

Министерство образования Российской Федерации
“МАТИ” – Российский Государственный Технологический
Университет им. К. Э. Циолковского.

Кафедра «Высшая математика»

Статистическая обработка результатов измерений

Методические указания к курсовому проектированию по курсу
«Математическая статистика»

Составители: И. Ф. Заварзина
И. А. Данилина
А. С. Ионова

Москва 2001

I. Задачи математической статистики

При изучении курса теории вероятностей предполагалось, что вероятности наступления отдельных событий известны. Считались известными законы распределения случайных величин или их числовые характеристики. Как правило, на практике вероятности наступления событий, законы распределения случайных величин или параметры этих законов распределения неизвестны. Для их определения (оценивания) необходимо производить эксперимент, специальные испытания.

При обработке эксперимента статистическими методами основные понятия теории вероятностей выступают как некоторые модели реальных закономерностей.

Основой статистических методов являются экспериментальные данные, часто называемые статистическими данными.

Одним из основных методов статистического наблюдения является выборочный метод. Рассмотрим основные понятия этого метода.

II. Генеральная и выборочная совокупность

Пусть для исследования закономерностей случайного явления произведено n опытов, в результате которых получен ряд наблюдений x_1, x_2, \dots, x_n . Требуется обработать этот ряд статистически. Для этого надо вначале построить математическую модель ряда наблюдений, т.е. указать, какие величины случайны, какие не случайны, какие зависимы, какие не зависимы и т.д.

Ставится задача оценить функцию распределения $F(x)$ исследуемой СВ X , т.е. построить уточненную вероятностную модель ряда наблюдений x_1, x_2, \dots, x_n , которая бы отражала в себе основные статистические особенности этого ряда.

Наиболее точные сведения о случайной величине X можно получить, производя максимально возможное количество измерений этой случайной величины.

Определение 1. *Генеральной совокупностью называется совокупность всех мыслимых наблюдений, которые могли бы быть сделаны при данном реальном комплексе условий измерений. Число членов, входящих в генеральную совокупность, называют объемом генеральной совокупности.*

Определение 2. *Выборочной совокупностью или просто выборкой объема n называется совокупность n объектов, отобранных из исследуемой генеральной совокупности.*

Определение 3. *Метод, состоящий в том, что на основании характеристик и свойств выборки x_1, x_2, \dots, x_n делаются заключения о*

числовых характеристиках и законе распределения СВ X, называется выборочным методом.

Для того чтобы сведения о законах распределения СВ X были объективны, необходимо, чтобы выборка была репрезентативной, т.е. представительной. Существуют специальные методы для этого.

III. Статистический ряд. Статистический закон распределения случайной величины

Предположим, что изучается дискретная или непрерывная случайная величина, закон распределения которой неизвестен. Для оценки закона распределения этой случайной величины и его числовых характеристик производится ряд независимых измерений x_1, x_2, \dots, x_n . Статистический материал представляют в виде таблицы, состоящей из двух строк, в первой из которых даны номера измерений, а во второй – результаты измерений.

i – номер измерения	1	2	...	
x_i – результат измерений	x_1	x_2	...	x_n

Такую таблицу называют простым статистическим рядом.

Для того чтобы правильно оценить закон распределения СВ X, производят группировку данных. Если X – дискретная СВ, то наблюдаемые значения располагаются в порядке возрастания и подсчитываются частоты m_i или частоты m_i/n появления одинаковых значений СВ X. В результате получаем сгруппированные статистические ряды:

x_i	x_1	x_2	...	x_k
m_i	m_1	m_2	...	m_k

Контроль: $\sum_{i=1}^k m_i = n$.

x_i	x_1	x_2	x_n
m_i/n	m_1/n	m_2/n	m_k/n

Контроль: $\sum_{i=1}^k m_i/n = 1$.

Если изучается непрерывная случайная величина, то группировка заключается в разбиении интервала наблюдаемых значений случайной величины на k частичных интервалов равной длины $[x_0; x_1 [$, $[x_1; x_2 [$, $[x_2; x_3 [$, $[x_{k-1}; x_k]$ и подсчете частоты или частоты m_i/n попадания наблюдаемых

значений в частичные интервалы. Количество интервалов выбирается произвольно, обычно не меньше 5 и не больше 15.

В результате составляется интервальный статистический ряд следующего вида:

СВХ	$[x_0; x_1 [$	$[x_1; x_2 [$	$[x_{k-1}; x_k]$
m_i/n	m_1/n	m_2/n	m_k/n

Контроль: $\sum_{i=1}^k m_i/n = 1.$

Определение. *Перечень наблюдаемых значений СВ X (или интервалов наблюдаемых значений) и соответствующих им частот m_i/n называется статистическим законом распределения случайной величины.*

Статистические законы позволяют визуально произвести оценку закона распределения исследуемой случайной величины.

IV. Эмпирическая функция распределения

Эмпирической функцией распределения случайной величины X называют функцию $F^*(x)$, определяющую для каждого значения x частоту события ($X < x$):

$$F^*(x) = n_x/n;$$

где n_x – число x_i , меньших x; n – объем выборки.

Из теоремы Бернулли следует, что при достаточно большом объеме выборки функции $F^*(x)$ и $F(x) = P(X < x)$ мало отличаются друг от друга.

Эмпирическая функция распределения обладает всеми свойствами интегральной функции распределения:

- 1) значения эмпирической функции $F^*(x)$ принадлежат отрезку $[0, 1]$;
- 2) $F^*(x)$ – неубывающая функция;
- 3) если x_1 – наименьшее, а x_n – наибольшее наблюдаемое значение, то $F^*(x) = 0$ при $x < x_1$ и $F^*(x) = 1$ при $x > x_n$.

Основное значение эмпирической функции распределения состоит в том, что она используется в качестве оценки функции распределения

$$F(x) = P(X < x).$$

Пример. Построить $F^*(x)$ по статистическому распределению СВ X:

x_i	2	3	5
m_i/n	0.75	0.20	0.05

Решение. Относительная частота события ($X < x$) равна $F^*(x)$. Следовательно,

$$F^*(x) = \begin{cases} 0 & x \leq 2; \\ 0.75 & \text{при } 2 < x \leq 3; \\ 0.95 & \text{при } 3 < x \leq 5; \\ 1 & x > 5. \end{cases}$$

График $F^*(x)$ изображен на рисунке 1.

Для наглядности сгруппированные статистические ряды изображают в виде графиков и диаграмм. Наиболее распространенными графиками являются полигон и гистограмма. Полигон применяется для изображения как дискретных, так и интервальных статистических рядов, гистограмма - для изображения только интервальных рядов.

Пример. Результаты исследования прочности 200 образцов бетона на сжатие представлены в виде интервального статистического ряда.

интервалы прочности кг/см ²	частоты m_i	частости m_i/n
190 – 200	10	0.05
200 – 210	26	0.13
210 – 220	56	0.28
220 – 230	64	0.32
230 – 240	30	0.15
240 – 250	14	0.07

$$n = \sum_i m_i = 200, \quad \sum_i m_i/n = 1.$$

На рисунке 2 представлена гистограмма. На оси абсцисс откладываются частичные интервалы наблюдаемых значений случайной величины X , на каждом из которых строим прямоугольник, площадь которого равна частости данного частичного интервала. Высота элементарного прямоугольника частостей равна m_i/nh , где h – длина интервала.

Если на гистограмме частостей соединить середины верхних сторон прямоугольников, то полученная замкнутая ломаная линия образует полигон распределения частостей.

V. Основные законы распределения случайных величин, используемых в математической статистике

П.1. Нормальное распределение

Нормальная модель распределения вероятностей играет исключительно важную роль в теории вероятностей и математической статистике. Главная особенность нормального распределения состоит в том, что оно является предельным, к которому приближаются другие распределения при соблюдении некоторых условий.

