Министерство образования и науки Российской Федерации
НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

С.В. НЕДЕЛЬКО, В.М. НЕДЕЛЬКО, Г.Н. МИРЕНКОВА

ТИПОВЫЕ ЗАДАЧИ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Утверждено Редакционно-издательским советом
университета в качестве учебного пособия

Новосибирск
2013

УДК 519.2(076.5)

Рецензенты:
канд. физ.-мат. наук, доцент *Э.Г. Соснина*
канд. физ.-мат. наук, доцент *А.П. Ковалевский*

Работа подготовлена кафедрой высшей математики
для студентов II курса РЭФ

Пособие содержит изложение методов решения ряда базовых задач математической статистики, а также задания для самостоятельной работы.

© Новосибирский государственный
технический университет, 2013 г.

Оглавление

1. ОСНОВНЫЕ ПОНЯТИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В настоящем разделе вводятся основные используемые понятия из теории вероятностей и математической статистики. Приведённые определения включены как справочный материал, и их рассмотрение не заменяет изучение теоретического курса.

Для чтения пособия желательно знакомство с теорией вероятностей и математической статистикой в рамках начальных курсов, например [1, 2, 4, 5]. Допустимо, если при первом прочтении часть материала данного параграфа останется непонятной. Это не препятствует изучению последующих разделов.

Базовым понятием теории вероятностей является *вероятностное пространство*, которое включает множество элементарных исходов Ω, σ-алгебру Λ событий и вероятностную меру P.

В качестве *множества элементарных исходов* может выступать любое конечное, счетное или континуальное множество Ω.

Алгебра событий есть некоторое множество, состоящее из подмножеств множества Ω и удовлетворяющее определённым условиям, которые мы приводить не будем.

Функция $P : \Lambda \to R$ называется *вероятностной мерой*, если:

1) $\forall A \in \Lambda,\ \ P(A) \geq 0$,
2) $P(\varnothing) = 0$, $P(\Omega) = 1$
3) Для попарно несовместных (непересекающихся)

$$A_i \in \Lambda,\ \ P\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Случайной величиной $X = X(\omega)$ называется измеримая функция $X : \Omega \to R$. Измеримость здесь означает существование функции распределения $F(x) = P(X(\omega) < x)$.

Для пояснения содержательного смысла введённых понятий рассмотрим пример. Предположим, что исследуется концентрация яда в железах змей некоторого вида. Для этого ловится определённое количество змей и изучается их яд.

В этом случае множество всех змей заданного вида можно отождествить с множеством Ω, а конкретную пойманную змею с элементарным исходом ω. Факт поимки определённой змеи является случайным событием. Событие – это подмножество исходов, поэтому рассматриваем подмножество, состоящее из одной змеи. Считаем, что для каждой змеи существует (определена) вероятность быть пойманной.

Поскольку концентрация яда – это функция змеи, она так же случайна. Для любого числа x можно определить вероятность того, что концентрация яда в железах случайной пойманной змеи не превзойдёт x. Эта вероятность, по определению, есть значение функции распределения в точке x.

Важным моментом является то, что Ω – это множество не только реально существующих змей, но и всех змей, которые, по нашим представлениям, могли бы существовать. Такое соглашение нужно для того, чтобы иметь возможность оперировать с непрерывными случайными величинами. В противном случае, если рассматривать только существующих змей, то множество Ω будет конечным, а все функции на этом множестве, в том числе случайные величины, — дискретными.

При пользовании данным пособием читатель может под случайной величиной понимать просто некоторую переменную X, для которой определена функция распределения. При этом под *функцией распределения* достаточно понимать некоторую функцию $F(x)$, такую что:

1) $F(-\infty)=0$;
2) $F(\infty)=1$;
3) $\forall a \le b,\ P(a \le X < b)=F(b)-F(a)\ge 0$;
4) $\forall a,\ F(a-0)=F(a)$.

Величина $P(a \le X < b)$ есть вероятность[1] попадания случайной величины X в интервал $[a,b)$.

Для системы случайных величин $X_1,\ldots,X_n$ определена совместная функция распределения $F(x_1,\ldots,x_n)=P(X_1<x_1,\ldots,X_n<x_n)$.

В математической статистике важную роль играют независимые случайные величины. Случайные величины называются *независимыми*, если их совместная функция распределения удовлетворяет условию $F(x_1,\ldots,x_n)=F_1(x_1)\cdot\ldots\cdot F_n(x_n)$.

Под *выборкой* понимается набор значений $(x_1,\ldots,x_n)$, полученных как n независимых *реализаций* случайной величины X. Число n называется *объёмом выборки*.

Например, случайной величиной может являться измеряемое значение некоторой физической величины, тогда в качестве выборки будут выступать результаты n независимых измерений.

В историческом значении термин «выборка» подразумевает выборку из так называемой *генеральной совокупности*. В примере со змеями генеральной совокупностью является множество всех змей, которых мы потенциально можем поймать, а выборку составляют фактически пойманные змеи.

Заметим, что термин «генеральная совокупность» для нас не актуален, т.е. мы его не будем в дальнейшем использовать. Для простоты можно отождествлять генеральную совокупность с множеством

[1] В рамках пособия свойство 3) можно использовать вместо определения вероятности.

элементарных исходов. В примере со змеями можно было бы генеральной совокупностью назвать множество реально существующих змей, но вряд ли такое понятие полезно для решения задач статистики. Дело в том, что фактически выборка производится, конечно, из множества реально существующих змей, но, поскольку мы не можем это множество описать и задать, приходится идеализированно считать, что выборка производится из множества «возможных» змей. Кроме того, сама совокупность реально существующих змей фактически является случайной выборкой из множества змей, которые потенциально могут появиться, поскольку рождение конкретной змеи — результат комбинации случайных факторов (полученный генетический материал, внешние условия).

Итак, под выборкой мы будем понимать просто набор случайных значений (реализаций) случайной величины и называть его *выборкой из распределения*.

Под набором мы здесь понимаем упорядоченную последовательность. Это может показаться неожиданным, поскольку на практике порядок значений в выборке, как правило, не важен. Например, если мы перемешаем пойманных змей в клетке, либо переставим результаты измерений свойств их яда — выводы не должны измениться. Тем не менее, выборку мы должны записывать как последовательность значений, а не как множество.

В математической статистике рассматриваются задачи, обратные по отношению к задачам теории вероятностей. Если в задачах теории вероятностей по известным распределениям вычисляются вероятности тех или иных событий, то в математической статистике по фактам наступления некоторых событий делаются выводы о распределениях.

Исходно имеется выборка, полученная в результате n-кратного повторения эксперимента. Распределение $F(x)$, которому подчиняются наблюдения, неизвестно. Требуется на основе информации, заложенной в выборке, насколько это возможно, восстановить закон распределения $F(x)$. Достоверно восстановить распределение, имея лишь n реализаций случайной величины, невозможно, и любые выводы в этом случае будут носить вероятностный характер.

Основные задачи математической статистики следующие:

1) восстановление закона распределения непараметрическими методами;

2) оценивание параметров распределения (параметрические методы);

3) проверка состоятельности гипотезы о законе распределения.

Для решения поставленных задач математическая статистика располагает рядом процедур и методов.

Непараметрический подход включает такие способы восстановления распределения, как построение эмпирической функции распределения и гистограммы, а также множество методов (потенциальных функций и др.), не рассматриваемых в данном пособии.

Представителями параметрических методов являются методы моментов и максимального правдоподобия для оценки параметров распределения. Некоторые другие методы читатель может найти в литературе. Также будет рассмотрено интервальное оценивание параметров распределения.

2. ЭМПИРИЧЕСКАЯ ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ. ГИСТОГРАММА

Эмпирическая функция распределения

На практике (истинная) функция распределения $F(x)$ случайной величины X является полностью или частично неизвестной. На основе информации, заложенной в выборке, можно построить приближение для неизвестной функции распределения.

Эмпирической функцией распределения называется функция $\tilde{F}(x) = \frac{m(x)}{n}$, где $m(x)$ – количество значений из выборки $(x_1, \ldots, x_n)$, меньших x, (т.е. число выборочных точек, расположенных на числовой оси левее заданного числа x).

Эмпирическая функция распределения $\tilde{F}(x)$ определяет частоту события, в то время как функция распределения $F(x) = P(X < x)$ определяет вероятность этого события.

Значения эмпирической функции распределения $\tilde{F}(x)$ изменяются от нуля до единицы, так же, как и у $F(x)$.

Построение эмпирической функции распределения является простейшим методом обработки наблюдаемых значений случайной величины.

По теореме Бернулли при большом числе испытаний частота события $X < x$ сходится по вероятности к вероятности этого события. Следовательно, при любом заданном x функция $\tilde{F}(x)$ сходится по вероятности к функции распределения. Справедливо и более сильное утверждение, а именно, теорема Гливенко–Кантелли.

Для построения графика функции $\tilde{F}(x)$ достаточно изобразить выборочные значения на числовой оси ox и нарисовать ступенчатую функцию, возрастающую в каждой точке выборки на $\frac{1}{n}$. В наименьшей выборочной точке и левее нее $\tilde{F}(x)$ равна нулю, поскольку отсутствуют еще меньшие выборочные точки. В наибольшей выборочной точке эмпирическая функция распределения делает последний скачок и достигает правее нее единицы. Этот способ построения $\tilde{F}(x)$ применим, когда все выборочные точки различны. Если же некоторое значение встречается в выборке r раз, то скачок в этой точке увеличивается также в r раз и составляет $\frac{r}{n}$.

Эмпирическая функция распределения в точках разрыва непрерывна слева, что показывается стрелками на «ступенях».

Функция $\tilde{F}(x)$ является дискретной случайной функцией, т.к. делает скачки в случайных точках, определяемых выборочными значениями.

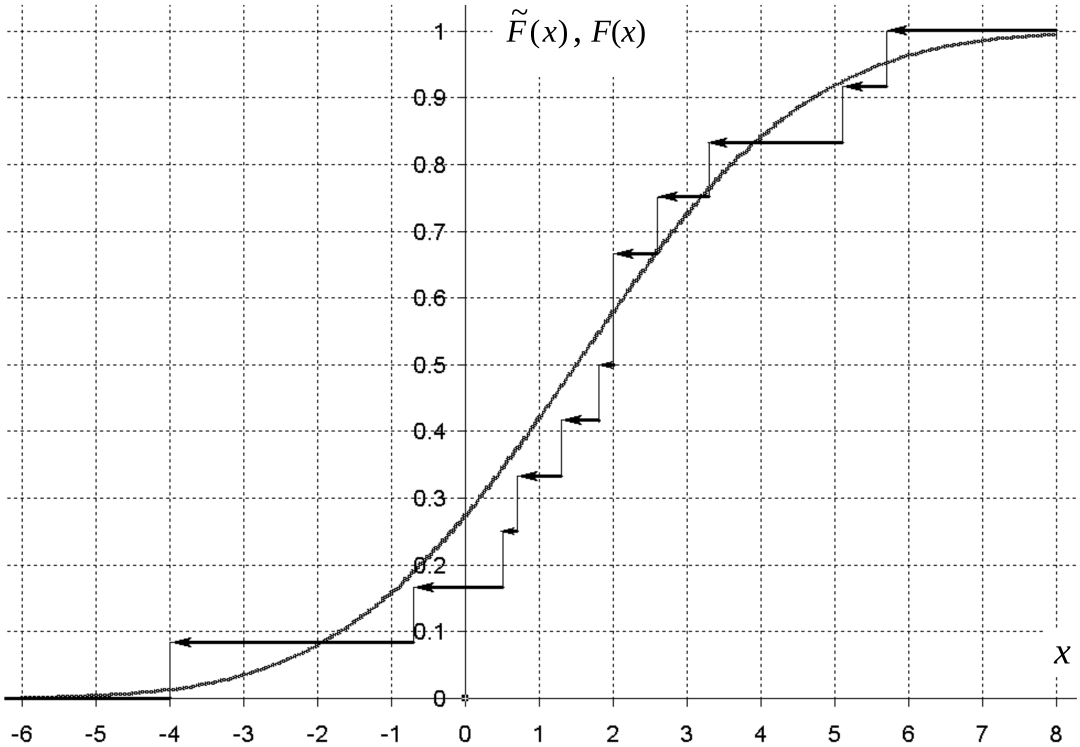

Рис. 1. Эмпирическая функция распределения для выборки из нормального распределения с параметрами m = 1,5 и σ = 2,5 и функция распределения

В первом задании практической работы 1 нужно построить график эмпирической функции распределения для данных из таблицы 4. Здесь и далее расчеты, графики и таблицы приводятся для выборки из варианта 0 таблиц 4 и 5.

На рис. 1 изображена эмпирическая функция распределения для выборки 1,8; 2; 3,3; 2,6; 1,3; –4; 0,5; 0,7; –0,7; 5,1; 5,7; 2 объема 12. Видно, например, что $\tilde{F}(3) = \frac{9}{12} = \frac{3}{4}$, поскольку левее точки 3 расположено 9 выборочных точек из 12.

Во всех точках выборки, кроме точки 2, $\widetilde{F}(x)$ возрастает на 1/12. В точку 2 попало две выборочные точки, поэтому скачок в ней вдвое больше, чем в точках, вошедших в выборку однократно, и равен 2/12=1/6.

Для сравнения на рис. 1 приведена функция распределения $F(x)$ нормально распределённой случайной величины с математическим ожиданием m = 1,5 и среднеквадратическим (стандартным) отклонением σ = 2,5 (параметры распределения указаны в последних колонках таблицы). Выборка является реализацией этой случайной величины.

