей брака в контрольных партиях, рассчитанных с позиций гипотез о распределении случайных величин-браков с позиций биномиального распределения, законов распределения Пуассона и

Гаусса. Обосновать свои выводы по выполненному варианту и устно (по возможности – кратко и по существу) сообщить их преподавателю.

На этом выполнение домашнего контрольного задания следует считать завершённым.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Управление качеством: Учебник для вузов / С.Д. Ильенкова, В.С. Мхитарян и др.; Под ред. С.Д. Ильенковой. – 2-е изд., перераб. И доп. – М.: ЮНИТИ-ДАНА, 2004. – 334 с.
2. Элементарное введение в теорию вероятностей. Б.В. Гнеденко, А.Я. Хинчин. Издание восьмое, исправленное. – М.: Наука, Гл. ред. Физ.-мат. Лит., 1976. – 168 с.
3. Тейлор Дж. Введение в теорию ошибок. Пер. с англ. – М.: Мир, 1985. – 272 с., ил.
4. Солонин С.И. Метод контрольных карт. Электронное текстовое издание. – Екатеринбург: ЦНОТ ИТОО УрФУ, 2014. – 213 с.
5. Громыко Г.Л. Общая теория статистики: Практикум. – М.: ИНФРА-М, 1999. – 139 с.
6. Ефимова М.Р., Петрова Е.В., Румянцев В.Н. Общая теория статистики: Учебник. Изд. 2-е, испр. и доп. – М.: ИНФРА-М, 2000. – 416 с.
7. Лютикас В.С. Школьнику о теории вероятностей. Учеб. Пособие по факультативному курсу дл учащихся 8 – 10 кл. – М.: «Просвещение», 1976. – 112 с.
8. Эконометрика: Учебник / Под ред. И.И. Елисеевой. – М.: Финансы и статистика, 2004. – 344 с.: ил.
9. Клейнер Г. Б., Тамбовцев В. Л., Качалов Р. М. Предприятия в нестабильной экономической среде: риски, стратегия, безопасность. – Москва: Экономика 2017. - 445с.
10. Интернет-ресурс: Правило трёх сигм – 18 Декабря 2012 – Примеры решения задач / reshim.su/blog/pravilo_trech...
11. Интернет-ресурс: Правило трёх сигм – 18 Декабря 2012 – Примеры решения задач / reshim.su/blog/pravilo_trech...
12. Ядов В.А. Социологическое исследование: методология, программа, методы. – М.: Наука, 1987. – 248 с.

13. Дубровин В.Т. Решебник по элементарной теории вероятностей: учебное пособие / В.Т. Дубровин, В.С. Желтухин, В.Ю. Чебакова. – Казань: Изд-во Казан. ун-та, 2015. – с. 118.
14. Карасев А.И. Теория вероятностей и математическая статистика. Учебник для экономических специальностей вузов. Изд. 3-е, перераб. и доп. – М.: «Статистика», 1977. – 280 с.
15. Деминг Э. Выход из кризиса: Новая парадигма управления людьми, системами и процессами / Эдвардс Деминг; Пер. с англ. – 10-е изд. – М.: Альпина Паблишер, 2019. – 417 с.
16. Интернет-ресурс: Закон распределения Релея. studopedia.ru/3_19722_... (обращение 04.12.2020 г.).
17. Интернет-ресурс: Закон распределения Вейбулла. vuzlit.ru/2112606/raspredelenie_... (обращение 04.12.2020 г.).
18. Интернет-ресурс: Равномерный закон распределения. studopedia.ru/4_32712_... (обращение 04.12.2020 г.).
19. Бурлаков А., Турбулентность экономических процессов: теоретические аспекты // Вопросы экономки. – 2018. - №11. С. 20.
20. Павлова Ю.Н. Управление предприятием в условиях нестабильности внешней среды // Социально-экономические явления и процессы. – 2016. – т. 10. - № 2. – С. 84.
21. Интернет-ресурс: «Проверка гипотезы о распределении Пуассона. Примеры решения задач. matem96.ru/orimer/primer_terverb... (обращение 04.02.2021 г.).
22. Интернет-ресурс: «Правило трёх сигм – 18 Декабря 2012 – Примеры решения задач / reshim.su/blog/pravilo_trech... (обращение 06.11.2020 г.).
23. Интернет-ресурс: «Факториал числа. – Калькулятор онлайн. kontrolnaya-rabota.ru/s/teoriya-... (обращение 02.02.2021 г.).
24. Шихалёв А.М., Хафизов И.И. Многокритериальные задачи принятия решений в управлении качеством: учебное пособие / А.М.

Шихалёв, И.И. Хафизов. – Казань: Издательство Казанского университета, 2020. – 92 с.

25. Общая теория статистики: Учебник / Т.В. Рябушкин, М.Р. Ефимова, И.М. Ипатова, Н.И. Яковлева. – М.: Финансы и статистика, 1981. – 279 с., ил.