Нормальные распределения часто встречаются на практике в самых различных областях. Принято считать, что все ошибки измерений, вес деталей, размер деталей, дальность полета артиллерийского снаряда и многие другие случайные величины имеют нормальное распределение.

Нормальное распределение задается функцией плотности вероятности:

$$f(x) = \frac{1}{s \sqrt{2\pi}} e^{-\frac{(x-a)^2}{2s^2}}, \quad (5.1)$$

где a – математическое ожидание случайной величины X , т.е. $M(X) = a$;

σ – среднее квадратичное отклонение СВ X , т.е. $\sqrt{D(X)} = \sigma$

($D(X)$ – дисперсия случайной величины).

Из формулы (5.1) видно, что нормальная модель зависит от двух параметров a и σ , поэтому ее называют двухпараметрической моделью распределения.

Если случайная величина X имеет нормальное распределение с параметрами $M(X) = a$ и $s = \sqrt{D(X)}$, то этот факт кратко записывают с помощью символической записи: СВ $X \in N(a, \sigma)$.

График функции плотности вероятности называют нормальной кривой или кривой Гаусса.

Эта кривая изображена на рисунке 3.

1⁰. $f(x)$ определена при всех $x \in \mathbb{R}$.

2⁰. Кривая нормального распределения симметрична относительно прямой $x = a$.

3⁰. Кривая Гаусса имеет максимум в точке $x = a$:

$$f(a) = \frac{1}{s \sqrt{2\pi}} .$$

4⁰. Кривая Гаусса имеет две точки перегиба:

$$x_1 = a - \sigma \text{ и } x_2 = a + \sigma.$$

5⁰. Площадь, заключенная между кривой Гаусса и осью абсцисс, равна 1; между осью абсцисс, кривой Гаусса и прямыми $a \pm 2\sigma$ равна $\approx 0,95$.

6⁰. При увеличении (уменьшении) параметра σ максимальная ордината уменьшается (увеличивается), см. рис. 4. Другими словами, параметр σ характеризует форму кривой, при неизменном положении центра кривой; так как площадь под кривой Гаусса всегда равна 1 ($\int_{-\infty}^{+\infty} f(x)dx=1$), то, если σ

увеличивается, то кривая становится плоско – вершинной, σ уменьшается – кривая Гаусса вытягивается вверх. Параметр σ иногда называют параметром масштаба.

7⁰. Если изменять математическое ожидание a при неизменном σ , то кривая Гаусса будет смещаться вдоль оси абсцисс, т.е. параметр $a = M(X)$ характеризует положение кривой при неизменной форме. Иногда параметр a называют параметром сдвига (см. рис. 5).

Если СВ $X \in N(a, \sigma)$, то случайная величина $U = \frac{x-a}{s}$ имеет нормальное распределение с параметром $M(U)=0$ и $\sigma(U) = 1$, т.е. $U \in N(0,1)$. Поэтому случайную величину $U = \frac{x-a}{s}$ называют нормированной или стандартизованной нормальной величиной. Плотность распределения вероятностей нормированной случайной величины U имеет вид:

$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \quad (5.2)$$

Функция распределения СВ $X \in N(a, \sigma)$ имеет следующий вид:

$$F(x) = \int_{-\infty}^x f(x) dx = \frac{1}{s \sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(x-a)^2}{2s^2}} dx. \quad (5.3)$$

Функция распределения нормализованной случайной величины

$$F(u) = P(U < u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-t^2/2} dt.$$

Для облегчения вычисления вероятности попадания СВ $X \in N(a, \sigma)$ в интервал $]\alpha, \beta[$ вводится нормированная функция Лапласа:

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$$

Тогда

$$P(a < x < b) = F(b) - F(a) \Leftrightarrow P(a < x < b) = \frac{1}{2} \left(\Phi\left(\frac{b-a}{s}\right) - \Phi\left(\frac{a-a}{s}\right) \right).$$

Используя нормированную функцию Лапласа, можно записать функцию распределения СВ $X \in N(a, \sigma)$ в виде:

$$F(x) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x-a}{s}\right).$$

П.2. Распределение χ^2 (хи – квадрат)

Рассмотрим случайную величину Y , распределенную по нормальному закону

$Y \in N(a, \sigma)$. Тогда случайная величина $U = \frac{Y-a}{s} = c^2$ распределена по нормальному закону с параметрами $M(U) = 0$ и $\sigma(U) = 1$, т.е. $U \in N(0, 1)$.

Квадрат такой стандартизованной случайной величины

$$U^2 = \left(\frac{Y-a}{s}\right)^2 = \chi^2$$

называется случайной величиной χ^2 (хи – квадрат) с одной степенью свободы.

Рассмотрим n независимых случайных величин Y_1, Y_2, \dots, Y_n , распределенных по нормальному закону с $M(Y_i) = a_i$ и средними квадратическими отклонениями $\sigma_i, i = \overline{1, n}$.

Образует для каждой из этих случайных величин стандартизованную случайную величину

$$U_i = \frac{Y_i - a_i}{s_i}, i = \overline{1, n}.$$

Сумма квадратов стандартизованных переменных

$c^2 = U_1^2 + U_2^2 + \dots + U_n^2 = \left(\frac{Y_1 - a_1}{s_1}\right)^2 + \left(\frac{Y_2 - a_2}{s_2}\right)^2 + \dots + \left(\frac{Y_n - a_n}{s_n}\right)^2$ называется случайной величиной χ^2 с $\nu = n$ степенями свободы.

Плотность распределения СВ χ^2 имеет вид:

$$f(\chi^2) = \begin{cases} 0, \\ \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}}, \text{ если } \chi^2 \geq 0. \end{cases}$$

Итак, распределение χ^2 зависит от одного параметра ν - числа степеней свободы.

Функция распределения χ^2 имеет вид:

$$F(\chi^2) = P(\chi^2 < \chi_0^2) = \begin{cases} 0, \\ \frac{1}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \int_0^{\chi_0^2} (\chi^2)^{\frac{\nu}{2}-1} e^{-\frac{\chi^2}{2}} d(\chi^2), \text{ если } \chi^2 \geq 0. \end{cases}$$

На рис. 6 и 7 изображены графики плотности вероятности и функции χ^2 - распределения.

В практике, как правило, используются не $f(\chi^2)$ и $F(\chi^2)$, а квантили χ^2 - распределения $C_{a,n}^2$. Квантилем $C_{a,n}^2$, отвечающим заданному уровню вероятности α , называется такое значение $\chi^2 = C_{a,n}^2$, при котором

$$P(\chi^2 > C_{a,n}^2) = \int_{C_{a,n}^2}^{\infty} f(\chi^2) d(\chi^2) = \alpha.$$

Нахождение квантиля, с геометрической точки зрения, заключается в том, чтобы выбрать такое значение $\chi^2 = C_{a,n}^2$, при котором площадь заштрихованной криволинейной трапеции (см. рис 6) была бы равна α .

II. 3. Распределение Стьюдента

Распределение Стьюдента (t - распределение) имеет важное значение при статистических вычислениях, связанных с нормальным законом, а именно тогда, когда среднее квадратическое отклонение σ неизвестно и подлежит определению по опытным данным.

Пусть Y, Y_1, Y_2, \dots, Y_n - независимые случайные величины, имеющие нормальное распределение с параметрами $M(Y) = M(Y_i) = 0$ и $\sigma_Y = \sigma_{Y_i} = 1$, $i = \overline{1, n}$.

Случайная величина

$$t = \frac{Y}{\sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2}} = \frac{Y}{\sqrt{\frac{1}{n} C_n^2}}, \quad (5.4)$$

являющаяся функцией нормально распределенных случайных величин, называется безразмерной дробью Стьюдента.