Гистограмма

Помимо приближения для функции распределения, по выборке можно построить приближения для неизвестной плотности распределения. Простейшим из таких приближений является гистограмма, которую можно использовать как самостоятельно, так и на предварительных этапах обработки данных. Восстанавливать распределение с помощью гистограммы можно, если число наблюдений достаточно велико (на практике требуется не менее пятидесяти выборочных точек). Если число наблюдений невелико, то лучше использовать иные методы обработки данных, например, построить эмпирическую функцию распределения.

Для построения гистограммы в выборке $(x_1,...,x_n)$ объема n находятся наименьшее $x_{\min}$ и наибольшее $x_{\max}$ наблюдаемые значения.

Затем диапазон $[x_{\min}, x_{\max}]$ наблюдаемых значений разбивается на несколько интервалов, которые могут быть как одинаковыми, так и разными по длине. Далее будут рассматриваться только одинаковые интервалы $[s_i, s_{i+1})$ длины $\Delta = s_{i+1} - s_i = \frac{x_{\max} - x_{\min}}{l}$, где l есть число интервалов. В соответствии с обозначениями $s_1 = x_{\min}$, $s_{l+1} = x_{\max}$.

Для определенности считаем, что левые границы строго включены в интервалы, а правые нестрого. Если значение из выборки совпадает с левой границей, то оно входит в интервал, если совпадает с правой

Таблица 1

i	1	2	3	4	5	6	7	8	9	10
(s_i, s_{i+1})	[10,14)	[14,18)	[18,22)	[22,26)	[26,30)	[30,34)	[34,38)	[38,42)	[42,46)	[46,50]
n_i	6	7	11	17	19	18	13	5	3	1
$\widetilde{p}_i$	0.06	0.07	0.11	0.17	0.19	0.18	0.13	0.05	0.03	0.01
$\frac{\widetilde{p}_i}{\Delta x_i}$	0.015	0.0175	0.0275	0.0425	0.0475	0.045	0.0325	0.0125	0.0075	0.0025

границей, то относится к следующему интервалу. В последний интервал включается также самая правая граница.

Границы интервалов можно выбирать и по другим правилам, например, границы крайних интервалов могут не совпадать с минимальным и максимальным выборочными значениями.

Для каждого интервала находится *эмпирическая частота* по формуле $\tilde{p}_i = \frac{n_i}{n}$. Здесь n_i есть количество значений выборки, попавшее в i-й интервал. Сумма эмпирических частот должна равняться единице.

Также вычисляются *относительные частоты* $\frac{\tilde{p}_i}{\Delta}$, т.е. значения эмпирических частот, отнесенные к длине интервала. Полученные данные заносятся в таблицу, называемую *статистическим рядом*, и графически представляются в виде гистограммы.

Гистограммой называется ступенчатая фигура из прямоугольников с основаниями на Δ и высотами, равными относительным частотам $\frac{\tilde{p}_i}{\Delta}$. Площадь гистограммы равна единице. Гистограмма является кусочно-постоянной аппроксимацией плотности вероятности.

Рис.2. Гистограмма для выборки из нормального распределения с параметрами m = 29; σ = 9 и плотность вероятности.

В первом задании практической работы 2 требуется построить гистограмму для данных из таблицы 5. В отличие от таблицы 4, выборка объёма 100 здесь задается не конкретными значениями точек, а интервально. Весь диапазон значений от $x_{\min} = 10$ до $x_{\max} = 50$ предполагается уже разбитым на l=10 интервалов длины 4. В шестой колонке таблицы перечислены через точку с запятой количества n_i выборочных точек в каждом интервале.

В таблице 1 представлен построенный статистический ряд.

На рисунке 2 совместно изображены гистограмма и плотность вероятности для случайной величины, реализациями которой является данная выборка (нормальная плотность с параметрами m = 29; σ = 9).

3. ТОЧЕЧНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

Пусть имеется выборка и предполагается, что теоретическое распределение зависит от некоторого параметра, который неизвестен и который нужно оценить по выборке. Параметр может быть одномерным или многомерным, т.е. имеется несколько параметров теоретического распределения. Оценкой неизвестного параметра называется любая функция от выборки, так или иначе приближающая этот параметр.

Оценки параметров распределения бывают точечными и интервальными (первые рассматриваются в настоящем разделе).

Оценка $\tilde{\theta}$ параметра θ называется *точечной*, если она указывает значение неизвестного параметра на числовой оси. Точечная оценка зависит от наблюдаемых значений и является случайной величиной. По одной и той же выборке можно построить множество различных (точечных и интервальных) оценок параметра.

Чтобы точечная оценка $\tilde{\theta}$ имела практическую ценность, она, по-возможности, должна быть:

1) *несмещённой*, т.е. математическое ожидание оценки должно равняться значению параметра;
2) *состоятельной*, т.е. должна сходиться по вероятности к оцениваемому параметру;
3) *эффективной*, т.е. должна иметь минимальную дисперсию из всех точечных несмещённых оценок.

Первые два требования к оценке являются обязательными, выполнение последнего требования – желательно. Несмещённость оценки означает, что в среднем её значение совпадает со значением параметра. Состоятельность оценки означает, что при увеличении объема выборки значение оценки должно сблизиться со значением параметра. Если же это не так, то подобная оценка неразумна.

Моментом k-го порядка (*начальным моментом* k-го порядка) случайной величины X называется математическое ожидание k-ой степени этой величины, k > 0.

Такие важные числовые параметры распределения, как математическое ожидание и дисперсия, неизвестны. Математическое ожидание является начальным моментом первого порядка, а дисперсия – центральным моментом второго порядка для распределения.

Неизвестные моменты распределения можно приблизить *эмпирическими моментами*, вычисленными по выборке.

Несмещённой, состоятельной и эффективной оценкой математического ожидания является *среднее выборочное значение* (или выборочный момент первого порядка)

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i .$$

В случае, когда математическое ожидание неизвестно, несмещённой оценкой дисперсии является величина

$$\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 ,$$

использующая оценку $\bar{x}$ математического ожидания m по выборке.

При известном математическом ожидании несмещённой оценкой дисперсии будет

$$\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - m\right)^2 .$$

Обе оценки для дисперсии являются состоятельными.

Во втором задании практической работы 1 среднее выборочное значение получается равным

$\bar{x} =$ (1,8 + 2 + 3,3 + 2,6 + 1,3 – 4 + 0,5 + 0,7 – 0,7 + 5,1 + 5,7 + 2)/12 = 1,69.

Оценка дисперсии $\tilde{\sigma}^2$ (математическое ожидание неизвестно) есть

$$1/(12-1)[(1,8-1,69)^2+(2-1,69)^2+(3,3-1,69)^2+(2,6-1,69)^2+(1,3-1,69)^2+ (-4-1,69)^2+(0,5-1,69)^2+(0,7-1,69)^2+(-0,7-1,69)^2+(5,1-1,69)^2+(5,7-1,69)^2+ (2-1,69)^2] = 6,54.$$

С учетом известного математического ожидания m = 1,5 получается оценка

$$\bar{\sigma}^2 = (1/12)[(1,8-1,5)^2+(2-1,5)^2+(3,3-1,5)^2+(2,6-1,5)^2+(1,3-1,5)^2+ (-4-1,5)^2+(0,5-1,5)^2+(0,7-1,5)^2+(-0,7-1,5)^2+(5,1-1,5)^2+(5,7-1,5)^2+ (2-1,5)^2] = 6,03.$$

Теперь можно найти оценки среднеквадратичного отклонения $\tilde{\sigma} = 2{,}56$ и $\bar{\sigma} = 2{,}46$, которые оказались достаточно близки к истинному (фактически использовавшемуся при генерации выборки) значению σ=2,5.

4. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ ДЛЯ МОМЕНТОВ. ДОВЕРИТЕЛЬНАЯ ВЕРОЯТНОСТЬ

Помимо точечных оценок параметров распределения существуют также интервальные оценки. При построении интервальной оценки нужно указать, в отличие от точечной оценки, некоторый интервал $(\tilde{\theta}_1, \tilde{\theta}_2)$ возможных значений параметра, называемый доверительным интервалом. Его границы $\tilde{\theta}_1$ и $\tilde{\theta}_2$ – случайные числа, поскольку являются функциями выборки. Понятно, что доверительный интервал тем лучше, чем он уже.

Доверительный интервал строится таким образом, чтобы вероятность того, что этот интервал «накроет» оцениваемый параметр, не зависела от истинного значения параметра, т.е.

$$P\left(\tilde{\theta}_1 < \theta < \tilde{\theta}_2\right) = \eta .$$

Эта вероятность η называется *доверительной вероятностью*.

Пусть $\tilde{\theta}$ есть несмещённая точечная оценка неизвестного параметра θ. Симметричный доверительный интервал имеет вид $(\tilde{\theta}_1 - \varepsilon, \tilde{\theta}_2 + \varepsilon)$ и с вероятностью η заключает в себе неизвестный параметр θ. Это условие записывается как

$$P\left(\left|\theta - \tilde{\theta}\right| < \varepsilon\right) = \eta \quad \text{или} \quad P\left(\tilde{\theta} - \varepsilon < \theta < \tilde{\theta} + \varepsilon\right) = \eta .$$

Для построения такого доверительного интервала для параметра θ требуется найти величину ε.

В пособии рассмотрены четыре задачи нахождения доверительных интервалов для параметров нормального распределения: нахождение математического ожидания при известной и неизвестной дисперсии и нахождение дисперсии при известном и неизвестном математическом ожидании. В зависимости от учебного плана часть этих задач может исключаться из задания.

Доверительные интервалы для математического ожидания

Пусть случайная величина подчинена нормальному распределению с математическим ожиданием m и среднеквадратичным (стандартным) отклонением σ.

Случай 1. Дисперсия известна.

Доверительная вероятность для математического ожидания m при известном среднеквадратичном отклонении σ вычисляется по формуле

$$\eta = P\left(\left|m - \bar{x}\right| < \varepsilon\right) = \Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma}\right), \text{ где } \Phi(x) = \frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}} dt \text{ есть функция}$$

Лапласа.

Значение $\frac{\varepsilon\sqrt{n}}{\sigma} = x$ представляет собой аргумент, при котором функция Лапласа равна требуемой доверительной вероятности η, и находится по таблице нормального распределения. Отсюда находится $\varepsilon = \frac{\sigma}{\sqrt{n}} x$, равное половине ширины доверительного интервала. Доверительный интервал имеет вид $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$. При большом объёме выборки n значение неизвестного параметра m можно локализовать довольно точно.

В задании 3 практической работы 1 требуется найти доверительный интервал для математического ожидания при известном значении σ=2,5 и доверительной вероятности η = 0,9.

По таблице для функции Лапласа (приложение 1) определяется аргумент x, при котором функция равна требуемой доверительной вероятности η = 0,9. Соответствующее значение $x \approx 1{,}64$. Отсюда $\varepsilon \approx 1{,}64 \frac{2{,}5}{\sqrt{12}} \approx 1{,}19$. Доверительный интервал, следовательно, получается (1,69–1,19;1,69+1,19)=(0,5;2,88), при этом используется ранее вычисленное значение $\bar{x} = 1{,}69$. Истинное значение математического ожидания равно 1,5.

Рассмотренный случай нахождения доверительного интервала является наиболее простым.

Случай 2. Дисперсия неизвестна.

При неизвестном значении σ доверительный интервал для математического ожидания строится иначе. Используется оценка $\tilde{\sigma}$ вместо σ, а также распределение Стьюдента вместо нормального распределения. В силу центральной предельной теоремы при достаточно больших объёмах выборки можно использовать указанные интервальные оценки и для случайной величины с неизвестным законом распределения. Более того, нормальное распределение можно использовать и при неизвестном σ, т.к. распределение Стьюдента стремится к нормальному распределению при больших значениях n.

В случае, когда σ неизвестно, нужно рассматривать величину $t = \frac{(\bar{x}-m)\sqrt{n}}{\tilde{\sigma}}$, где $\tilde{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$. Случайная величина t распределена по закону Стьюдента с числом степеней свободы $k = n-1$. Вероятности $P(|t| \le \hat{t})$ при различных значениях $\hat{t}$ и числа степеней свободы можно найти в таблицах (приложение 3). Доверительный интервал задается в виде $(\bar{x} - \varepsilon', \bar{x} + \varepsilon')$, где $\varepsilon' = \hat{t}\,\frac{\tilde{\sigma}}{\sqrt{n}}$.

Следует иметь в виду, что часто табулируется не величина $P_k(\hat{t}) = P(|t| \le \hat{t})$, а величина $\breve{P}_k(\hat{t}) = P(t \le \hat{t})$. В этом случае можно пересчитать одну в другую через соотношение $\breve{P}_k(\hat{t}) = 1 - \frac{1 - P_k(\hat{t})}{2}$. Тогда пороговое значение $\hat{t}$ находится по доверительной вероятности η как $\hat{t} = \breve{P}_k^{-1}\left(1 - \frac{1-\eta}{2}\right)$.

Для данных из таблицы 4 доверительная вероятность $\eta = 0{,}9$, число степеней свободы $k = 12 - 1 = 11$. Для использования таблицы $\breve{P}_k(\hat{t})$ приложения 3 предварительно нужно пересчитать.
$\breve{P}_{11}(\hat{t}) = 1 - \frac{1 - P_{11}(\hat{t})}{2} = 1 - \frac{1-\eta}{2} = 0{,}95$. Далее по таблице определяется значение $\hat{t} \approx 1{,}8$. С учётом ранее найденного $\tilde{\sigma} \approx 2{,}56$ находится $\varepsilon' \approx 1{,}8 \frac{2{,}56}{\sqrt{12}} \approx 1{,}33$, откуда получается доверительный интервал $(0{,}37; 3{,}02)$.