26. Шрейдер Ю.А. Равенство, сходство, порядок: М.: Наука, Гл. редакция физ. – мат. литературы, 1971. – с. 255.

27. Интернет-ресурс: «Законы распределения случайных величин»: lektsii.org/3-122119.htm... (обращение 29.04.2021 г.).

28. Интернет-ресурс: «Законы распределения случайных величин – ПРИКЛАДНАЯ МАТЕМАТИКА»: studme.org/140788/matematika... (обращение 29.04.2021 г.).

29. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. пособие для студ. вузов / В. Е. Гмурман. - М.: Высш. шк., 2004. - 480 с.

31. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учеб. пособие для студ. вузов / В. Е. Гмурман. -- М.: Высш. шк., 2004. - 400 с.

32. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. Пособие для бакалавров, студентов вузов / В.Е. Гмурман. – 12-е изд. – М.: Юрайт, 2013. – 478 с.

33. Павлушков И.В. Основы высшей математики и математической статистики: учебник для вузов / Павлушков И.В., Розовский Л.В., Капulyцевич Л.Е. и др. – 2-е изд., испр. – М.: ГЭОТАР-Медиа, 2005. – 424 с.

34. Теория вероятностей и математическая статистика: учебник для бакалавров / В.Н. Калинина. – 2-е изд., перераб. и доп. – М.: Юрайт, 2015. – 472 с.

35. Белько И.В. Теория вероятностей и математическая статистика. Примеры и задачи: учеб. пособие для вузов / И. В. Белько, Г. П. Свирид; под ред. К. К. Кузьмина. -- Мн.: Новое знание, 2002. -- 250 с.

36. Ефимова М.Р., Бычкова С.Г. Практикум по социальной статистике:

Учеб. пособие / Под ред. М.Р. Ефимовой. – М.: Финансы и статистика, 2005. – 448с.

37. Ефимова М.Р., Бычкова С.Г. Социальная статистика: Учеб. пособие / Под ред. М.Р. Ефимовой. – М.: Финансы и статистика, 2005. – 560с.

38. Октябрьский П.Я. Статистика: Учебник. – М.: ТК Велби, Изд-во Проспект, 2003. – 328с.

39. Статистика: Учеб. пособие / А.В. Багат, М.М. Конкина, В.М. Симчера и др.; Под ред. В.М. Симчеры. – М.: Финансы и статистика, 2005. – 368с.

40. Статистика: Учеб. пособие / Харченко Л.П., Ионин В.Г. и др.; Под ред. канд. экон. наук В.Г. Ионина.– Изд. 2-е перераб. и доп. - М.: ИНФРА-М, 2005. – 384с.

41. Bernardo, José M., Smith, Adrian F. M. (2009). Bayesian Theory. John Wiley & Sons. 608 p.

42. Bryc, Wlodzimierz (1995). The Normal Distribution: Characterizations with Applications. Springer-Verlag. 132 p.

43. Geary R.C. (1936). The distribution of the "Student's" ratio for the non-normal samples". Supplement to the Journal of the Royal Statistical Society, 3 (2); pp. 178–184.

44. Lukacs, Eugene (1942). "A Characterization of the Normal Distribution". Annals of Mathematical Statistics. 13 (1); pp. 91–93.

45. Das, Abhranil, Wilson S Geisler (2021). "A method to integrate and classify normal distributions". URL: <https://arxiv.org/abs/2012.14331v6>.

46. Quine, M.P. (1993). "On three characterisations of the normal distribution". Probability and Mathematical Statistics. 14 (2); 257–263.

47. Pearson, Karl (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space". Philosophical Magazine. 6. 2 (11); pp. 559–572.

48. Pearson, Karl (1905). "Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson'. A rejoinder". Biometrika. 4 (1); pp. 169–212.