Плотность распределения случайной величины t имеет вид:

$$f(t) = S(t, n) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{pn} \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < +\infty \quad (5.5),$$

где v - число слагаемых в подкоренном выражении дроби Стьюдента, т.е. $v = n$. Такое обозначение числа степеней свободы общепринято в математической статистике.

Из формулы (5.5) видно, что распределение СВ t зависит только от одного параметра – числа степеней свободы v , равного числу слагаемых в подкоренном выражении дроби Стьюдента (5.4).

Известно, что математическое ожидание и дисперсия СВ t соответственно равны

$$M(t) = 0; \quad D(t) = \frac{n}{n-2}; \quad (n > 2).$$

На рис. 8 изображен график плотности распределения Стьюдента при различных степенях свободы. Замечаем, что при увеличении числа степеней свободы v он приближается к кривой Гаусса.

В статистических расчетах используются квантили t – распределения $t_{\frac{a}{2}; n}$. Значения квантилей находятся из решения уравнения:

$$P(|t| > t_{\frac{a}{2}; n}) = 2 \int_{t_{\frac{a}{2}; n}}^{\infty} f(t) dt = a$$

С геометрической точки зрения, нахождение квантилей $t_{\frac{a}{2}; n}$ заключается в том выборе значения $t = t_{\frac{a}{2}; n}$, при котором суммарная площадь заштрихованных на рис. 9 криволинейных трапеций была бы равна α .

VI. Точечные оценки параметров нормального распределения

Пусть СВ X имеет нормальное распределение: $X \in N(a, \sigma)$. Параметры a , σ нормального распределения, как правило, неизвестны. С целью их определения производится эксперимент, в результате которого фиксируется n значений случайной величины X : x_1, x_2, \dots, x_n .

Результаты измерения x_1, x_2, \dots, x_n рассматривают как выборку объема n из бесконечной генеральной совокупности. На основании этой выборки необходимо «оценить» (найти приближенные значения) двух параметров – математического ожидания a и среднего квадратического отклонения σ .

Вообще говоря, по результатам выборки, какого бы большого размера она ни была, нельзя определить точные значения неизвестных параметров a и σ , но можно найти их приближенные значения $\hat{a}, \hat{\sigma}$, которые называются оценками.

Для нахождения приближенных значений $\hat{a}, \hat{\sigma}$, неизвестных параметров a и σ нормального закона будем рассматривать функции вида:

$\hat{a} = \hat{a}(x_1, x_2, \dots, x_n)$, $\hat{\sigma} = \hat{\sigma}(x_1, x_2, \dots, x_n)$, которые называются выборочными функциями или статистиками.

Задача оценки неизвестных параметров a и σ сводится к нахождению таких статистик $\hat{a} = \hat{a}(x_1, x_2, \dots, x_n)$, $\hat{\sigma} = \hat{\sigma}(x_1, x_2, \dots, x_n)$, которые могут быть использованы для приближенного определения значений неизвестных параметров a и σ .

Оценки параметров подразделяются на точечные и интервальные. Точечная оценка параметра θ (где под θ будем понимать либо a , либо σ) определяется одним числом $\hat{q} = \hat{q}(x_1, x_2, \dots, x_n)$.

Интервальной оценкой называют оценку, которая определяется двумя числами \hat{q}_1 и \hat{q}_2 – концами интервала, накрывающего оцениваемый параметр θ .

Можно показать, что если СВ $X \in N(a, \sigma)$, то точечные оценки неизвестных параметров a и σ находятся по формулам:

$$\hat{a} = M(x) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}; \quad (6.1),$$

$$\hat{\sigma} = S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}; \quad (6.2).$$

Эти оценки обладают свойствами несмещенности, состоятельности и эффективности.

VII. Интервальные оценки параметров нормального распределения

Пусть $X \in N(a, \sigma)$, причем a и σ неизвестны. Для нахождения точечных оценок a и σ из генеральной совокупности извлечена выборка объемом n . Пусть на основании этой выборки найдены точечные несмещенные оценки неизвестных параметров a и σ по формулам (6.1) и (6.2).

Точечные оценки, найденные по выборке объемом n , не позволяют непосредственно ответить на вопрос, какую ошибку мы допускаем, принимая вместо точного значения неизвестного параметра a или σ его приближенные значения $\hat{a}, \hat{\sigma}$.

Поэтому во многих случаях выгоднее пользоваться интервальной оценкой, основанной на определении некоторого интервала, внутри которого с определенной вероятностью находится неизвестное значение параметра a (или σ).

Пусть найденная по результатам выборки объема n статистическая характеристика $\hat{q} = \hat{q}(x_1, x_2, \dots, x_n)$ является точечной оценкой неизвестного параметра θ . Чем меньше разность $|q - \hat{q}|$, тем лучше качество оценки, тем она точнее. Таким образом, положительное число ϵ характеризует точность оценки

$$|q - \hat{q}| < \epsilon. \quad (7.1)$$

Однако статистический метод не позволяет категорически утверждать, что оценка удовлетворяет неравенству (7.1) в смысле математического анализа. Можно только говорить о вероятности $(1-\alpha)$, с которой это неравенство выполняется.

Доверительной вероятностью оценки называют вероятность $(1-\alpha)$ выполнения неравенства $|q - \hat{q}| < \epsilon$. Обычно доверительная вероятность

оценки задается заранее. Наиболее часто полагают $(1-\alpha) = 0,95; 0,99; 0,9973$. Доверительная вероятность точечной оценки показывает, что при извлечении выборки объема n из одной и той же генеральной совокупности в $(1-\alpha)$ 100% случаях параметр θ будет покрываться данным интервалом.

Пусть вероятность того, что $|q - \hat{q}| < e$ равна $(1-\alpha)$

$$P(|q - \hat{q}| < e) = 1 - \alpha. \quad (7.2)$$

Преобразуем формулу (7.2)

$$P(\hat{q} - e < q < \hat{q} + e) = 1 - \alpha. \quad (7.3)$$

Последняя формула показывает, что неизвестный параметр θ заключен внутри интервала $[\hat{q} - e, \hat{q} + e]$. Этот интервал называется доверительным.

Итак, доверительный интервал $[\hat{q} - e, \hat{q} + e]$ покрывает неизвестный параметр θ с заданной надежностью $(1-\alpha)$.

В практических приложениях важную роль играет длина доверительного интервала. Чем меньше длина доверительного интервала $(\hat{q} - e, \hat{q} + e)$, тем точнее оценка.

Из формулы (7.3) длина доверительного интервала равна $2e$. Из этой формулы видно, что длина доверительного интервала $2e$ определяется двумя величинами: доверительной вероятностью $(1-\alpha)$ и объемом выборки n . Таким образом, e , $(1-\alpha)$ и n тесно взаимосвязаны и, задавая определенные значения двум из них, можно определить величину третьей.

Если σ известно, то доверительный интервал, покрывающий неизвестное математическое ожидание с заданной доверительной вероятностью $(1-\alpha)$, имеет следующий вид:

$$\bar{x} - u_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < a < \bar{x} + u_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \quad (7.4)$$

где \bar{x} - средняя арифметическая результатов измерений; $\left(\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \right)$

n – объем выборки;

$u_{\frac{\alpha}{2}}$ – квантиль нормированного нормального распределения, определяемый по

доверительной вероятности $(1-\alpha)$;

$e = u_{\frac{\alpha}{2}} \frac{d}{\sqrt{n}}$ – точность (предельная погрешность) точечной оценки математического ожидания.