В случае неизвестной дисперсии доверительный интервал для математического ожидания получился шире, чем был при известной дисперсии. Это естественно объясняется наличием меньшей информации о распределении. Чем больше известной информации использует исследователь, тем более точные оценки он может получить.

Доверительные интервалы для дисперсии

Случай 3. Математическое ожидание известно.

При построении доверительного интервала для дисперсии (а также для стандартного отклонения) в случае известного математического ожидания рассматривается случайная величина $w = n\frac{\bar{\sigma}^2}{\sigma^2}$, где

$\bar{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2$ – несмещённая оценка дисперсии при известном математическом ожидании.

Величина w имеет распределение χ^2 с $k = n$ степенями свободы.

Далее вводится функция $V_k(\hat{w}) = P(w \ge \hat{w})$.

Искомый доверительный интервал задаётся как (σ_1, σ_2), где

$$\sigma_1 = \bar{\sigma}\sqrt{\frac{n}{a}}, \quad \sigma_2 = \bar{\sigma}\sqrt{\frac{n}{b}},$$

а значения a и b находятся по таблице распределения χ^2 в приложении 4 (читается как «хи-квадрат») в виде

$$a = V_k^{-1}\left(\frac{1-\eta}{2}\right), \quad b = V_k^{-1}\left(1 - \frac{1-\eta}{2}\right).$$

Для данных из таблицы 4, а также для m = 1,5 и $\overline{\sigma} = 2{,}46$ получается:

$$a = V_{12}^{-1}\left(\tfrac{1-0{,}9}{2}\right) = V_{12}^{-1}(0{,}05) \approx 21{,}03,$$

$$b = V_{12}^{-1}\left(1-\tfrac{1-0{,}9}{2}\right) = V_{12}^{-1}(0{,}95) \approx 5{,}23.$$

Поэтому $\sigma_1 \approx 2{,}46\sqrt{\frac{12}{21{,}03}} \approx 1{,}86$, а $\sigma_2 \approx 2{,}46\sqrt{\frac{12}{5{,}23}} \approx 3{,}72$.

Тем самым находится доверительный интервал $(1{,}86; 3{,}72)$ для стандартного отклонения σ, истинное значение которого равно 2,5. Границы доверительного интервала для дисперсии получаются возведением в квадрат полученных границ 1,86 и 3,72.

Случай 4. Математическое ожидание неизвестно.

При построении доверительного интервала для дисперсии в случае неизвестного математического ожидания рассматривается случайная величина $s = (n-1)\dfrac{\tilde{\sigma}^2}{\sigma^2}$, где $\tilde{\sigma}^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2$.

Величина s имеет распределение χ^2 с $k = n - 1$ степенями свободы.

Границы доверительного интервала (σ_1', σ_2') есть

$$\sigma_1' = \tilde{\sigma}\sqrt{\tfrac{n}{a}}, \qquad \sigma_2' = \tilde{\sigma}\sqrt{\tfrac{n}{b}},$$

а величины a и b находятся по таблице распределения χ^2 как

$$a = V_k^{-1}\left(\tfrac{1-\eta}{2}\right), \qquad b = V_k^{-1}\left(1-\tfrac{1-\eta}{2}\right).$$

Вычисления для данных из таблицы 4 с учётом оценки $\tilde{\sigma} = 2{,}56$ будут такие:

$$a = V_{11}^{-1}\left(\tfrac{1-0{,}9}{2}\right) = V_{11}^{-1}(0{,}05) \approx 19{,}66,$$

$$b = V_{11}^{-1}\left(1-\tfrac{1-0{,}9}{2}\right) = V_{11}^{-1}(0{,}95) \approx 4{,}57,$$

$$\sigma_1' \approx 2{,}56\sqrt{\tfrac{11}{19{,}66}} \approx 1{,}91, \qquad \sigma_2' \approx 2{,}56\sqrt{\tfrac{11}{4{,}57}} \approx 3{,}97.$$

Искомый доверительный интервал для стандартного отклонения есть (1,91; 3,97). Аналогично предыдущему случаю находится также интервал для дисперсии.

Построение доверительных интервалов для оцениваемых параметров других распределений читатель может найти в дополнительной литературе.

5. ОЦЕНИВАНИЕ ПАРАМЕТРОВ. МЕТОД МОМЕНТОВ

Пусть исследователем выбран некоторый теоретический закон распределения с одним либо несколькими неизвестными параметрами. Наиболее распространёнными методами нахождения оценок параметров распределения являются метод моментов и метод максимального правдоподобия. Оценка параметров распределения по методу моментов сводится к следующему. Во-первых, параметры распределения выражаются через моменты распределения. Во-вторых, в полученных соотношениях теоретические моменты заменяются их выборочными оценками.

Метод моментов заключается в том, что числовые параметры теоретического распределения выражаются через моменты распределения, оцененные по выборке.

Пусть случайная величина X подчинена нормальному закону с плотностью $f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}$. Тогда параметры распределения m и σ выбираются так, что $m=\bar{x}$, а $\sigma^2=\tilde{\sigma}^2$.

Плотность равномерно распределённой случайной величины

$$f(x)=\begin{cases}\frac{1}{b-a},\ a\le x\le b\\ 0,\quad x\notin[a,b]\end{cases}.$$

Параметрами равномерного распределения являются границы a и b отрезка $[a,b]$, на котором плотность принимает ненулевое значение. Эти параметры необходимо выразить через моменты распределения.

Первый момент (математическое ожидание) есть

$$M(x)=\int_{-\infty}^{\infty}xf(x)dx=\int_a^b\frac{x}{b-a}dx=\frac{a+b}{2}.$$

Второй центральный момент (дисперсия) выражается формулой $D(x)=M(x-M(x))^2$ и преобразуется как

$$D(x)=M(x^2)-(M(x))^2=\int_{-\infty}^{\infty}x^2f(x)dx-(M(x))^2=\int_a^b\frac{x^2}{b-a}dx-\left(\frac{a+b}{2}\right)^2=\frac{(b-a)^2}{12}.$$

Следовательно, система уравнений для нахождения параметров

$$\begin{cases}b+a=2M(x)\\ b-a=\sqrt{12D(x)}=2\sigma\sqrt{3}\end{cases}.$$

Оценивая параметры методом моментов, заменяем $M(x)$ на $\bar{x}$, а $D(x)$ заменяем на $\tilde{\sigma}^2$, далее находим оценки

$$\tilde{a}=\bar{x}-\tilde{\sigma}\sqrt{3},\quad \tilde{b}=\bar{x}+\tilde{\sigma}\sqrt{3}.$$

В задании 2 практической работы 2 требуется найти методом моментов оценки параметров нормального и равномерного распределений. Данные таблицы 5 представляют собой интервальное распределение

Таблица 2

i	$[s_i, s_{i+1})$	n_i	$\tilde{x}_i$	$\tilde{p}_i$	$\tilde{x}_i \tilde{p}_i$	$\tilde{x}_i - \bar{x}$	$(\tilde{x}_i - \bar{x})^2$	$(\tilde{x}_i - \bar{x})^2 \tilde{p}_i$
1	[10,14)	6	12	0,06	0,72	–15,68	245,86	14,75
2	[14,18)	7	16	0,07	1,12	–11,68	136,42	9,55
3	[18,22)	11	20	0,11	2,2	–7,68	58,98	6,49
4	[22,26)	17	24	0,17	4,08	–3,68	13,54	2,3
5	[26,30)	19	28	0,19	5,32	0,32	0,1	0,019
6	[30,34)	18	32	0,18	5,76	4,32	18,66	3,36
7	[34,38)	13	36	0,13	4,68	8,32	69,22	9
8	[38,42)	5	40	0,05	2	12,32	151,78	7,59
9	[42,46)	3	44	0,03	1,32	16,32	266,34	7,99
10	[46,50]	1	48	0,01	0,48	20,32	412,9	4,13
сумма		100		1	27,68			65,18

выборки объёма 100 на *l* интервалов. Поскольку конкретные значения выборочных точек не указаны, для оценивания моментов (математического ожидания и

дисперсии) распределения по выборке нужно выполнить ряд вычислений для каждого из заданных интервалов. Результаты всех вычислений оформляются в виде таблицы.

Сначала для каждого из интервалов $[s_i, s_{i+1})$ фиксируется среднее значение границ

$\tilde{x}_i = \dfrac{s_i + s_{i+1}}{2}$, которому приписывается число n_i выборочных точек, попавших в интервал. Таким способом исходная интервально заданная выборка представляется в виде выборки с повторяющимися n_i раз значениями в точках $\tilde{x}_i$, *i=1, ,l.*

Затем находятся относительные частоты $\tilde{p}_i = \frac{n_i}{n}$, а также величины $\tilde{x}_i \tilde{p}_i$. Среднее выборочное значение, являющееся оценкой математического ожидания, находится как сумма $\bar{x} = \sum\limits_i \tilde{x}_i \tilde{p}_i$, что для нашего примера составляет $\bar{x} \approx 27{,}68$.

Далее последовательно находятся значения $\tilde{x}_i - \bar{x}$, $(\tilde{x}_i - \bar{x})\tilde{p}_i$ и $(\tilde{x}_i - \bar{x})^2\,\tilde{p}_i$ для каждого интервала (последние три столбца таблицы). Результаты всех вычислений занесены в таблицу 2.

Оценкой дисперсии является величина

$$\tilde{\sigma}^2 = \sum_i (\tilde{x}_i - \bar{x})^2\,\tilde{p}_i\,.$$

Сумма чисел в последнем столбце есть оценка дисперсии $\tilde{\sigma}^2 \approx 65{,}18$, откуда находится среднеквадратичное отклонение $\tilde{\sigma} \approx \sqrt{65{,}18} \approx 8{,}07$.

Оцененные по выборке моменты $\bar{x}$ и $\tilde{\sigma}$ являются оценками параметров нормального распределения m и σ методом моментов.

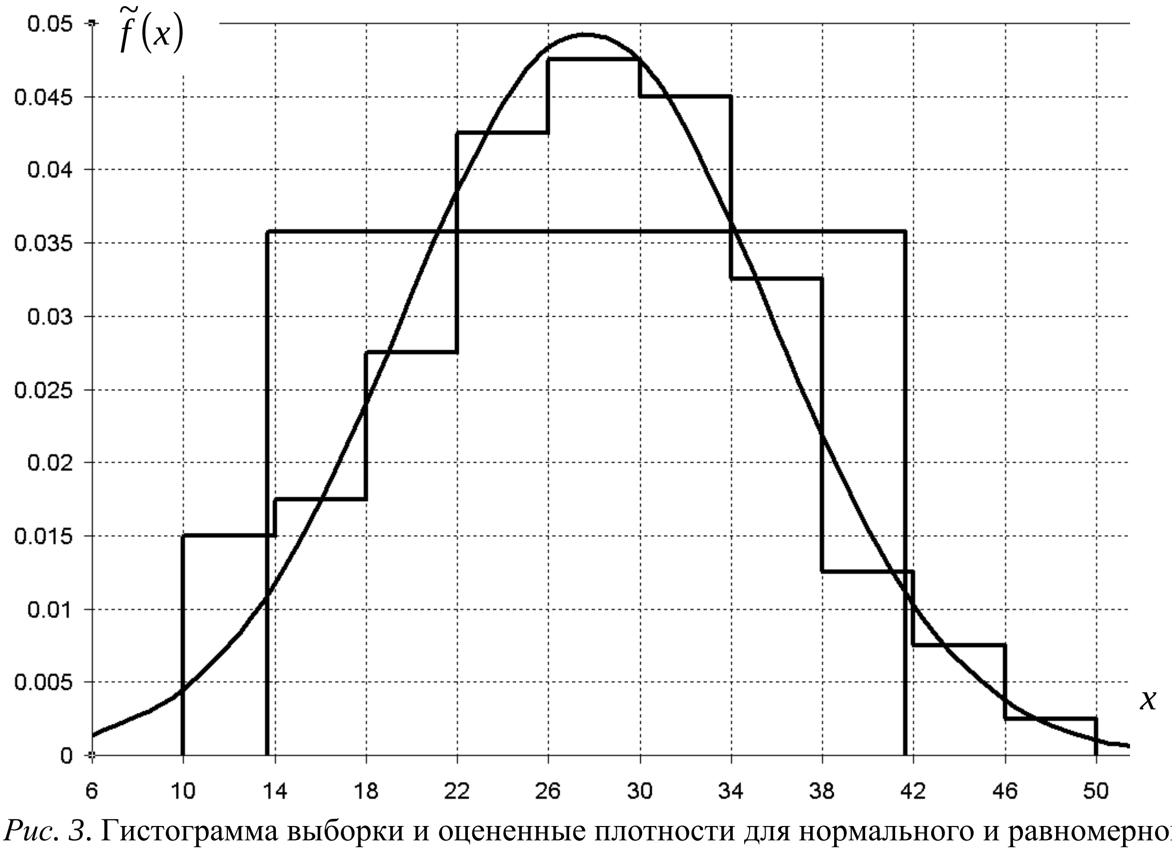

Рис. 3. Гистограмма выборки и оцененные плотности для нормального и равномерного распределений

Оценки параметров a и b равномерного распределения вычисляются по формулам

$$\tilde{a} = \bar{x} - \tilde{\sigma}\sqrt{3} \approx 13{,}7 \text{ и } \tilde{b} = \bar{x} + \tilde{\sigma}\sqrt{3} \approx 41{,}66\,.$$

Оценка плотности равна $\tilde{f}(x) = 0{,}0358$ для $x \in [\tilde{a}, \tilde{b}\,]$.