49. Stigler, Stephen M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Harvard University Press. 400 p.
50. Bertsekas, Dimitri P., Tsitsiklis, John N. (2002). *Introduction to Probability*. Belmont, Mass.: Athena Scientific. 413 p.
51. McCullagh, Peter; Nelder, John (1989). *Generalized Linear Models*, volume 37 of *Monographs on Statistics and Applied Probability*. Chapman and Hall/CRC Press, London, 2 ed. 532 p.
52. Haight, Frank A. (1967). *Handbook of the Poisson Distribution*. New York, NY, USA: John Wiley & Sons. 168 p.
53. Johnson, Norman L.; Kemp, Adrienne W.; Kotz, Samuel (2005). "Poisson Distribution", *Univariate Discrete Distributions* (3rd ed.), New York: John Wiley & Sons, pp. 156–207.
54. Ugarte, Maria Dolores; Militino, Ana F.; Arnholt, Alan T. (2016). *Probability and Statistics with R* (2nd ed.). Boca Raton, USA: CRC Press. 726 p.
55. Short, Michael (2013), "Improved Inequalities for the Poisson and Binomial Distribution and Upper Tail Quantile Functions", *ISRN Probability and Statistics*, 2013, article ID 412958, 6 p. DOI:10.1155/2013/412958.
56. Feller, W. (1968). *An Introduction to Probability Theory and Its Applications* (3rd ed.). New York: Wiley. 526 p.
57. Wadsworth, G. P. (1960). *Introduction to Probability and Random Variables*. New York: McGraw-Hill. 292 p.
58. Jowett, G. H. (1963). "The Relationship Between the Binomial and F Distributions". *Journal of the Royal Statistical Society D*. 13 (1); pp. 55–57.
59. Arratia, R., Gordon, L. (1989). Tutorial on large deviations for the binomial distribution. *Bulletin of Mathematical Biology*, 51, pp. 125–131.
60. André Nicolas (2019). Finding mode in Binomial distribution. URL: <https://math.stackexchange.com/q/117940>.
61. D. Ahle, Thomas (2021). Sharp and Simple Bounds for the raw Moments of the Binomial and Poisson Distributions. URL: <https://arxiv.org/abs/2103.17027>.

62. Crosby, Ph. Let's Talk Quality: 96 questions you always wanted to ask Phil Crosby. McGraw-Hill Publishing, New York, 1989. 208 p.

63. Joseph M. Juran, A. Blanton Godfrey. Juran's Quality Handbook, Fifth Edition, McGraw-Hill, New York, 1999. 1872 p.

64. Wawak, S., Rogala, P. and Dahlgaard-Park, S.M. (2020). "Research trends in quality management in years 2000-2019", *International Journal of Quality and Service Sciences*, ahead-of-print. <https://doi.org/10.1108/IJQSS-12-2019-0133>.

65. Mohammad Reza Sadeghi, Moghadam, Hossein Safari, Narjes Yousefi (2019). Clustering quality management models and methods: systematic literature review and text-mining analysis approach, *Total Quality Management & Business Excellence*, 32(1); pp. 1-24. URL: DOI: 10.1080/14783363.2018.1540927.

66. C. Siyun, S. Wen, S. Simin and Z. Lanbo (2013). "The conceptions, principles and systems for automotive materials purchasing quality decision and management based on total quality management theory," 6th International Conference on Information Management, Innovation Management and Industrial Engineering, pp. 117-121, URL:<https://ieeexplore.ieee.org/document/6702889>. DOI: 10.1109/ICIM.2013.6702889.

67. I. Nikolova-Jahn (2018). "Conditions for Applying the Principles of Total Quality Management to the Requirements of the Fourth Industrial Revolution," 2018 International Conference on High Technology for Sustainable Development (HiTech), pp. 1-4, URL: <https://ieeexplore.ieee.org/document/8566296>. DOI: 10.1109/HiTech.2018.8566296.

68. P. B. Sakthivel, R. Raju Pages. Conceptualizing total quality management in engineering education and developing a TQM educational excellence model/*Total Quality Management & Business Excellence* Volume 17, 2006 - Issue 7, pp. 913-934, URL: <https://doi.org/10.1080/14783360600595476>.

69. Gopal K. Kanji. Total quality management and statistical understanding/*Total Quality Management* Volume 5, 1994 - Issue 3, pp. 105-114. URL: <https://doi.org/10.1080/09544129400000030>.

70. Gianni Campatelli, Paolo Citti, Andrea Meneghin. Development of a simplified approach based on the EFQM model and Six Sigma for the implementation of TQM principles in a university administration / Total Quality Management & Business Excellence, Volume 22, 2011 - Issue 7: Business Excellence and Quality Management Systems, pp. 691-704. URL: <https://doi.org/10.1080/14783363.2011.585755>.

71. Namish Mehta, Prakash Verma. Total quality management implementation in engineering education in India: an interpretive structural modelling approach / Total Quality Management & Business Excellence, Volume 25, 2014 - Issue 1-2, pp. 124-140, URL: <https://doi.org/10.1080/14783363.2013.791113>.

Учебное издание

Шихалёв Анатолий Михайллович,

Воронцов Дмитрий Петрович

**КОНТРОЛЬ КАЧЕСТВА ИЗДЕЛИЯ НА ОСНОВЕ
СТАТИСТИЧЕСКИХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ**

Дизайн обложки

М.А. Ахметов

Подписано в печать ____ . ____ .2021.

Бумага офсетная. Печать цифровая.

Формат 60x84 1/16. Гарнитура «TimesNewRoman». Усл. печ. л. .

Тираж экз. Заказ

Отпечатано с готового оригинал-макета

в типографии Издательства Казанского университета

420008, г. Казань, ул. Профессора Нужина, 1/37

тел. (843) 233-73-59, 233-73-28