Для наиболее употребительных значений доверительной вероятности $(1-\alpha)$ квантили стандартизованного нормального распределения приведены в сокращенной таблице:

$(1-\alpha)$	$u_{\frac{\alpha}{2}}$
0,90	1,64
0,95	1,96
0,99	2,58
0,9973	3,00
0,999	3,37

Анализируя формулу доверительного интервала, задаваемого системой неравенств (7.4), можно заметить, что:

- а) увеличение объема выборки n приводит к уменьшению длины доверительного интервала;
- б) увеличение доверительной вероятности $(1-\alpha)$ приводит к увеличению длины доверительного интервала, т.е. к уменьшению точности $\varepsilon = u_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$;
- в) если задать точность ε и доверительную вероятность $(1-\alpha)$, то из соотношения $\varepsilon = u_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$ можно найти минимальный объем выборки, который обеспечивает заданную точность.

Если же σ неизвестно, тогда доверительный интервал, накрывающий неизвестное математическое ожидание a СВ $X \in N(a, \sigma)$, имеет следующий вид:

$$\bar{x} - t_{\frac{\alpha}{2}; n-1} \frac{S}{\sqrt{n}} < a < \bar{x} + t_{\frac{\alpha}{2}; n-1} \frac{S}{\sqrt{n}}, \quad (7.5)$$

где $t_{\frac{\alpha}{2}; n-1}$ - квантиль распределения Стьюдента, определяемый по таблицам

по заданной доверительной вероятности $P = (1-\alpha)$ и числу степеней свободы $\nu = n-1$ (n – объем выборки);

\bar{x}, S - точечные несмещенные оценки параметров нормального распределения;

$e = t_{\frac{\alpha}{2}; n-1} \frac{S}{\sqrt{n}}$ - предельная погрешность точечного оценивания математического ожидания СВ $X \in N(a, \sigma)$ при неизвестном σ обладает теми же свойствами, что и при известном σ .

Доверительный интервал для среднего квадратического отклонения σ задается системой неравенств

$$S \sqrt{\frac{n-1}{c_{\frac{\alpha}{2}; n-1}^2}} < \sigma < S \sqrt{\frac{n-1}{c_{1-\frac{\alpha}{2}; n-1}^2}}, \quad (7.6)$$

где $c_{\frac{\alpha}{2}; n-1}^2$; $c_{1-\frac{\alpha}{2}; n-1}^2$ - квантили χ^2 распределения, определенные по таблице распределения χ^2 по заданной доверительной вероятности $(1-\alpha)$ и числу степеней свободы $\nu = n-1$.

Значение величин $g_1 = \sqrt{\frac{n-1}{C_{\frac{a}{2}; n-1}^2}}$ и $g_2 = \sqrt{\frac{n-1}{C_{1-\frac{a}{2}; n-1}^2}}$ приведены в таблице.

VIII. Примеры обработки результатов эксперимента

Измерена максимальная емкость шести конденсаторов, выбранных из большого числа конденсаторов.

Получены следующие результаты (в пф) 4,45; 4,40; 4,42; 4,45; 4,38; 4,42. Предполагая, что результаты измерений имеют нормальное распределение, требуется:

- 1) найти точечные несмешанные оценки математического ожидания и среднего квадратического отклонения;
- 2) записать плотность вероятности и функцию распределения СВ X (емкости конденсаторов);
- 3) найти доверительный интервал, накрывающий математическое ожидание емкости конденсаторов с заданной доверительной вероятностью $(1-\alpha) = 0,95$, считая σ неизвестной;
- 4) найти доверительный интервал, накрывающий неизвестное среднее квадратичное отклонение σ с заданной доверительной вероятностью $(1-\alpha) = 0,95$;
- 5) принимая доверительную вероятность $P = 1-\alpha = 0,99$, найти предельную погрешность, с которой $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ оценивает математическое ожидание a емкости конденсаторов;
- 6) найти минимальное число конденсаторов, емкость которых надо измерить, чтобы с доверительной вероятностью $(1-\alpha) = 0,95$ можно было бы утверждать, что, принимая среднее арифметическое \bar{x} за математическое ожидание емкости конденсаторов, мы совершаем погрешность, не превышающую $\varepsilon = 0,5\sigma$, считая $\sigma = S$;
- 7) вычислить $P(4,41 < x < 4,43)$.

Решение:

$$1) \hat{a} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{6} \sum_{i=1}^6 x_i = 4.42 \text{ (пф – микрофарад)}$$

$$\hat{s} = S = \sqrt{\frac{\sum_{i=1}^6 (x_i - \bar{x})^2}{6-1}} = 0.028 \text{ пф.}$$

2) Следовательно, плотность вероятности СВ X (емкость конденсаторов) имеет вид:

$$f(x) = \frac{1}{0.028\sqrt{2p}} \exp\left(-\frac{(x-4.42)^2}{0.016}\right).$$

Функция распределения емкости конденсаторов имеет вид:

$$F(x) = \int_{-\infty}^x f(x) dx = \frac{1}{0.028\sqrt{2p}} \int_{-\infty}^x \exp\left(-\frac{(x-4.42)^2}{0.016}\right) dx.$$

Используя нормированную функцию Лапласа

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt,$$

можно записать

$$F(x) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x-\bar{x}}{S}\right) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x-4.42}{0.028}\right).$$

- 3) Найдем интервальные оценки параметров нормального распределения емкости конденсаторов. Для нахождения доверительного интервала, накрывающего математическое ожидание, найдем по таблице квантилей распределение Стьюдента по заданной доверительной вероятности $P = 1-\alpha = 0,95$ и числу степеней свободы $\nu = n-1 = 6-1 = 5$ квантиль $t_{\frac{\alpha}{2}; n} = t_{0.025; 5} = 2.571$.

Вычислим предельную погрешность интервального оценивания математического ожидания

$$e = t_{\frac{\alpha}{2}; n} \frac{S}{\sqrt{n}} = 2.571 \frac{0.028}{\sqrt{6}} = 0.029 \text{ (пф).}$$

Искомый доверительный интервал, накрывающий математическое ожидание емкости конденсаторов с заданной доверительной вероятностью $P = 0,95$, равен:

$$\begin{aligned} \bar{x} - e < a < \bar{x} + e; \\ 4,42 - 0,029 < a < 4,42 + 0,029; \\ 4,391 < a < 4,449. \end{aligned}$$

Смысл полученного результата:

если будет произведено достаточно большое число выборок по 6 конденсаторов из бесконечно большой по численности партии конденсаторов, то в 95% случаев из них доверительный интервал накроет неизвестное математическое ожидание и только в 5% математическое ожидание может выйти за границы доверительного интервала.

- 4) Для нахождения доверительного интервала, накрывающего неизвестное среднее квадратическое отклонение σ с заданной

доверительной вероятностью $(1-\alpha) = 0,95$, найдем по заданной доверительной вероятности $0,95$ и числу степеней свободы $\nu = n-1 = 6-1 = 5$ два числа γ_1 и γ_2 , т.е. $\gamma_1 = 0,624$ и $\gamma_2 = 2,45$. Искомый доверительный интервал равен:

$$\gamma_1 S < \sigma < \gamma_2 S;$$

$$0.624 \cdot 0.028 < \sigma < 2.45 \cdot 0.028;$$

$$0.017 < \sigma < 0.068.$$

5) Если задать доверительную вероятность $P = 1-\alpha = 0,99$, то предельная погрешность, с которой среднее арифметическое емкости конденсаторов \bar{x} оценивает неизвестное математическое ожидание, равна:

$$e = t_{\frac{\alpha}{2}; n} \frac{S}{\sqrt{n}} = t_{0.005} \frac{0.028}{\sqrt{6}} = 4.032 \frac{0.028}{\sqrt{6}} = 0.046.$$

6) Найдем минимальное число конденсаторов, емкость которых необходимо измерить, чтобы с доверительной вероятностью $P = 1-\alpha = 0,95$ можно было бы утверждать, что, принимая среднее арифметическое \bar{x} за математическое ожидание емкости конденсаторов, мы совершаем погрешность, не превышающую $0,2\sigma = 0,0056$, считая σ известным и равным $0,028$.