На рис. 3 изображена гистограмма, на которую наложены нормальная и равномерная плотности, построенные с учётом оцененных методом моментов параметров распределения.

6. МЕТОД МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Пусть, как и ранее, имеется теоретический закон распределения с параметром θ (возможно, многомерным), подлежащим оценке.

Основную идею метода максимального правдоподобия можно пояснить на таком примере. Стрелок 10 раз выстрелил по мишени, из них 6 раз он попал в мишень и 4 раза промахнулся. Вероятность p попадания в мишень при одном выстреле неизвестна, но относительно нее можно строить различные предположения. Из значений p=0,1, p=0,6 и p=0,95 наиболее правдоподобным после завершения стрельбы является второе значение. При любом из этих предположений стрелок может 6 раз попасть в мишень и 4 раза промахнуться, но наиболее вероятен такой результат при p=0,6.

Идея метода максимального правдоподобия состоит в следующем: подбирается такое значение параметра θ, при котором вероятность получить имеющуюся выборку максимальна. Иначе говоря, подбирается наиболее правдоподобное значение параметра.

Пусть имеется функция распределения $F_\theta(x)$ случайной величины X, зависящая от параметра θ, и пусть есть выборка $(x_1,..., x_n)$ из этого распределения. Соответствующая закону $F_\theta(x)$ плотность вероятности есть $f_\theta(x)$.

Функцией правдоподобия называется функция

$$f_\theta(x_1,..., x_n) = \prod_{i=1}^{n} f_\theta(x_i).$$

В случае дискретного распределения функцией правдоподобия будет просто вероятность появления выборки

$$P_\theta(x_1,..., x_n) = \prod_{i=1}^{n} P_\theta(x_i).$$

Оценка θ^* параметра θ является *оценкой максимального правдоподобия*, если в точке θ^* достигается максимум функции правдоподобия (при фиксированном значении выборки $(x_1,..., x_n)$), т.е.

$$f_{\theta^*}(x_1,..., x_n) = \max_\theta f_\theta(x_1,..., x_n).$$

Функцию правдоподобия при фиксированном значении параметра θ можно рассматривать как плотность (или ряд распределения) выборочного вектора $(x_1,..., x_n)$.

При нахождении оценки максимального правдоподобия часто бывает удобно предварительно перейти к *логарифмической функции правдоподобия*

$$L_\theta(x_1,...,x_n) = \ln f_\theta(x_1,...,x_n) = \sum_{i=1}^{n} \ln f_\theta(x_i).$$

Ясно, что максимум функции $L_\theta(x_1,...,x_n)$ также достигается в точке θ^*, при этом с суммой величин работать удобнее, чем с их произведением.

Если распределение имеет несколько параметров, что можно записать вектором $\theta = (\theta_1,...,\theta_d)$, то исследовать на максимум функцию правдоподобия нужно как функцию d переменных.

Если функция правдоподобия достигает максимального значения в нескольких точках, то все они, по определению, считаются оценками максимального правдоподобия.

В некоторых случаях оценки по методу моментов совпадают с оценками максимального правдоподобия. Как правило, оценки максимального правдоподобия являются состоятельными и асимптотически несмещёнными.

Асимптотически несмещённая оценка это оценка, математическое ожидание которой сходится к истинному значению оцениваемого параметра при $n \to \infty$, т.е. при неограниченном росте объёма выборки.

В случае нормального распределения оценки параметров m и σ^2, найденные методом максимального правдоподобия, совпадают с оценками, полученными методом моментов. Это можно показать следующим образом.

Логарифмическая функция правдоподобия равна

$$L_{m,\sigma^2}(x_1,...,x_n) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2 .$$

Точка, в которой достигается максимум функции L, находится из системы уравнений

$$\frac{\partial L_{m,\sigma^2}(x_1,...,x_n)}{\partial m} = 0, \qquad \frac{\partial L_{m,\sigma^2}(x_1,...,x_n)}{\partial(\sigma^2)} = 0, \text{ т.е.}$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - m) = 0, \qquad -\frac{n}{\sigma^2} + \frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i - m)^2 = 0.$$

Решением этой системы являются $m^* = \bar{x}$ и $(\sigma^2)^* = \overline{x^2} - (\bar{x})^2$, т.е. среднее выборочное значение и второй выборочный центральный момент (оценка дисперсии).

Таким образом, во втором задании практической работы 2 оценками максимального правдоподобия для нормального распределения будут $\bar{x} \approx 27{,}68$ и $\tilde{\sigma} \approx \sqrt{65{,}18} \approx 8{,}07$, совпадающие с оценками по методу моментов.

В случае равномерного распределения максимально правдоподобные оценки границ есть минимальное и максимальное значения выборочных точек. Для интервального распределения выборки можно приближённо брать левую и правую границы всего выборочного интервала, $x_{\min}$ и $x_{\max}$.

Плотность равномерно распределённой случайной величины

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & x \notin [a,b] \end{cases}.$$

Функция правдоподобия есть

$$f_{a,b}(x_1,...,x_n) = \begin{cases} \frac{1}{(b-a)^n}, & a \le x_i \le b, i = 1,2,...,n \\ 0, & \textit{иначе} \end{cases}.$$

Иначе это можно записать

$$f_{a,b}(x_1,...,x_n) = \begin{cases} \frac{1}{(b-a)^n}, & \textit{если } a < x_{\min}, b > x_{\max} \\ 0, & \textit{иначе} \end{cases}.$$

Отсюда следует, что функция правдоподобия строго положительна и своего наибольшего значения достигает при таких значениях параметров, при которых разность *b–a* принимает наименьшее значение, но при этом $a < x_{\min} \le x_{\max} < b$. Это достигается при $b = x_{\max}$, $a = x_{\min}$, что дает искомую оценку максимального правдоподобия.

Рассмотренный пример показывает, что не всегда требуется прибегать к дифференцированию функции правдоподобия для нахождения точки её максимума. Ясно также, что ненужным здесь был бы переход к логарифмической функции правдоподобия.

Для примера из практической работы 2 оценками максимального правдоподобия параметров равномерного распределения являются величины 10 и 50, т.е. $x_{\min}$ и $x_{\max}$. Тогда оценка равномерной плотности будет $\tilde{f}(x)$ = 0,025.

На практике нередко оказывается, что оценки по методу моментов получаются хуже оценок максимального правдоподобия для тех же параметров.

7. КРИТЕРИИ СОГЛАСИЯ

Пусть имеется выборка из генеральной совокупности, закон распределения которой неизвестен, и пусть найдены точечные оценки для математического ожидания и дисперсии. Естественно в этом случае строить различные предположения, или гипотезы, относительно вида истинного (теоретического) распределения. Чтобы проверить согласованность наблюдений с гипотезой, можно сравнивать графики эмпирической функции распределения и функции распределения или гистограмму с графиком плотности вероятности. Однако эмпирическая функция распределения из-за случайности выборки может существенно отличаться от истинной функции распределения, даже если гипотеза верна. Задача исследователя состоит в том, чтобы определить, какие отличия эмпирического и теоретического законов не противоречат гипотезе, а какие являются значимыми и гипотезу отвергают.

С помощью *критериев согласия* оценивают, насколько полученные наблюдения согласуются с гипотезой о распределении.

Критерий согласия χ^2

Критерий согласия Пирсона χ^2 (читается «хи-квадрат») наиболее часто используется среди критериев согласия. Меру расхождения эмпирического и теоретического законов распределения этот критерий представляет в виде величины

$$\tilde{\chi}^2 = n\sum_{i=1}^{l}\frac{(\tilde{p}_i - p_i)^2}{p_i} = \sum_{i=1}^{l}\frac{(n_i - np_i)^2}{np_i} .$$

Здесь n есть объём выборки, l – число интервалов, на которые разбивается вся область значений случайной величины, $\tilde{p}_i = \frac{n_i}{n}$ – эмпирические частоты, а p_i – теоретические вероятности попадания в i-й интервал с границами s_i, s_{i+1}, которые находятся по формуле

$$p_i = \int_{s_i}^{s_{i+1}} f(x)dx = F(s_{i+1}) - F(s_i),\ i = 1, 2, \ldots, l ,$$

где $F(x)$ есть функция распределения.

Для нормального теоретического закона

$p_i = \frac{1}{2}[\Phi(\frac{s_{i+1}-\bar{x}}{\tilde{\sigma}}) - \Phi(\frac{s_i-\bar{x}}{\tilde{\sigma}})]$, где $\Phi(x) = \frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ – функция Лапласа.

Для равномерного теоретического закона вероятности есть $p_i = \frac{1}{l}$.

Если проверяемая гипотеза верна, то значение χ^2 является приближённо реализацией случайной величины, имеющей распределение χ^2 с $k = l - d - 1$ степенями свободы, где d – число параметров теоретического распределения. Для нормального и равномерного распределений $d = 2$, поскольку параметрами нормального распределения являются математическое ожидание и среднеквадратическое отклонение, а для равномерного – концы отрезка, где задана ненулевая плотность.

Очевидно, что чем меньше отличаются теоретические вероятности от соответствующих эмпирических частот, тем меньше величина χ^2. Слишком большая мера расхождения означает, что гипотеза о распределении неверна. Какой должна быть величина χ^2, чтобы принять гипотезу, зависит от задаваемого уровня значимости α. *Уровень значимости* α есть вероятность того, что случайная величина χ^2 попадает в *критическую область* $[\hat{\chi}^2,\infty)$, т.е. в область значений, где рассматриваемая гипотеза отвергается.

Критические значения $\hat{\chi}^2$, зависящие от α и k, находятся из условия $\alpha = P(\chi^2 \geq \hat{\chi}^2)$ и табулированы. Величина $\eta = 1-\alpha$ есть вероятность попадания случайной величины χ^2 в *область допустимых значений* $(0, \hat{\chi}^2)$, т.е. в область, при попадании в которую гипотеза принимается.

Схема применения критерия согласия Пирсона

1. На основе выборки вычисляется наблюдаемое значение $\tilde{\chi}^2$ величины χ^2 по приведённой выше формуле.
При заданном уровне значимости α по таблице критических точек распределения χ^2 (приложение 2) находится $\hat{\chi}^2$.

2. Если $\tilde{\chi}^2 < \hat{\chi}^2$, т.е. $\tilde{\chi}^2$ попало в область допустимых значений, то оснований отвергнуть гипотезу нет. В этом случае отклонение эмпирических частот от теоретических вероятностей определяется лишь случайностью выборки и считается незначимым при данном уровне значимости α.

3. Если $\tilde{\chi}^2 \geq \hat{\chi}^2$, то выдвинутая гипотеза о теоретическом распределении отвергается. Здесь расхождение эмпирических частот и теоретических вероятностей является значимым при данном уровне значимости α и не носит случайный характер.

Критерий согласия Пирсона, как и все критерии согласия, позволяет отвергнуть неподходящие гипотезы, но доказательством справедливости гипотезы не служит. Гипотеза принимается не как истинная, а как согласующаяся с результатами наблюдений.

Применять критерий χ^2 можно также другим способом. Известно $\tilde{\chi}^2$ и число степеней свободы k. По любой таблице распределения χ^2 с подходящими входами определяется вероятность p того, что случайная величина, распределенная по закону χ^2, будет больше $\tilde{\chi}^2$. Если найденная вероятность меньше уровня значимости, то результат опыта считается противоречащим выбранному теоретическому закону. Если она сравнительно велика (больше уровня значимости), то расхождение теоретического и эмпирического распределения считается незначимым.

Такая схема дает оценку вероятности полученного уклонения, что позволяет численно выразить меру соответствия выборки гипотезе. Полученная вероятность называется *достигнутым уровнем значимости* (p-value), будем обозначать её $\tilde{\alpha}$.

Использование критерия Пирсона

В задании 4 практической работы 2 нужно проверить по критерию χ^2 гипотезы о согласии выборки с нормальным и равномерным распределениями. Параметры обоих законов распределения оценены ранее методом моментов.

Готовое интервальное задание выборки в таблице 5 можно использовать в качестве разбиения области значений случайной величины на интервалы. Поскольку в соответствии с рекомендациями по использованию критерия требуется, чтобы в интервал попало не менее пяти выборочных точек, то предварительно восьмой, девятый и десятый интервалы объединяются в один (под восьмым номером). Тогда новое число точек в восьмом интервале будет $n_8 = 5+3+1 = 9$. Новое число интервалов $l' = 8$. При достаточном числе выборочных точек в интервалах изменений не требуется. Помимо этого крайние интервалы необходимо продлить соответственно до $-\infty$ и до $+\infty$, в соответствии с требованиями критерия.

Для нормального и равномерного законов вычисляются вероятности p_i, величины np_i и значения χ^2 отдельно для каждого интервала. Все полученные значения занесены в таблицу 3.

Для нормального распределения вероятности вычислены по приведённой выше формуле, с использованием табулированых значений функции Лапласа. Сумма величин χ^2 по столбцу дает искомое значение критерия $\tilde{\chi}_N^2 \approx 1{,}059$. Здесь N в позиции нижнего индекса указывает на нормальный закон.

Для равномерного закона распределения вероятности p_i попадания в каждый из 10 интервалов одинаковы и равны 1/10. Но поскольку последние три интервала объединялись в один, то для восьмого интервала $p_8 = 3 \times 0{,}1 = 0{,}3$. Значение критерия $\tilde{\chi}_U^2 \approx 37{,}6$ также получено суммированием по столбцу. Буквой U обозначено равномерное распределение.