Искомый объем выборки найдем из соотношения

$$n = \frac{S^2 u_a^2}{e^2} = \frac{0.028^2 (1.96)^2}{(0.0056)^2} \geq 96 \text{ конденсаторов.}$$

$$\begin{aligned} 7) P(4.41 < X < 4.43) &= \frac{1}{2} \left[\Phi\left(\frac{4.43 - 4.42}{0.028}\right) - \Phi\left(\frac{4.41 - 4.42}{0.028}\right) \right] = \\ &= \frac{1}{2} [\Phi(0.357) - \Phi(-0.357)] = \Phi(0.357) = 0.279. \end{aligned}$$

Задания для самостоятельного решения

Ниже приводятся результаты измерений некоторой физической величины, которые будут рассматриваться как n реализаций случайной величины X . Предполагая, что СВ X имеет нормальное распределение, требуется:

1. Найти точечные несмещенные оценки математического ожидания a и среднего квадратического отклонения σ .
2. Записать плотность вероятности и функцию распределения СВ X .
3. Найти доверительный интервал, накрывающий математическое ожидание СВ X с заданной доверительной вероятностью $P = 1-\alpha = 0,95$, считая σ неизвестным.

4. Найти доверительный интервал, накрывающий среднее квадратическое отклонение σ с заданной вероятностью $P = 1 - \alpha = 0,95$.
5. Принимая $P = 1 - \alpha = 0,99$, найти предельную погрешность, с которой среднее арифметическое оценивает неизвестное математическое ожидание СВ X .
6. Найти минимальное число измерений, которое нужно произвести, чтобы с доверительной вероятностью $P = 1 - \alpha = 0,95$ можно было бы утверждать, что, принимая $M(X) = \bar{x}$, мы совершаем погрешность, не превышающую $\varepsilon = 0,2S$.
7. Вычислить:

$$P(a < X < b) = \frac{1}{2} \left[\Phi \left(\frac{b - \bar{x}}{S} \right) - \Phi \left(\frac{a - \bar{x}}{S} \right) \right].$$

Задачи для самостоятельного решения

Задача 1.

СВ X – сопротивление резистора в кило омах.

i – номер резистора	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i – сопротивление резистора (ком)	4,8	6,2	6,0	5,9	5,6	4,9	6,0	6,1	5,5	5,8	5,7	5,1	5,5	6,2	5,4

$P(X < 5) = ?$

Задача 2.

СВ X – еженедельные затраты времени (в часах) на посещение библиотеки, определяемые путем анкетирования:

i – номер анкеты	1	2	3	4	5	6	7	8	9	10	11	12
x_i – затраты времени (ч)	2,2	4,5	3,8	5,0	3,0	6,0	12,0	8,0	16,2	15,0	2,0	1,0

$P(8 < X < 14) = ?$

Задача 3.

СВ X – индуктивность катушки в мГн.

i – номер катушки	1	2	3	4	5	6	7	8	9	10
x_i – индуктивность (мГн)	8,345	8,346	8,348	8,342	8,343	8,345	8,343	8,347	8,344	8,347

$P(8,345 < X < 8,349) = ?$

IX. Критерий согласия χ^2

Предположим, что по виду гистограммы или полигона частостей или из каких - либо других соображений удастся выдвинуть гипотезу о множестве функций определенного вида (нормальных, показательных, биномиальных и т. п.), к которому может принадлежать функция распределения исследуемой СВ X . Критерий χ^2 Пирсона позволяет производить проверку согласия эмпирической функции распределения $F^*(x)$ с гипотетической функцией распределения $F(x)$.

Для этого придерживаются следующей последовательности действий:

1) на основании гипотетической функции $F(x)$ вычисляют вероятность попадания СВ X в частичные интервалы $[x_{i-1}, x_i[$:

$$p_i = P(x_{i-1} \leq X < x_x) = F(x_i) - F(x_{i-1}); i = 1, 2, \dots, k;$$

2) умножая полученные вероятности p_i на объем выборки n , получают теоретические частоты np_i частичных интервалов $[x_{i-1}, x_i[$, т.е. частоты, которые следует ожидать, если гипотеза справедлива;

3) вычисляют выборочную статистику (критерий) χ^2 :

$$\chi^2_{\text{набл.}} = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}.$$

Можно показать, что если гипотеза верна, то при $n \rightarrow \infty$ распределение выборочной статистики, независимо от вида функции $F(x)$, стремится к распределению χ^2 с $\nu = k - r - 1$ степенями свободы (k – число частичных интервалов, r – число параметров гипотетической функции $F(x)$, оцениваемых по данным выборки).

Критерий χ^2 сконструирован таким образом, что чем ближе к нулю наблюдаемое значение критерия χ^2 , тем вероятнее, что гипотеза справедлива. Поэтому для проведения гипотезы применяется критерий χ^2 с правосторонней критической областью. Необходимо найти по таблицам квантилей χ^2 – распределения по заданному уровню значимости α и числу степеней свободы $\nu = k - r - 1$ критическое значение $c_{\alpha, \nu}^2$, удовлетворяющее условию $P(\chi^2 \geq \chi_{\alpha, \nu}^2) = \alpha$.

Если $\chi^2_{\text{набл.}} \geq \chi_{\alpha, \nu}^2$, то считается, что гипотетическая функция $F(x)$ не согласуется с результатами эксперимента. Если $\chi^2_{\text{набл.}} < \chi_{\alpha, \nu}^2$, то считается, что гипотетическая функция $F(x)$ согласуется с результатами эксперимента.

Замечание. При применении критерия χ^2 необходимо, чтобы в каждом частичном интервале было не менее 5 элементов. Если число элементов (частота) меньше 5, то рекомендуется объединять такие частичные интервалы с соседними.

Х. Курсовая работа

Каждому студенту в соответствии со своим номером варианта требуется:

- 1) записать исходную выборку в виде таблицы;
- 2) построить статистический ряд;
- 3) записать сгруппированную выборку в виде таблицы;
- 4) построить график эмпирической функции распределения;
- 5) построить гистограмму;
- 6) проверить гипотезу о нормальном законе распределения случайной величины X и записать вычисления в таблицу;
- 7) построить график плотности случайной величины X .

При выполнении работы принять уровень значимости $\alpha = 0,05$, отрезок $[24,5; 54,5]$, число интервалов $k = 10$. Варианты индивидуальных заданий приведены в таблице.

i – му варианту соответствуют элементы выборки, расположенные в 15 – ти следующих строчках таблицы, начиная с i – й (объем выборки при этом $n = 150$).

1	48	39	43	44	34	34	32	43	40	46
2	25	31	34	49	39	37	45	49	31	49
3	43	46	34	35	42	32	41	34	42	42
4	38	40	46	47	34	42	38	40	38	36
5	30	43	41	40	40	35	35	41	38	45
6	37	42	38	36	44	39	32	48	43	39
7	43	30	32	36	42	34	49	48	49	50
8	37	30	44	48	44	35	45	34	33	41
9	43	45	50	34	33	39	41	39	46	31
10	40	52	44	39	35	45	33	42	42	36
11	44	51	45	39	34	44	40	37	43	32
12	33	42	40	35	37	43	48	48	50	32
13	40	48	45	43	36	36	42	40	37	30
14	44	50	46	39	41	48	44	42	36	51
15	44	50	47	37	33	34	42	43	43	47
16	33	48	38	42	45	32	34	44	39	45
17	48	26	31	34	38	36	46	49	40	48
18	42	47	35	34	41	33	41	35	43	42
19	39	37	47	47	33	42	37	39	39	37
20	43	41	30	39	38	36	36	34	42	46
21	39	44	37	35	43	38	33	47	45	38
22	37	48	38	52	40	45	44	42	38	40
23	44	46	37	34	41	37	41	39	30	38
24	32	41	48	36	51	36	33	39	45	40

25	34	41	38	34	33	27	51	45	27	38
26	42	37	46	41	47	36	30	45	41	40
27	37	37	39	42	48	41	36	39	33	47
28	43	49	27	31	41	46	40	36	36	42
29	41	46	33	37	47	35	31	29	30	36
30	48	38	37	34	40	34	36	50	48	39
31	30	38	43	41	44	45	38	37	46	50
32	41	48	41	43	47	37	42	34	32	44
33	37	48	46	41	41	37	37	48	49	46
34	38	44	50	37	47	27	48	37	46	38
35	48	47	38	52	34	36	34	41	41	32
36	31	43	34	46	37	40	41	39	32	42
37	47	33	51	41	40	45	37	36	27	36
38	37	42	46	35	34	38	45	36	20	40
39	34	48	30	51	33	41	44	42	39	39
40	45	45	41	40	36	27	50	44	41	48
41	36	36	32	32	36	49	27	45	30	35
42	40	38	45	40	40	50	42	37	50	39
43	43	38	30	59	42	41	33	42	38	44
44	44	41	47	52	51	38	50	39	50	48
45	49	43	52	50	30	30	26	50	27	49
46	27	49	46	39	47	26	49	52	29	44
47	51	53	48	49	53	45	27	43	48	44

Порядок выполнения работы

1. По данной выборке объема n строится статистический ряд

y_1	y_2	...	y_e
n_1	n_2	...	n_e

где $y_1 < y_2 < \dots < y_e$ элементы выборки, записанные в порядке возрастания, n_i – частоты появления одинаковых значений СВ X .

2. На основе статистического ряда строится сгруппированная выборка.

Для этого задается определенный отрезок $[a, b]$, внутри которого расположены все элементы исследуемой выборки, число интервалов k , на которое делится этот отрезок. Находятся длины интервалов

$h = \frac{b-a}{k}$, концы интервалов $x_i = a + (i-1)h$, середины интервалов

$z_i = \frac{1}{2}(x_i + x_{i+1})$ и соответствующие эмпирические частоты m_i (m_i –

число элементов выборки, попавших в i – й интервал), $i = 1, 2, \dots k$.

Результаты вычислений заносятся в таблицу:

номер интервала	границы интервала	середины интервалов	эмпирические частоты
i	x_i, x_{i+1}	z_i	m_i
1			
2			
⋮			
⋮			
k			

3. Строится график эмпирической функции распределения

$$F^*(x) = \begin{cases} 0 & x \leq z_1 \\ \frac{1}{n}(m_1 + \dots + m_i), & \text{при } z_i < x \leq z_{i+1}, (i=1,2,\dots,k-1). \\ 1 & x > z_k \end{cases}$$

4. Строится гистограмма – фигура, состоящая из прямоугольников с основаниями $[x_i, x_{i+1}]$ и высотами $\frac{m_i}{nh}$.

5. Находится выборочное среднее $\bar{x} = \frac{1}{n} \sum_{i=1}^k m_i z_i$, исправленная

выборочная дисперсия $S^2 = \frac{1}{n-1} \sum_{i=1}^k m_i (z_i - \bar{x})^2$; исправленное

выборочное среднее квадратическое отклонение $S = \sqrt{S^2}$.

6. Проверяется гипотеза о нормальном распределении СВ X с математическим ожиданием $a = \bar{x}$ и средним квадратическим отклонением $s = S$ с помощью критерия χ^2 Пирсона.

Для этого вычисляют теоретические частоты попадания СВ X в i – й интервал np_i ,

$$\text{где } p_i = P\{x_i \leq X < x_{i+1}\} = \Phi\left(\frac{x_{i+1} - \bar{x}}{S}\right) - \Phi\left(\frac{x_i - \bar{x}}{S}\right).$$

Значения функции Лапласа $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$ находятся по таблице.

Если при некотором i эмпирическая или теоретическая частота меньше 5, тогда этот интервал объединяют с соседним, при этом теоретические и эмпирические частоты суммируются. После объединения получают r интервалов ($r \leq k$).

Составляется статистика χ^2 Пирсона

$$\chi^2_{\text{набл.}} = \sum_{i=1}^r \frac{(m_i - np_i)^2}{np_i}.$$

Затем по закону уровня значимости α и числу степеней свободы $\nu = r-3$ находится критическая точка $c_{\alpha, \nu}^2$ по таблице квантилей распределения χ^2 . Если $\chi^2_{\text{набл.}} > \chi^2_{\alpha, \nu}$, то гипотеза отвергается. Если $\chi^2_{\text{набл.}} \leq \chi^2_{\alpha, \nu}$, гипотеза принимается.

7. Строится график плотности вероятности

$$f(x) = \frac{1}{\sqrt{2\pi}S} e^{-\frac{(x-\bar{x})^2}{2S^2}}$$

случайной величины X , распределенной по нормальному закону.

Пример выполнения курсовой работы

1	37	30	44	48	44	35	45	34	33	41
2	43	45	50	34	33	39	41	39	46	31
3	40	52	44	39	35	45	33	42	42	36
4	44	51	45	39	34	44	40	37	43	32
5	33	42	40	35	37	43	48	48	50	32
6	40	48	45	43	36	36	42	40	37	30
7	44	50	46	39	41	48	44	42	36	51
8	44	50	47	37	33	34	42	43	43	47
9	33	48	38	42	45	32	34	44	39	45
10	48	26	31	34	38	36	46	49	40	48
11	42	47	35	34	41	33	41	35	43	42
12	39	37	47	47	33	42	37	39	39	37
13	43	41	30	39	38	36	36	34	42	46
14	39	44	37	35	43	38	33	47	45	38
15	37	48	38	52	40	45	44	42	38	40

Для данной выборки объема $n = 150$ построим статистический ряд, где $Y_1 < Y_2 < \dots < Y_m$ – элементы выборки, записанные в порядке возрастания, n_i – число повторений элемента Y_i в выборке.

26	30	31	32	33	34	35	36	37	38	39	40
1	3	2	3	9	8	6	7	10	7	11	8

41	42	43	44	45	46	47	48	49	50	51	52
6	12	9	11	9	4	6	9	1	4	2	2

№	Граница интервала $x_i; x_{i+1}$	Середины интервалов Z_i	Эмпирические частоты m_i
1	24.5 27.5	26	1
2	27.5 30.5	29	3
3	30.5 33.5	32	14
4	33.5 36.5	35	21
5	36.5 39.5	38	28
6	39.5 42.5	41	26
7	42.5 45.5	44	29
8	45.5 48.5	47	19
9	48.5 51.5	50	7
10	51.5 54.5	53	2