По таблице критических точек распределения χ^2 (приложение 2) находим значение $\hat{\chi}^2$ = 9,24, соответствующее уровню значимости (критической вероятности) α =0,1 и числу степеней свободы k = 8–2–1=5.

Поскольку для нормального закона $\tilde{\chi}_N^2 \approx 1{,}059 < 9{,}24 = \hat{\chi}^2$, то $\tilde{\chi}_N^2$ попадает в область допустимых значений. Поэтому нет оснований отвергнуть гипотезу о нормальности теоретического закона распределения. Расхождение между эмпирическим и теоретическим распределениями незначимо и носит случайный характер.

В случае равномерного распределения значение критерия $\tilde{\chi}_U^2 \approx 37{,}6 > 9{,}24 = \hat{\chi}^2$, следовательно, $\tilde{\chi}_U^2$ попадает в критическую область. Следовательно, гипотеза о равномерном законе распределения отвергается, как не согласующаяся с результатами наблюдений. При данном уровне значимости 0,1 расхождение между эмпирическим и теоретическим законами распределения не является случайным.

Таблица 3

i	интервал	n_i	нормальный закон			равномерный закон		
			p_i	np_i	$\tilde{\chi}_N^2$	p_i	np_i	$\tilde{\chi}_U^2$
1	(–∞,14)	6	0,045	4,509	0,493	0,1	10	1,6
2	[14,18)	7	0,070	7,017	0,000	0,1	10	0,9
3	[18,22)	11	0,126	12,559	0,194	0,1	10	0,1
4	[22,26)	17	0,177	17,672	0,026	0,1	10	4,9
5	[26,30)	19	0,196	19,551	0,016	0,1	10	8,1
6	[30,34)	18	0,170	17,005	0,058	0,1	10	6,4
7	[34,38)	13	0,116	11,629	0,162	0,1	10	0,9
8	[38, +∞)	9	0,101	10,057	0,111	0,3	30	14,7
сумма		100	1	100	1,059	1	100	37,6

Полезно также рассмотреть применение критерия согласия χ^2 с другой стороны, оценивая достигнутый уровень значимости, т.е. вероятность получения отклонения, большего чем $\tilde{\chi}^2$. Число степеней свободы k = 5, $\tilde{\chi}_N^2 \approx 1{,}059$ для нормального и $\tilde{\chi}_U^2 \approx 37{,}6$ для равномерного законов.

В таблице вероятностей для распределения χ^2 (приложение 4) по аргументу 1,059 и k = 5 находится $\tilde{\alpha}_N = P\left(\chi^2 \geq \tilde{\chi}_N^2\right) \approx V_5(1{,}059) \approx 0{,}958$ (ближайшее табличное значение 0,9626 для аргумента 1). Вероятность

0,958 > 0,1 = α (уровня значимости), т.е. вероятность получения такого либо большего отклонения сравнительно велика. Поэтому гипотеза о том, что выборка из нормального распределения, принимается.

Для равномерного закона вероятность есть $\tilde{\alpha}_U = P\left(\chi^2 \geq \tilde{\chi}_U^2\right) \approx V_5(37{,}6) \approx 4{,}5 \cdot 10^{-7} < \alpha$. В таблице вероятность для аргумента 30 указана нулевой, поскольку её значение меньше 10^{-4}, что выходит за рамки точности таблицы. Для аргументов, больших 30, в этом случае можно прибегнуть к дополнительным таблицам либо учесть, что при увеличении аргумента происходит дальнейшее уменьшение вероятности. Найденная вероятность меньше уровня значимости, т.е. при равномерном законе распределения вероятность получить отклонение больше $\tilde{\chi}_U^2$ очень мала. Поэтому гипотеза о равномерном законе распределения отвергается, как не согласующаяся с результатами наблюдений.

8. КРИТЕРИЙ ОТНОШЕНИЯ ПРАВДОПОДОБИЯ

В предыдущем разделе при проверке гипотезы о распределении мы не задавались вопросом о том, каким может быть распределение, если проверяемая гипотеза неверна. В данном разделе делаются предположения об альтернативных распределениях.

Общий вид критерия

Рассмотрим случай, когда у нас есть две гипотезы о распределении, при этом одна из гипотез обязательно верна. Гипотезы будем предполагать *простыми*, что означает, что каждая гипотеза фиксирует не только вид распределения, но и конкретные значения параметров. Заметим, что в предыдущем разделе гипотеза задавала только вид распределения, а параметры оценивались по выборке — такие гипотезы называются *сложными*.

Первую гипотезу принято называть *основной* и обозначать H_0, а вторую — *альтернативной* и обозначать H_1. Такая терминология, очевидно, подразумевает, что гипотезы в каком-то смысле не «равноправны». Однако эта неравноценность гипотез проявляется только в их содержательной интерпретации, с математической точки зрения нет разницы, какую из гипотез обозначить основной.

Хотя из критериев согласия мы рассмотрели только критерий χ^2, различных критериев согласия существует довольно много, и все они так или иначе используются на практике. При этом неизвестно, какой из критериев лучше, а какой хуже, более того, в общем случае неизвестно, как такие критерии сравнивать. Однако в случае двух простых гипотез известен наилучший (в нескольких смыслах, которые мы уточнять не будем) статистический критерий, а именно, критерий отношения правдоподобия.

Как следует из названия, критерий использует отношение двух функций правдоподобия. Для удобства будем рассматривать логарифм этого отношения, который есть разность логарифмических функций правдоподобия

$$\zeta(x_1,\ldots,x_n)=\ln\frac{f_0(x_1,\ldots,x_n)}{f_1(x_1,\ldots,x_n)}=\sum_{i=1}^{n}\ln f_0(x_i)-\sum_{i=1}^{n}\ln f_1(x_i),$$

где $f_0(\cdot)$ и $f_1(\cdot)$ – плотности вероятности для гипотез H_0 и H_1 соответственно.

Собственно критерий отношения правдоподобия заключается в том, что мы принимаем гипотезу H_0, при $\zeta > C$, и отвергаем в противном случае. Здесь C – некоторый пороговый параметр.

Очевидно, что относительно гипотезы мы можем ошибиться. Ситуация, когда мы ошибочно отвергли первую (основную) гипотезу называется *ошибкой первого рода*, ситуация, когда мы ошибочно отвергли вторую (альтернативную) гипотезу называется *ошибкой второго рода*.

Параметр C находится из условия, что вероятность ошибки первого рода равна заданному уровню значимости α, т.е.

$$P_0(\zeta \le C)=\alpha,$$

где вероятность вычисляется по распределению, соответствующему гипотезе H_0.

Величина

$$P_1(\zeta \le C)=\beta,$$

где вероятность вычисляется по распределению, соответствующему гипотезе H_1, называется *мощностью* критерия.

Заметим, что $1-\beta$ есть вероятность ошибки второго рода.

Если в последние выражения вместо C подставить значение $\tilde{\zeta}$, которое есть ζ, вычисленное по имеющейся выборке, то получим $\tilde{\alpha}$ и $\tilde{\beta}$, которые будем называть соответственно достигнутым уровнем значимости и достигнутой мощностью критерия.

Случай нормальных распределений

Критерий отношения правдоподобия достаточно прост в применении, но в общем случае требует численных расчётов или статистического моделирования на компьютере. Получить аналитические выражения удаётся в сравнительно редких случаях, два из которых мы сейчас рассмотрим.

Пусть при гипотезе H_k распределение имеет вид

$$f_k(x) = \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{(x-m_k)^2}{2\sigma_k^2}}.$$

Подставив его в логарифм отношения правдоподобия, получаем

$$\zeta(x_1,\ldots,x_n) = \sum_{i=1}^{n} \frac{(x_i - m_1)^2}{2\sigma_1^2} - \sum_{i=1}^{n} \frac{(x_i - m_0)^2}{2\sigma_0^2} + \ln\frac{\sigma_1}{\sigma_0}.$$

Однако это выражение ещё «недостаточно простое». Рассмотрим дополнительные ограничения.

Пусть $\sigma_0 = \sigma_1 = \sigma$, а $m_0 \neq m_1$. Тогда

$$\zeta(x_1,\ldots,x_n) = \sum_{i=1}^{n} \frac{(x_i - m_1)^2 - (x_i - m_0)^2}{2\sigma^2} = \sum_{i=1}^{n} \frac{x_i(m_0 - m_1)}{\sigma^2} + n \cdot \frac{m_1^2 - m_0^2}{2\sigma^2}.$$

Поскольку критерий не меняется (при соответствующем пересчёте порога C) при любом строго монотонном преобразовании функции $\zeta(x_1,\ldots,x_n)$, можем убрать несущественные константы, получив

$$\zeta'(x_1,\ldots,x_n) = \operatorname{sign}(m_0 - m_1)\cdot \tfrac{1}{n}\sum_{i=1}^{n} x_i, \quad \text{где} \quad \operatorname{sign}(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases}.$$

Величина $\frac{1}{n}\sum_{i=1}^{n} x_i$ как сумма независимых нормальных случайных величин имеет нормальное распределение с параметрами m_k, $\frac{\sigma}{\sqrt{n}}$. Это позволяет легко вычислять значения параметра C, а также уровень значимости и мощность критерия.

Пусть теперь $m_0 = m_1 = m$, а $\sigma_0 \neq \sigma_1$. Тогда

$$\zeta(x_1,\ldots,x_n) = \sum_{i=1}^{n} (x_i - m)^2 \left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2} \right) + \ln\frac{\sigma_1}{\sigma_0}.$$

Убирая несущественные константы, получаем

$$\zeta''(x_1,\ldots,x_n) = \operatorname{sign}(\sigma_0 - \sigma_1)\cdot \sum_{i=1}^{n} (x_i - m)^2.$$

Случайная величина $\sum_{i=1}^{n}\left(\frac{x_i - m}{\sigma_k}\right)^2$ подчиняется распределению χ^2 с n степенями свободы. Напомним, что распределение χ^2 по определению есть распределение, которому подчиняется сумма квадратов независимых случайных величин, каждая из которых распределена нормально с параметрами 0 и 1, а величины $\frac{x_i - m}{\sigma_k}$ как раз таковыми являются.

Полученный результат позволяет, пользуясь таблицами для распределения χ^2, вычислять значения параметра *C*, а также уровень значимости и мощность критерия.

Пример вычислений

Рассмотрим решение задания варианта 0.

Для случая нормальных распределений с равными дисперсиями имеем набор параметров: $m_0' = 1{,}5$, $m_1' = 1{,}7$, $\sigma' = 2{,}5$. Уровень значимости задан $\alpha = 0{,}1$. Объём выборки берём из таблицы 4, откуда $n = 12$.

Найдём константу C' из условия $P_0(\zeta' \le C') = \alpha$.

Учитывая, что $\operatorname{sign}(m_0' - m_1') = \operatorname{sign}(-0{,}2) = -1$, получаем

$\zeta'(x_1,\dots,x_n) = -\frac{1}{n}\sum_{i=1}^{n} x_i$. Преобразуем условие

$$P_0(\zeta' \le C') = P_0\left(-\frac{1}{n}\sum_{i=1}^{n} x_i \le C'\right) = P_0\left(-\frac{\sqrt{n}}{n\sigma'}\sum_{i=1}^{n} x_i + \frac{m_0'}{\sigma'}\sqrt{n} \le \frac{C'+m_0'}{\sigma'}\sqrt{n}\right) = \alpha.$$

Величина $-\frac{\sqrt{n}}{n\sigma'}\sum_{i=1}^{n} x_i + \frac{m_0'}{\sigma'}\sqrt{n}$ имеет нормальное распределение с параметрами 0 и 1, откуда $\frac{C'+m_0'}{\sigma'}\sqrt{n} = x$, где x – значение аргумента, при котором функция нормального распределения равна α. Поскольку функция нормального распределения $F(x)$ связана с функцией Лапласа соотношением $F(x) = \frac{1}{2}\cdot(1+\Phi(x))$ или $\Phi(x) = 2F(x) - 1$, нам нужно по таблице из приложения 1 найти *x*, при котором $\Phi(x) = 2\alpha - 1 = -0{,}8$. Учитывая, что $\Phi(-x) = -\Phi(x)$, находим $x \approx -1{,}2816$.

Получаем $C' = \dfrac{x\sigma'}{\sqrt{n}} - m_0' \approx -2{,}42$.

Находим мощность критерия

$$\beta' = P_1(\zeta' \le C') = P_1\left(-\frac{1}{n}\sum_{i=1}^{n} x_i \le C'\right) = F\left(\frac{C'+m_1'}{\sigma'}\sqrt{n}\right) = F\left(x + \frac{m'-m_0'}{\sigma'}\sqrt{n}\right) \approx 0{,}158.$$

Достигнутые уровень значимости и мощность критерия вычисляются по формулам для α и β', при $C' = -\overline{x}$, где $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \approx 1{,}69$ находится по выборке из таблицы 4.

Находим

$$\widetilde{\alpha}' = P_0(\zeta' \le -\overline{x}) = F\left(\frac{-\overline{x}+m_0'}{\sigma'}\sqrt{n}\right) \approx 0{,}395,$$

$$\widetilde{\beta}' = P_1(\zeta' \le -\overline{x}) = F\left(\frac{-\overline{x}+m_1'}{\sigma'}\sqrt{n}\right) \approx 0{,}505.$$

Поскольку достигнутый уровень значимости не меньше α, гипотезу H_0 не отвергаем, несмотря на то, что гипотеза H_1 для данной выборки более правдоподобна. Если принять гипотезу H_1 основой, то достигнутый уровень значимости будет равен $1-\tilde{\beta}' \approx 0{,}405$.