$$X = 40.34$$

$$X_H = 2.465$$

$$k = 4$$

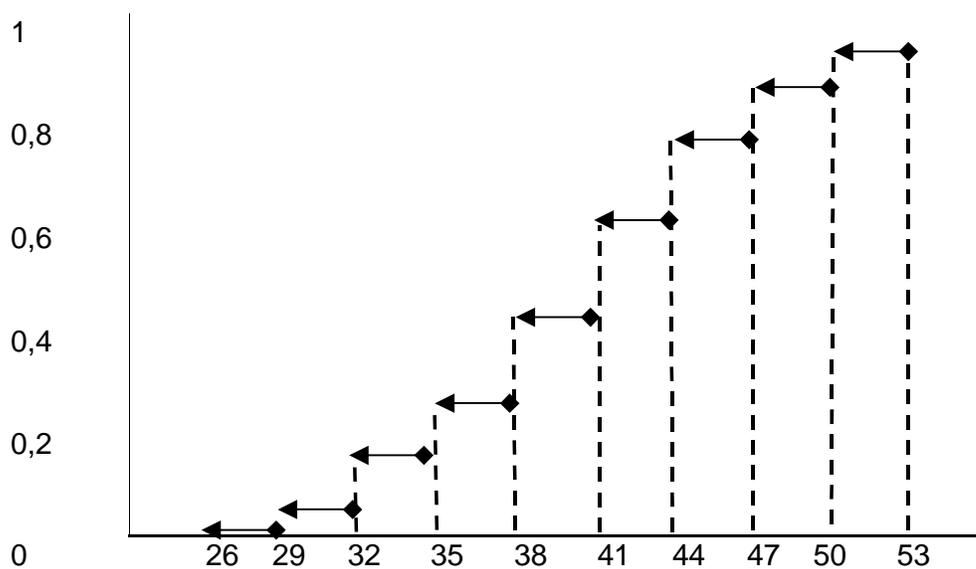
$$S = 5.51$$

$$X_H(\alpha, k) = 9.5$$

$$r = 7$$

Так как $X_H(\alpha, k) \geq X_H$, то гипотеза о нормальном распределении принимается, результаты занесены в таблицу.

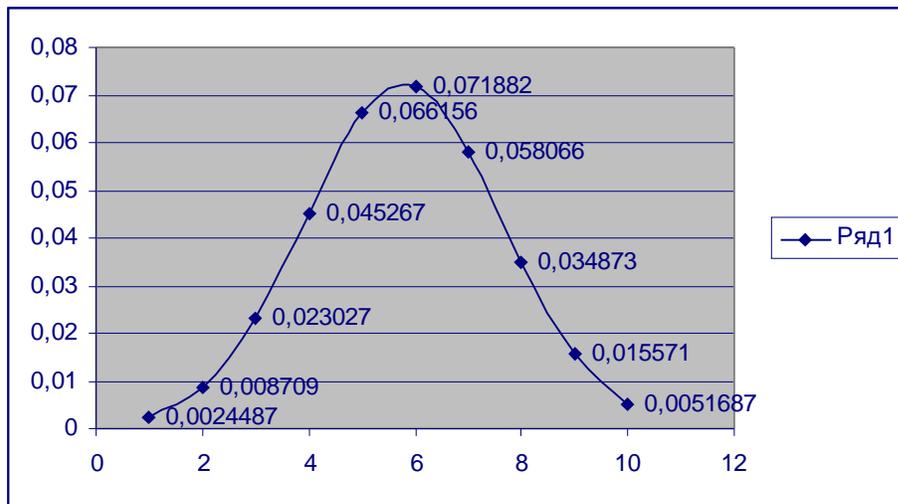
I	$x_i; x_{i+1}$	m_i	p_i	np_i	$(m_i - np_i)^2$	$\frac{(m_i - np_i)^2}{np_i}$
1	24.5 27.5	1	0.0078	1.17		
2	27.5 30.5	3	0.0208	3.12		
3	30.5 33.5	14	0.0768	11.52	4.49	0.302
4	33.5 36.5	21	0.1376	20.64	0.12	0.005
5	36.5 39.5	28	0.1998	29.97	3.88	0.129
6	39.5 42.5	26	0.2113	31.69	32.43	1.043
7	42.5 45.5	29	0.1721	25.81	10.14	0.392
8	45.5 48.5	19	0.1068	16.02	8.88	0.554
9	48.5 51.5	7	0.0477	7.15		
10	51.5 54.5	2	0.0165	2.47	0.39	0.04



Эмпирическая функция распределения

$$F^*(x) = \begin{cases} 0 & x \leq z_1 \\ \frac{1}{n}(m_1 + \dots + m_i) & \text{при } z_i < x \leq z_{i+1}, (i=1,2,\dots,k-1). \\ 1 & x > z_k \end{cases}$$

График плотности вероятности СВ X.



Приложение 2

1. Нормальное распределение

Плотность вероятностей нормированного нормального распределения:

$$u \rightarrow N(0,1) \quad f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

<i>U</i>	0	1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9	10	11
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0203	0203	0198	0194	0189	0184	0180
2,5	0175	0181	0167	0158	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0122	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0071	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0053	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

2. Нормальное распределение

Значение функции:

$$\Phi(u_i) = \frac{2}{\sqrt{2p}} \int_0^{u_i} e^{-\frac{x^2}{2}} dx = P(|u_i| < u_i).$$

Целые и десятичные доли u_i	Сотые доли u_i									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,3999	0,0478	0,0558	0,0638	0,0717
0,1	0797	0876	0955	1034	1113	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2960	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7994	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9090
1,7	9109	9127	9146	9164	9181	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9426	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9756	9762	9768	9774	9780
2,3	9786	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9841	9845	9849	9853	9857	9861	9865	9869	9872
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9910	9912	9915	9917	9920	9922	9924	9926	9928
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9951	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9979	0,9979	0,9979	0,9980
3,1	9981	9981	9984	9983	9983	9984	9984	9985	9985	9986
3,2	9986	9987	9987	9988	9988	9989	9989	9989	9990	9990
3,3	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993
3,4	9993	9994	9994	9994	9994	9994	9995	9995	9995	9995
3,5	9995	9996	9996	9996	9996	9996	9996	9996	9997	9997
3,6	9997	9997	9997	9997	9997	9997	9998	9998	9998	9998
3,7	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3,8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3,9	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999

Целые и десятичные доли u_i	Сотые доли u_i									
4,0	0,999936	9999	9999	9999	9999	9999	9999	9999	9999	9999
4,5	0,999994	-	-	-	-	-	-	-	-	-
5,0	0,99999994	-	-	-	-	-	-	-	-	-

Распределения Стьюдента

Значения $t_{a,n}$ удовлетворяют условию

$$P(t \geq t_{a,n}) = \int_{t_{a,n}}^{\infty} S(t,n) dt = a.$$

В таблице приведены значения квантилей $t_{a,n}$ в зависимости от числа степеней свободы n и вероятности a .

$\alpha \backslash v$	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
1	0,325	0,727	1,376	3,078	6,314	12,71	31,82	63,66	318,3	636,6
2	0,289	0,617	1,061	1,886	2,920	4,303	6,965	9,925	22,33	31,60
3	0,277	0,584	0,978	1,638	2,353	3,182	4,541	5,841	10,22	12,94
4	0,271	0,569	0,941	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	0,267	0,559	0,920	1,476	2,015	2,571	3,365	5,032	5,893	6,859
6	0,265	0,553	0,906	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	0,263	0,549	0,896	1,415	1,895	2,365	2,998	3,499	4,785	5,405
8	0,262	0,546	0,889	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	0,261	0,543	0,883	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	0,260	0,542	0,879	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	0,260	0,540	0,876	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	0,259	0,539	0,873	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	0,259	0,538	0,870	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	0,258	0,537	0,868	1,345	1,761	2,145	2,624	3,977	3,787	4,140
15	0,258	0,536	0,866	1,341	1,753	2,131	2,602	2,947	3,733	4,073
16	0,258	0,535	0,865	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	0,257	0,534	0,863	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	0,257	0,534	0,862	1,330	1,734	2,101	2,552	2,878	3,611	3,922
19	0,257	0,533	0,861	1,328	1,729	2,093	2,539	2,861	3,597	3,883
20	0,257	0,533	0,860	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	0,257	0,533	0,859	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	0,256	0,532	0,858	1,1321	1,717	2,074	2,508	2,819	3,505	3,792
23	0,256	0,532	0,858	1,319	1,714	2,069	2,500	2,807	3,485	3,767