Видим, что вероятности ошибок первого и второго рода велики. Это объясняется тем, что разность математических ожиданий для гипотез мала по отношению к стандартному отклонению выборочного среднего, поэтому гипотезы близки и неразличимы с приемлемым уровнем достоверности.

Для случая нормальных распределений с равными математическими ожиданиями имеем набор параметров: $m''=1{,}5$, $\sigma_0''=2{,}3$, $\sigma_1''=2{,}5$.

Поскольку $\operatorname{sign}(\sigma_0''-\sigma_1'')=-1$, $\zeta''(x_1,\ldots,x_n)=-\sum_{i=1}^{n}(x_i-m'')^2$. Для нахождения C'' преобразуем уравнение

$$P_0(\zeta''\le C'')=P_0\left(-\sum_{i=1}^{n}(x_i-m'')^2\le C''\right)=P_0\left(\tfrac{1}{\sigma_0''^2}\sum_{i=1}^{n}(x_i-m'')^2\ge-\tfrac{C''}{\sigma_0''^2}\right)=\alpha.$$

Величина $\sum_{i=1}^{n}\left(\dfrac{x_i-m''}{\sigma_0''}\right)^2$ имеет распределение χ^2. Из таблицы приложения 2 находим для $k=n=12$ и $\alpha=0{,}1$ значение $\hat{w}\approx18{,}5$ и вычисляем $C''=-\hat{w}\sigma_0''^2\approx-98{,}1$. Мощность критерия есть

$$\beta''=P_1(\zeta''\le C'')=P_1\left(\tfrac{1}{\sigma_1''^2}\sum_{i=1}^{n}(x_i-m'')^2\ge-\tfrac{C''}{\sigma_1''^2}\right)\approx0{,}205.$$

Данное значение получено из таблицы приложения 4 для $k=n=12$ и $\hat{w}=\frac{C''}{\sigma_1''^2}\approx15{,}7$.

Для вычисления достигнутых уровня значимости и мощности критерия вместо C'' подставляем вычисленное по выборке из таблицы 4 значение $\tilde{C}''=-\frac{1}{\sigma_0''}\sum_{i=1}^{n}(x_i-m'')^2=-72{,}41$.

Пользуясь таблицей приложения 4, находим

$$\tilde{\alpha}''=P_0\left(\tfrac{1}{\sigma_0''^2}\sum_{i=1}^{n}(x_i-m'')^2\ge-\tfrac{\tilde{C}''}{\sigma_0''^2}\approx13{,}7\right)\approx0{,}32,$$

$$\tilde{\beta}''=P_1\left(\tfrac{1}{\sigma_1''^2}\sum_{i=1}^{n}(x_i-m'')^2\ge-\tfrac{\tilde{C}''}{\sigma_1''^2}\approx11{,}6\right)\approx0{,}48.$$

Поскольку достигнутый уровень значимости не меньше α, гипотезу H_0 не отвергаем, при этом гипотеза H_1 также не может быть отвергнута, поскольку, если принять её основой, то достигнутый уровень значимости будет $1-\tilde{\beta}'' \approx 0{,}52$. Делаем вывод, что имеющегося объёма выборки недостаточно, чтобы с требуемой достоверностью различать гипотезы.

Объём выборки достаточен для различения гипотез, если мощность критерия оказывается не меньше, чем $1-\alpha$.

9. ЗАДАНИЯ ПРАКТИЧЕСКИХ РАБОТ

Практическая работа 1

Для выполнения заданий использовать данные из таблицы 4.

1. Построить график эмпирической функции распределения. Построить график функции распределения.
2. Найти оценку математического ожидания, а также несмещённые оценки дисперсии при известном и неизвестном математическом ожидании.
3. Построить доверительные интервалы для математического ожидания и стандартного отклонения при известном и неизвестном втором параметре соответственно.

Практическая работа 2

Для выполнения заданий использовать данные из таблицы 5.

1. Построить гистограмму.
2. Найти оценки параметров нормального и равномерного распределений по методу моментов. Найти оценки максимального правдоподобия для параметров обоих законов распределения.
3. Наложить графики оцененных методом моментов плотностей на гистограмму.
4. Проверить по критерию согласия χ^2 гипотезы о согласии выборки с нормальным и равномерным распределениями.

Практическая работа 3

Для выполнения заданий использовать данные из таблиц 6 и 4.

1. Найти пороговое значение статистики и мощность критерия отношения правдоподобия для заданного уровня значимости.
2. Найти достигнутые уровень значимости и мощность критерия для выборки из таблицы 4.
3. Сделать вывод от принятии гипотезы и степени достоверности.

Задания выполняются для двух случаев: случая различных математических ожиданий при равенстве дисперсий (в таблице 6 соответствующие параметры отмечены штрихом) и случая различных дисперсий при равенстве математических ожиданий (в таблице 6 соответствующие параметры отмечены двумя штрихами).

Таблица 4

№	n	выборка	η	m	σ
0	12	1,8; 2; 3,3; 2,6; 1,3; -4; 0,5; 0,7; -0,7; 5,1; 5,7; 2	0,9	1,5	2,5
1	15	-6; -4,4; -2; -7,6; -0,4; 0,1; -3,7; -5,4; -0,8; -3,9; -5,3; -0,3; -4,8; -8,6; -0,9	0,95	-4	2
2	10	25; 27; 25; 20; 30; 25; 20; 23; 26; 22	0,9	25,2	3,4
3	12	-93; 52; 192; 79; 150; 102; -147; -165; -16; 105; 144; 162	0,95	10,5	102,1
4	14	73; -16; -66; 38; -85; -11; 24; 93; 112; 1; 15; 36; -7; 70	0,99	14	45
5	11	0; 9; 2,4; -1,7; 5; 5,4; 3,4; 5,7; 12,5; 4,5; 6,9	0,9	5	5
6	10	3,2; -0,5; 2; 1,6; 1,1; 2,7; 1,6; 0; 1,7; 4,8	0,9	1	2
7	12	8,33; 8,36; 8,23; 8,42; 7,95; 8,16; 8,32; 8,21; 8,27; 8,08; 8,09; 8,02	0,95	8,2	0,1
8	10	0,108; 0,093; 0,11; 0,117; 0,12; 0,089; 0,113; 0,111; 0,092; 0,091	0,9	0,1	0,01
9	15	-0,87; 0,4; -2,7; -0,01; 1,25; -0,9; 0,58; -0,61; 1,25; -1,04; -0,62; -0,19; -0,16; 2,31; 1,05	0,95	0	1,4
10	11	298; 322; 331; 346; 299; 337; 318; 313; 329; 304; 317	0,9	315	17
11	17	25; 29; 19; 26; 23; 16; 20; 22; 24; 18; 18; 30; 19; 26; 24; 24; 19	0,99	22	4
12	12	102; 39; -111; 87; 150; -76; 164; 151; 60; 127; 149; 94	0,9	100	100
13	10	-1,38; -2,21; -0,8; -0,1; -0,21; -0,54; -0,98; -3,05; -0,08; -0,19	0,9	-1	1
14	14	0,75; 0,34; 0,8; 0,86; 0,55; 0,43; 0,34; 0,84; 1,04; 0,58; 1,25; 0,76; 0,82; 1,16	0,95	0,8	0,25
15	11	-1,56; -1,73; -1,72; -1,54; -0,7; -1,58; -1,04; -1,18; -1,83; -1,51; -1,99	0,95	-1,5	0,3
16	12	3,59; 2,48; 2,35; 3,94; 3,58; 2,99; 3,75; 3,42; 3,33; 3,97; 2,98; 4,15	0,9	3,3	0,5
17	15	9,4; 7,3; 9,4; 8,4; 8,7; 5,7; 6,3; 5,3; -1,2; 5,2; 3,6; 4,3; 11,5; 6,5; 8,2	0,95	7	3
18	17	0,7; 1,4; -0,7; -1,5; 0,1; -0,1; -1; 0,1; 1,4; 0,9; -0,8; -0,3; 3; 0,3; -0,7; 0,1; 1,2	0,9	0	1
19	11	19,2; 20,1; 20,6; 19,4; 19,7; 19,1; 19,9; 18,7; 20,2; 19,8; 19,4	0,9	19,5	0,9
20	10	39,3; 21; 33,8; 34,9; 27,9; 20,6; 42,9; 30,2; 22,2; 23	0,9	32,2	7,3
21	10	20,5; 8,2; 20; 12,6; 14,1; 13,4; 12,1; 16; 19,9; 19,4	0,9	12	5
22	12	-9,9; -12,9; -11,7; -11,6; -12,3; -12; -12,7; -11,3; -15,7; -9,5; -10,4; -13,1	0,95	-12	2
23	15	416; 381; 383; 419; 428; 408; 397; 393; 400; 413; 405; 409; 404; 404; 400	0,9	400	20
24	12	59; 60; 70; 69; 63; 62; 71; 69; 67; 67; 76; 58	0,95	65	6,5
25	11	-3; 7; 6; 3; 18; 3; 19; 7; 5; 0; 25	0,9	9	12
26	12	-20; 79; 18; 57; 10; 94; -48; 3; -16; 115; 88; -8	0,95	-5	100
27	15	0,16; -0,09; 0,05; -0,1; 0,23; -0,18; 0,11; -0,11; 0,08; 0,09; 0,06; 0,11; -0,27; -0,1; 0,18	0,99	0	0,15
28	12	0,4; -13,5; 4,3; -4,8; -4,5; -2,5; -3,9; 0,9; 2,3; -4; -11,6; 6,7	0,9	-1	5
29	10	87; 96; 77; 91; 82; 92; 80; 111; 81; 84	0,9	90	10
30	11	68; 126; -60; 89; 15; -65; -59; -99; 19; 66; 26	0,95	50	100

Таблица 5

№	n	l	x_{min}	x_{max}	n_i	α
0	100	10	10	50	6; 7; 11; 17; 19; 18; 13; 5; 3; 1	0,1
1	162	8	-100	100	17; 20; 19; 25; 20; 22; 21; 18	0,1
2	150	10	-50	50	16; 14; 15; 16; 16; 14; 15; 15; 15; 14	0,05
3	240	12	22	46	5; 8; 15; 23; 33; 37; 38; 30; 23; 15; 7; 6	0,05
4	224	10	1	10	23; 22; 24; 23; 24; 23; 20; 23; 21; 21	0,1
5	100	12	5	17	2; 6; 6; 8; 15; 18; 12; 14; 8; 7; 2; 2	0,05
6	102	8	0	50	1; 5; 16; 28; 30; 16; 5; 1	0,05
7	100	12	-6	6	2; 3; 6; 10; 12; 16; 15; 14; 11; 5; 4; 2	0,1
8	300	10	-5	5	5; 13; 26; 46; 58; 60; 46; 27; 13; 6	0,05
9	190	8	-50	50	4; 15; 32; 38; 39; 32; 23; 7	0,1
10	295	10	0	50	5; 13; 26; 45; 56; 60; 44; 28; 13; 5	0,1
11	294	12	3	15	24; 25; 23; 27; 26; 25; 25; 23; 24; 23; 23; 26	0,05
12	200	12	10	50	4; 11; 12; 15; 30; 36; 24; 27; 16; 14; 8; 3	0,1
13	100	10	30	60	8; 9; 12; 11; 13; 9; 10; 9; 9; 10	0,1
14	250	8	10	50	1; 11; 38; 71; 74; 40; 12; 3	0,05
15	150	8	100	900	18; 20; 19; 17; 21; 19; 18; 18	0,1
16	500	12	-50	50	2; 8; 21; 44; 78; 96; 100; 73; 45; 22; 8; 3	0,05
17	90	10	8	28	3; 6; 9; 14; 14; 14; 13; 8; 6; 3	0,1
18	204	8	-8	0	2; 9; 32; 59; 59; 32; 9; 2	0,05
19	500	8	-100	100	4; 24; 79; 147; 138; 81; 23; 4	0,05
20	100	10	-100	200	5; 7; 10; 17; 20; 18; 13; 5; 3; 2	0,1
21	150	10	0	20	14; 14; 14; 16; 15; 16; 14; 17; 14; 16	0,05
22	200	10	1	21	21; 20; 19; 22; 20; 22; 21; 18; 18; 19	0,1
23	204	8	50	90	2; 9; 33; 59; 58; 32; 9; 2	0,1
24	300	12	0	50	8; 10; 19; 29; 41; 44; 46; 39; 29; 18; 10; 7	0,05
25	305	12	0	24	24; 25; 25; 24; 25; 27; 25; 23; 27; 27; 27; 26	0,05
26	100	10	-100	100	6; 5; 11; 17; 19; 16; 15; 5; 3; 3	0,1
27	220	8	-50	50	16; 24; 32; 41; 37; 32; 24; 14	0,05
28	215	10	1	6	19; 23; 20; 20; 23; 23; 23; 22; 22; 20	0,1
29	170	8	3,1	3,9	23; 21; 19; 22; 20; 22; 24; 19	0,1
30	155	10	-100	100	17; 16; 14; 15; 16; 16; 16; 15; 14; 16	0,05

Таблица 6

№	m_0'	m_1'	σ'	m''	σ_0''	σ_1''	α
0	1,5	1,7	2,5	1,5	2,3	2,5	0,1
1	–4	–2	2	–4	3	2	0,05
2	25	20	3,4	25,2	3,5	4	0,05
3	0	100	102,1	10,5	100	80	0,1
4	14	0	45	14	45	50	0,05
5	5	7	5	5	5	4	0,1
6	2	0	2	1	2	4	0,1
7	8,2	8,3	0,1	8,2	0,15	0,1	0,05
8	0,1	0,09	0,01	0,1	0,01	0,02	0,1
9	–1	0	1,4	0	1,4	1	0,05
10	320	300	17	315	15	17	0,05
11	20	24	4	22	4	3	0,1
12	100	50	100	100	100	120	0,05
13	–1	0	1	–1	1,5	1	0,1
14	0,8	0,5	0,25	0,8	0,25	0,4	0,1
15	–1,5	–1	0,3	–1,5	0,3	0,2	0,05
16	3,5	3	0,5	3,3	0,5	1	0,1
17	7	8	3	7	3	1,5	0,05
18	0	–1	1	0	0,8	1,1	0,05
19	20	21	0,9	19,5	1	0,5	0,1
20	32	30	7,3	32,2	7	10	0,05
21	12	15	5	12	4	3	0,1
22	–12	–13	2	–12	2	3	0,1
23	400	420	20	400	25	15	0,05
24	65	60	6,5	65	5	7	0,1
25	7	10	12	9	12	10	0,05
26	0	–5	100	–5	90	110	0,05
27	0	0,1	0,15	0	0,15	0,1	0,1
28	0	–1	5	–1	5	8	0,05
29	90	100	10	90	11	10	0,1
30	50	–50	100	50	100	150	0,1

10. ОТВЕТЫ

Ответы для практической работы 1 приведены в таблице 7.

Таблица 7

№	$\bar{x}$	$\bar{\sigma}$	$\tilde{\sigma}$	ε	ε'	σ_1	σ_2	σ_1'	σ_2'
0	1,69	2,45	2,56	1,19	1,33	1,85	3,71	1,91	3,97
1	–3,60	2,68	2,77	1,01	1,54	1,98	4,14	2,03	4,37
2	24,30	2,97	3,13	1,77	1,81	2,19	4,73	2,28	5,15
3	47,1	118,3	123,5	57,8	78,5	84,8	195,2	87,5	209,7
4	19,8	54,4	56,5	31,0	45,5	36,4	100,9	37,3	107,9
5	4,83	3,76	3,95	2,48	2,16	2,81	5,83	2,92	6,28
6	1,82	1,45	1,52	1,04	0,88	1,07	2,30	1,11	2,51
7	8,20	0,14	0,15	0,06	0,09	0,10	0,23	0,10	0,25
8	0,10	0,011	0,012	0,005	0,007	0,008	0,018	0,009	0,020
9	–0,02	1,18	1,22	0,71	0,68	0,87	1,82	0,89	1,92
10	319,5	14,8	15,5	8,4	8,5	11,1	23,0	11,5	24,7
11	22,47	3,93	4,05	2,50	2,87	2,71	6,78	2,77	7,14
12	78,0	85,4	89,2	47,5	46,2	64,5	129,4	66,7	138,3
13	–0,95	0,95	1,00	0,52	0,58	0,70	1,51	0,73	1,64
14	0,75	0,27	0,28	0,13	0,16	0,20	0,43	0,20	0,46
15	–1,49	0,36	0,38	0,18	0,25	0,25	0,61	0,26	0,66
16	3,38	0,55	0,58	0,24	0,30	0,42	0,84	0,43	0,89
17	6,57	2,94	3,05	1,52	1,69	2,18	4,56	2,23	4,81
18	0,24	1,08	1,11	0,40	0,47	0,85	1,51	0,87	1,58
19	19,65	0,52	0,55	0,45	0,30	0,39	0,81	0,41	0,88
20	29,58	7,57	7,98	3,80	4,63	5,60	12,06	5,82	13,13
21	15,62	3,99	4,21	2,60	2,44	2,95	6,36	3,07	6,93
22	–11,93	1,59	1,66	1,13	1,05	1,14	2,62	1,17	2,81
23	404,0	12,22	12,65	8,49	5,75	9,47	17,56	9,72	18,46
24	65,92	5,28	5,52	3,68	3,51	3,79	8,72	3,91	9,37
25	8,18	8,31	8,72	5,95	4,76	6,21	12,89	6,44	13,88
26	31,0	51,0	53,3	56,6	33,8	36,6	84,2	37,7	90,5
27	0,01	0,14	0,15	0,10	0,11	0,10	0,25	0,10	0,27
28	–2,52	5,71	5,96	2,37	3,09	4,31	8,65	4,46	9,25
29	88,10	9,51	10,03	5,20	5,81	7,03	15,15	7,31	16,50
30	11,5	69,9	73,3	59,1	49,2	49,5	118,6	51,2	128,6

Ответы для практической работы 2 приведены в таблице 8.

Таблица 8

№	$\bar{x}$	$\tilde{\sigma}$	$\tilde{a}$	$\tilde{b}$	l'	$\hat{\chi}^2$	$\tilde{\chi}_N^2$	$\tilde{\alpha}_N$	$\tilde{\chi}_U^2$	$\tilde{\alpha}_U$
0	27,68	8,07	13,70	41,66	8	9,24	1,06	0,958	37,60	0,000
1	1,23	55,14	–94,27	96,74	8	9,24	8,56	0,128	2,15	0,828
2	–0,53	28,60	–50,07	49,01	10	14,07	9,66	0,209	0,40	1,000
3	33,98	4,92	25,45	42,50	12	16,92	0,52	1,000	84,20	0,000
4	5,42	2,56	0,98	9,86	10	12,02	14,24	0,047	0,73	0,998
5	10,81	2,47	6,53	15,09	9	12,59	2,64	0,853	35,04	0,000
6	25,06	8,15	10,94	39,18	6	7,81	0,14	0,986	73,06	0,000
7	0,06	2,44	–4,16	4,28	9	10,64	0,39	0,999	35,70	0,000
8	0,03	1,92	–3,30	3,36	10	14,07	0,23	1,000	133,33	0,000
9	2,04	20,89	–34,14	38,21	7	7,78	3,68	0,452	53,01	0,000
10	25,09	9,60	8,46	41,72	10	12,02	0,52	0,999	130,25	0,000
11	8,97	3,44	3,00	14,93	12	16,92	21,67	0,010	0,86	1,000
12	29,67	8,33	15,23	44,10	10	12,02	5,76	0,568	69,70	0,000
13	44,94	8,32	30,53	59,35	10	12,02	6,01	0,539	2,20	0,948
14	30,28	6,35	19,27	41,29	6	7,81	0,27	0,966	189,86	0,000
15	498,0	227,1	104,7	891,3	8	9,24	9,43	0,093	0,61	0,987
16	0,10	16,44	–28,38	28,58	10	14,07	0,40	1,000	355,01	0,000
17	17,91	4,43	10,23	25,59	8	9,24	0,78	0,978	19,22	0,002
18	–4,00	1,28	–6,22	–1,78	6	7,81	0,27	0,965	154,08	0,000
19	–0,20	32,07	–55,75	55,35	6	7,81	0,70	0,874	367,38	0,000
20	35,90	60,96	–69,69	141,49	8	9,24	0,74	0,980	38,93	0,000
21	10,19	5,70	0,31	20,06	10	14,07	9,08	0,247	0,80	0,997
22	10,81	5,68	0,97	20,65	10	12,02	11,12	0,133	1,00	0,995
23	69,95	6,43	58,82	81,08	6	6,25	0,29	0,962	152,04	0,000
24	24,85	10,42	6,80	42,90	12	16,92	0,41	1,000	97,36	0,000
25	12,20	6,91	0,23	24,17	12	16,92	27,27	0,001	0,82	1,000
26	–8,00	42,19	–81,08	65,08	8	9,24	1,76	0,881	35,33	0,000
27	–0,51	24,29	–42,58	41,56	8	11,07	3,21	0,667	23,71	0,000
28	3,52	1,41	1,08	5,97	10	12,02	15,97	0,025	1,05	0,994
29	3,50	0,23	3,10	3,90	8	9,24	13,22	0,021	1,11	0,954
30	–0,97	57,80	–101,1	99,14	10	14,07	9,93	0,193	0,55	0,999

Ответы для практической работы 3 приведены в таблице 9.

Таблица 9

№	C'	β'	$\tilde{\alpha}'$	$\tilde{\beta}'$	C''	β''	$\tilde{\alpha}''$	$\tilde{\beta}''$
0	–2,42	0,158	0,395	0,505	–98,13	0,205	0,321	0,480
1	3,15	0,987	0,219	0,999	65,35	0,640	0,338	0,975
2	23,23	0,999	0,258	1,000	–224,3	0,172	0,643	0,814
3	–37,77	0,983	0,055	0,964	63038	0,371	0,896	0,996
4	–5,78	0,315	0,685	0,950	–47962	0,158	0,109	0,268
5	–6,93	0,518	0,546	0,925	139,4	0,352	0,143	0,447
6	1,19	0,970	0,388	0,998	–63,95	0,947	0,734	0,998
7	–8,25	0,966	0,454	1,000	0,118	0,535	0,420	0,976
8	0,096	0,970	0,918	1,000	–0,0016	0,947	0,148	0,962
9	0,405	0,869	0,003	0,519	14,23	0,492	0,221	0,857
10	311,6	0,988	0,458	1,000	–4427	0,168	0,388	0,613
11	–21,24	0,998	0,005	0,943	161,4	0,607	0,520	0,970
12	52,52	0,535	0,223	0,834	–210261	0,264	0,674	0,890
13	0,595	0,970	0,442	0,999	10,95	0,638	0,053	0,468
14	0,714	0,999	0,234	1,000	–1,317	0,877	0,250	0,946
15	1,351	1,000	0,452	1,000	0,412	0,496	0,849	1,000
16	3,32	0,985	0,198	0,996	–4,64	0,969	0,245	0,988
17	–8,27	0,362	0,709	0,967	65,35	0,984	0,531	1,000
18	–0,40	0,993	0,840	1,000	–17,66	0,625	0,013	0,444
19	–20,35	0,992	0,904	1,000	5,58	0,978	0,013	0,709
20	28,20	0,218	0,147	0,428	–897,0	0,535	0,218	0,779
21	–14,03	0,731	0,011	0,347	77,84	0,434	0,948	1,000
22	–12,74	0,674	0,552	0,969	–74,20	0,766	0,818	0,992
23	–408,5	0,987	0,219	0,999	4538	0,834	0,002	0,249
24	62,60	0,917	0,687	0,999	–463,7	0,663	0,314	0,855
25	–12,95	0,207	0,372	0,692	658,8	0,169	0,086	0,257
26	–47,48	0,071	0,859	0,894	–170311	0,296	0,927	0,986
27	–0,05	0,903	0,352	0,986	0,192	0,797	0,428	0,989
28	–2,37	0,171	0,041	0,147	–525,7	0,768	0,159	0,886
29	–94,05	0,970	0,726	1,000	588,7	0,175	0,349	0,506
30	11,36	0,979	0,101	0,979	–172750	0,742	0,799	0,989

ПРИЛОЖЕНИЕ 1

Нормальное распределение

Функция Лапласа $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int\limits_0^x e^{-\frac{t^2}{2}} dt$

x	**0,0**	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**
0,	0	797	1585	2358	3108	3829	4515	5161	5763	6319
1,	6827	7287	7699	8064	8385	8664	8904	9109	9231	9426
2,	9545	9643	9722	9786	9836	9876	9907	9931	9949	9963
3,	9973	9981	9986	9990	9993	9995	9997	9998	9999	9999

В таблице приведены значения $\Phi(x) \cdot 10^4$, В первом столбце указаны целые, а в верхней строке — десятые доли аргумента x,

Таблица некоторых значений x, отвечающих часто используемым значениям функции $\Phi(x)$:

$\Phi(x)$	0,8	0,9	0,95	0,99	0,999
x	1,2816	1,6449	1,9600	2,5758	3,2905

ПРИЛОЖЕНИЕ 2

Критические точки распределения хи–квадрат

Закон распределения $\chi^2: V_k(\hat{w}) = P(w \geq \hat{w}) = \frac{1}{\Gamma\left(\frac{k}{2}\right)2^{\frac{k}{2}}}\int\limits_{\hat{w}}^{\infty} w^{\frac{k}{2}-1}e^{-\frac{w}{2}}\,dw$

k \ α	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995
1	7,88	6,63	5,02	3,84	2,71	1,32	0,455	0,102	0,016	0,0039	0,00098	0,00016	0,000039
2	10,6	9,21	7,38	5,99	4,61	2,77	1,39	0,575	0,211	0,103	0,051	0,020	0,010
3	12,8	11,3	9,35	7,81	6,25	4,11	2,37	1,21	0,584	0,352	0,216	0,115	0,072
4	14,9	13,3	11,1	9,49	7,78	5,39	3,36	1,92	1,06	0,711	0,484	0,297	0,207
5	16,7	15,1	12,8	11,1	9,24	6,63	4,35	2,67	1,61	1,15	0,831	0,554	0,412
6	18,5	16,8	14,4	12,6	10,6	7,84	5,35	3,45	2,20	1,64	1,24	0,872	0,676
7	20,3	18,5	16,0	14,1	12,0	9,04	6,35	4,25	2,83	2,17	1,69	1,24	0,99
8	22,0	20,1	17,5	15,5	13,4	10,2	7,34	5,07	3,49	2,73	2,18	1,65	1,34
9	23,6	21,7	19,0	16,9	14,7	11,4	8,34	5,90	4,17	3,33	2,70	2,09	1,73
10	25,2	23,2	20,5	18,3	16,0	12,5	9,34	6,74	4,87	3,94	3,25	2,56	2,16
11	26,8	24,7	21,9	19,7	17,3	13,7	10,3	7,58	5,58	4,57	3,82	3,05	2,60
12	28,3	26,2	23,3	21,0	18,5	14,8	11,3	8,44	6,30	5,23	4,40	3,57	3,07
13	29,8	27,7	24,7	22,4	19,8	16,0	12,3	9,30	7,04	5,89	5,01	4,11	3,57
14	31,3	29,1	26,1	23,7	21,1	17,1	13,3	10,2	7,79	6,57	5,63	4,66	4,07
15	32,8	30,6	27,5	25,0	22,3	18,2	14,3	11,0	8,55	7,26	6,26	5,23	4,60
16	34,3	32,0	28,8	26,3	23,5	19,4	15,3	11,9	9,31	7,96	6,91	5,81	5,14
17	35,7	33,4	30,2	27,6	24,8	20,5	16,3	12,8	10,1	8,67	7,56	6,41	5,70
18	37,2	34,8	31,5	28,9	26,0	21,6	17,3	13,7	10,9	9,39	8,23	7,01	6,26
19	38,6	36,2	32,9	30,1	27,2	22,7	18,3	14,6	11,7	10,1	8,91	7,63	6,84
20	40,0	37,6	34,2	31,4	28,4	23,8	19,3	15,5	12,4	10,9	9,59	8,26	7,43
21	41,4	38,9	35,5	32,7	29,6	24,9	20,3	16,3	13,2	11,6	10,3	8,90	8,03
22	42,8	40,3	36,8	33,9	30,8	26,0	21,3	17,2	14,0	12,3	11,0	9,54	8,64
23	44,2	41,6	38,1	35,2	32,0	27,1	22,3	18,1	14,8	13,1	11,7	10,2	9,26
24	45,6	43,0	39,4	36,4	33,2	28,2	23,3	19,0	15,7	13,8	12,4	10,9	9,89
25	46,9	44,3	40,6	37,7	34,4	29,3	24,3	19,9	16,5	14,6	13,1	11,5	10,5
26	48,3	45,6	41,9	38,9	35,6	30,4	25,3	20,8	17,3	15,4	13,8	12,2	11,2
27	49,6	47,0	43,2	40,1	36,7	31,5	26,3	21,7	18,1	16,2	14,6	12,9	11,8
28	51,0	48,3	44,5	41,3	37,9	32,6	27,3	22,7	18,9	16,9	15,3	13,6	12,5
29	52,3	49,6	45,7	42,6	39,1	33,7	28,3	23,6	19,8	17,7	16,0	14,3	13,1
30	53,7	50,9	47,0	43,8	40,3	34,8	29,3	24,5	20,6	18,5	16,8	15,0	13,8

В таблице приведены значения $\hat{w}$, представляющие собой величины $\chi^2_{крит.}$,соответствующие уровню значимости α (критической вероятности).

ПРИЛОЖЕНИЕ 3

Распределение Стьюдента

Закон распределения Стьюдента: $\breve{P}_k(\hat{t}) = \dfrac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \int\limits_{-\infty}^{\hat{t}} (1+\frac{t^2}{k})^{-\frac{k+1}{k}}\, dt$

$\hat{t}$ \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	∞
0	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000
0,1	5317	5353	5367	5374	5379	5382	5384	5386	5387	5388	5389	5390	5391	5391	5392	5398
0,2	5628	5700	5729	5744	5753	5760	5764	5768	5770	5773	5774	5776	5777	5778	5779	5793
0,3	5928	6038	6081	6104	6119	6129	6136	6141	6145	6148	6151	6153	6155	6157	6159	6179
0,4	6211	6361	6420	6452	6472	6485	6495	6502	6508	6512	6516	6519	6522	6524	6526	6554
0,5	6476	6667	6743	6783	6809	6826	6838	6847	6855	6861	6865	6869	6873	6876	6878	6915
0,6	6720	6953	7046	7096	7127	7148	7163	7174	7183	7191	7197	7202	7206	7210	7213	7257
0,7	6944	7218	7328	7387	7424	7449	7467	7481	7492	7501	7508	7514	7519	7523	7527	7580
0,8	7148	7462	7589	7657	7700	7729	7750	7766	7778	7788	7797	7804	7810	7815	7819	7881
0,9	7333	7684	7828	7905	7953	7986	8010	8028	8042	8054	8063	8071	8078	8083	8088	8159
1	7500	7887	8045	8130	8184	8220	8247	8267	8283	8296	8306	8315	8322	8329	8334	8413
1,1	7651	8070	8242	8335	8393	8433	8461	8483	8501	8514	8526	8535	8544	8551	8557	8643
1,2	7789	8235	8419	8518	8581	8623	8654	8678	8696	8711	8723	8734	8742	8750	8756	8849
1,3	7913	8384	8578	8683	8748	8793	8826	8851	8870	8886	8899	8910	8919	8927	8934	9032
1,4	8026	8518	8720	8829	8898	8945	8979	9005	9025	9041	9055	9066	9075	9084	9091	9192
1,5	8128	8638	8847	8960	9030	9079	9114	9140	9161	9177	9191	9203	9212	9221	9228	9332
1,6	8222	8746	8960	9076	9148	9196	9232	9259	9280	9297	9310	9322	9332	9340	9348	9452
1,7	8307	8844	9062	9178	9251	9300	9335	9362	9383	9400	9414	9426	9435	9444	9451	9554
1,8	8386	8932	9152	9269	9341	9390	9426	9452	9473	9490	9503	9515	9525	9533	9540	9641
1,9	8458	9011	9232	9349	9421	9469	9504	9530	9551	9567	9580	9591	9601	9609	9616	9713
2	8524	9082	9303	9419	9490	9538	9572	9597	9617	9633	9646	9657	9666	9674	9680	9772
2,1	8585	9147	9367	9482	9551	9598	9631	9655	9674	9690	9702	9712	9721	9728	9735	9821
2,2	8642	9206	9424	9537	9605	9649	9681	9705	9723	9738	9750	9759	9768	9774	9781	9861
2,3	8695	9259	9475	9585	9651	9694	9725	9748	9765	9779	9790	9799	9807	9813	9819	9893
2,4	8743	9308	9521	9628	9692	9734	9763	9784	9801	9813	9824	9832	9840	9846	9851	9918
2,5	8789	9352	9561	9666	9728	9767	9795	9815	9831	9843	9852	9860	9867	9873	9877	9938
2,6	8831	9392	9598	9700	9759	9797	9823	9842	9856	9868	9877	9884	9890	9895	9900	9953
2,7	8871	9429	9631	9730	9786	9822	9847	9865	9878	9888	9897	9903	9909	9914	9918	9965
2,8	8908	9463	9661	9756	9810	9844	9867	9884	9896	9906	9914	9920	9925	9929	9933	9974
2,9	8943	9494	9687	9779	9831	9863	9885	9901	9912	9921	9928	9933	9938	9942	9945	9981
3	8976	9523	9712	9800	9850	9880	9900	9915	9925	9933	9940	9945	9949	9952	9955	9987
3,1	9007	9549	9734	9819	9866	9894	9913	9927	9936	9944	9949	9954	9958	9961	9963	9990
3,2	9036	9573	9753	9835	9880	9907	9925	9937	9946	9953	9958	9962	9965	9968	9970	9993
3,3	9063	9596	9771	9850	9893	9918	9934	9946	9954	9960	9965	9968	9971	9974	9976	9995
3,4	9089	9617	9788	9864	9904	9928	9943	9953	9961	9966	9970	9974	9976	9978	9980	9997
3,5	9114	9636	9803	9876	9914	9936	9950	9960	9966	9971	9975	9978	9980	9982	9984	9998
3,7	9160	9670	9829	9896	9930	9950	9962	9970	9975	9979	9982	9985	9987	9988	9989	9999
4	9220	9714	9860	9919	9948	9964	9974	9980	9984	9987	9990	9991	9992	9993	9994	10000

В таблице приведены значения $\breve{P}_k\left(\hat{t}\right)\cdot 10^4$.

ПРИЛОЖЕНИЕ 4

Распределение хи–квадрат

Закон распределения $\chi^2: V_k\left(\hat{w}\right) = P(w \geq \hat{w}) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} \int\limits_{\hat{w}}^{\infty} w^{\frac{k}{2}-1} e^{-\frac{w}{2}}\, dw,$

$\hat{w}$ \ k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0,5	4795	7788	9189	9735	9921	9978	9994	9999	10000	10000	10000	10000	10000	10000	10000	10000
1	3173	6065	8013	9098	9626	9856	9948	9982	9994	9998	9999	10000	10000	10000	10000	10000
1,5	2207	4724	6823	8266	9131	9595	9823	9927	9971	9989	9996	9999	10000	10000	10000	10000
2	1573	3679	5724	7358	8491	9197	9598	9810	9915	9963	9985	9994	9998	9999	10000	10000
2,5	1138	2865	4753	6446	7765	8685	9271	9617	9809	9909	9958	9982	9992	9997	9999	10000
3	833	2231	3916	5578	7000	8088	8850	9344	9643	9814	9907	9955	9979	9991	9996	9998
3,5	614	1738	3208	4779	6234	7440	8352	8992	9411	9671	9823	9909	9954	9978	9990	9995
4	455	1353	2615	4060	5494	6767	7798	8571	9114	9473	9699	9834	9912	9955	9977	9989
4,5	339	1054	2123	3425	4799	6093	7207	8094	8755	9220	9529	9726	9846	9916	9956	9977
5	253	821	1718	2873	4159	5438	6600	7576	8343	8912	9312	9580	9752	9858	9921	9958
5,5	190	639	1386	2397	3579	4815	5992	7030	7887	8554	9046	9392	9625	9776	9870	9927
6	143	498	1116	1991	3062	4232	5397	6472	7399	8153	8734	9161	9462	9665	9797	9881
6,5	108	388	897	1648	2606	3696	4827	5914	6890	7717	8380	8888	9261	9523	9701	9817
7	82	302	719	1359	2206	3208	4289	5366	6371	7254	7991	8576	9022	9347	9576	9733
7,5	62	235	576	1117	1860	2771	3787	4838	5852	6775	7573	8229	8746	9137	9423	9624
8	47	183	460	916	1562	2381	3326	4335	5341	6288	7133	7851	8436	8893	9238	9489
8,5	36	143	367	749	1307	2037	2906	3862	4846	5801	6679	7449	8096	8617	9022	9326
9	27	111	293	611	1091	1736	2527	3423	4373	5321	6219	7029	7729	8311	8775	9134
9,5	21	87	233	497	907	1473	2187	3019	3925	4854	5758	6597	7342	7978	8500	8914
10	16	67	186	404	752	1247	1886	2650	3505	4405	5304	6160	6939	7622	8197	8666
10,5	12	52	148	328	622	1051	1620	2317	3115	3978	4860	5722	6526	7248	7872	8392
11	9	41	117	266	514	884	1386	2017	2757	3575	4433	5289	6108	6860	7526	8095
11,5	7	32	93	215	423	741	1182	1749	2430	3199	4024	4866	5690	6464	7164	7776
12	5	25	74	174	348	620	1006	1512	2133	2851	3636	4457	5276	6063	6790	7440
12,5	4	19	59	140	285	517	853	1303	1866	2530	3273	4064	4871	5662	6409	7089
13	3	15	46	113	234	430	721	1118	1626	2237	2933	3690	4478	5265	6023	6728
13,5	2	12	37	91	191	357	608	958	1413	1970	2619	3338	4100	4876	5637	6359
14	2	9	29	73	156	296	512	818	1223	1730	2330	3007	3738	4497	5255	5987
14,5	1	7	23	59	127	245	430	696	1056	1514	2065	2699	3396	4132	4880	5615
15	1	6	18	47	104	203	360	591	909	1321	1825	2414	3074	3782	4514	5246
15,5	1	4	14	38	84	167	301	501	781	1149	1607	2152	2772	3449	4160	4884
16	1	3	11	30	68	138	251	424	669	996	1411	1912	2491	3134	3821	4530
17	0	2	7	19	45	93	174	301	487	744	1079	1496	1993	2562	3189	3856
18	0	1	4	12	29	62	120	212	352	550	816	1157	1575	2068	2627	3239
19	0	1	3	8	19	42	82	149	252	403	611	885	1231	1649	2137	2687
20	0	0	2	5	12	28	56	103	179	293	453	671	952	1301	1719	2202
25	0	0	0	1	1	3	8	16	30	53	91	148	231	346	499	698
30	0	0	0	0	0	0	1	2	4	9	16	28	47	76	119	180

В таблице приведены значения $V_k\left(\hat{w}\right) \cdot 10^4$.

Литература

1, *Неделько В.М., Ступина Т.А.* Основы теории вероятностей и математической статистики в примерах и задачах: учеб. пособие — Новосибирск: Изд–во НГУ, 2006, — 82с.
2, *Неделько В.М.* Основы теории вероятностей. Учебное пособие, НГТУ, 2011, — 116 с.
3, *Миренкова Г.Н., Неделько С.В., Неделько В.М., Тренёва Т.В.* Основы математической статистики: учеб.–метод. пособие, — Новосибирск, Изд–во НГТУ, 2008, — 36с.
4, *Бородихин В.М., Ковалевский А.П.* Высшая математика, Новосиб. гос. техн. ун–т, – Т. 4.2. Теория вероятностей и математическая статистика: учеб. пособие — Новосибирск, 2005, — 256с.
5, *Гмурман В.Е.* Теория вероятностей и математическая статистика. —М,: Высш. Шк., 2005, — 480с.