$\alpha \backslash v$	0,40	0,30	0,20	0,10	0,050	0,025	0,010	0,005	0,001	0,0005
24	0,256	0,531	0,857	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	0,256	0,531	0,856	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	0,256	0,531	0,856	1,315	1,706	2,056	2,479	2,779	3,435	3,707
27	0,256	0,531	0,855	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	0,256	0,530	0,855	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	0,256	0,530	0,854	1,311	1,699	2,045	2,462	2,756	3,396	3,659
30	0,256	0,530	0,854	1,310	1,697	2,042	2,457	2,750	3,385	3,646
40	0,256	0,529	0,851	1,303	1,684	2,021	2,423	2,704	3,307	3,551
50	0,255	0,528	0,849	1,298	1,676	2,009	2,403	2,678	3,262	3,495
60	0,255	0,527	0,848	1,296	1,671	2,000	2,390	2,660	3,232	3,460
80	0,254	0,527	0,846	1,292	1,664	1,990	2,374	2,639	3,195	3,415
100	0,254	0,526	0,846	1,290	1,660	1,984	2,365	2,622	3,174	3,389
200	0,254	0,525	0,843	1,286	1,653	1,972	2,345	2,601	3,131	3,339
500	0,253	0,525	0,842	1,283	1,648	1,965	2,334	2,586	3,106	3,310
∞	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	3,090	3,291

4.с²- распределение

В таблице представлены значения квантилей $C^2_{a,n}$ в зависимости от числа степеней свободы n и вероятности a .

$\alpha \backslash v$	0,20	0,10	0,05	0,02	0,01	0,001
1	1,642	2,706	3,841	5,412	6,635	10,827
2	3,219	4,605	5,991	7,824	9,210	13815
3	4,642	6,251	7,815	9,837	11,345	16,266
4	5,989	7,779	9,488	11,668	13,277	18,467
5	7,289	9,236	11,070	13,388	15,086	20,515
6	8,558	10,645	12,592	15,033	16,812	22,457
7	9,803	12,017	14,067	16,622	18,475	24,322
8	11,030	13,362	15,507	18,168	20,090	26,125
9	12,242	14,684	16,919	19,679	21,666	27,877
10	13,442	15,987	18,307	21,161	23,209	29,588
11	14,631	17,275	19,675	22,618	24,725	31,264
12	15,812	18,549	21,026	24,054	26,217	32,909
13	16,985	19,812	22,362	25,472	27,688	34,528
14	18,151	21,064	23,685	26,783	29,141	36,123
15	19,311	22,307	24,996	28,259	30,578	37,697
16	20,465	23,542	26,296	29,633	32,000	39,252
17	21,615	24,769	27,587	30,995	33,409	40,790
18	22,760	25,989	28,869	32,346	34,805	42,312
19	23,900	27,204	30,144	33,687	36,191	43,820

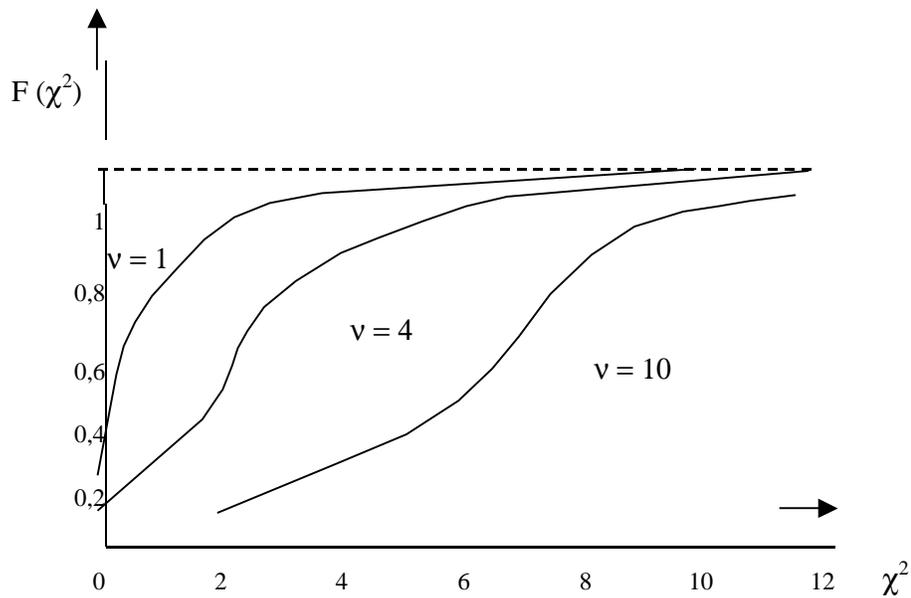
α v	0,20	0,10	0,05	0,02	0,01	0,001
20	25,038	28,412	31,410	35,020	37,556	45,315
21	26,171	29,615	32,671	36,343	38,932	46,797
22	27,301	30,813	33,924	37,659	40,289	48,268
23	28,429	32,007	35,172	38,968	41,638	49,728
24	29,559	33,196	36,415	40,270	42,980	51,179
25	30,675	34,382	37,652	41,556	44,314	52,620
26	31,795	35,563	38,885	42,856	45,642	54,052
27	32,912	36,741	40,113	44,140	46,963	55,476
28	34,027	37,916	41,337	45,419	48,278	56,893
29	35,139	39,087	42,557	46,693	49,588	58,302
30	36,250	40,256	43,773	47,962	50,892	59,703

5. Доверительные интервалы для S
Нижние g_1 и верхние g_2 границы доверительного интервала.

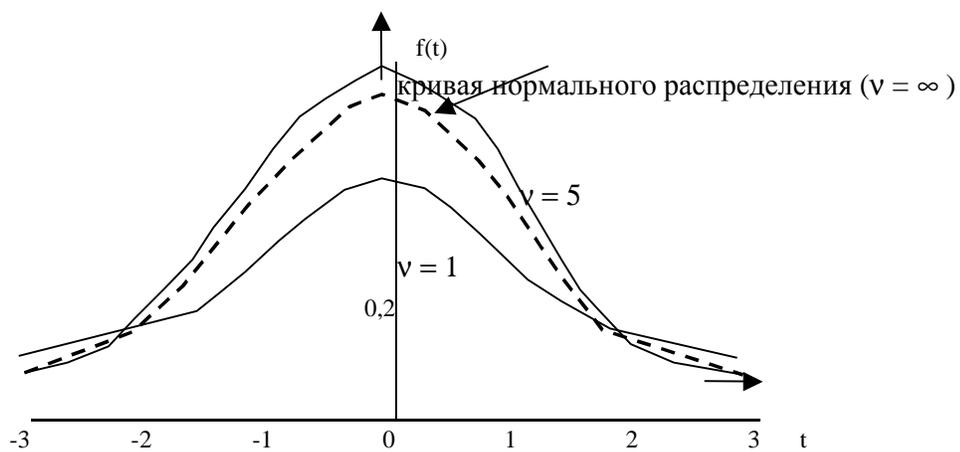
$$g_1 s < S < g_2 s \quad (s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2})$$

P v=n-1	0,99		0,98		0,95		0,90	
	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2
1	0,356	159	0,388	79,8	0,446	31,9	0,510	15,9
2	0,434	14,1	0,466	9,97	0,521	6,28	0,578	4,40
3	0,483	6,47	0,514	5,11	0,566	3,73	0,620	2,92
4	0,519	4,39	0,549	3,67	0,599	2,87	0,649	2,37
5	0,546	3,48	0,576	3,00	0,624	2,45	0,672	2,09
6	0,569	2,98	0,597	2,62	0,644	2,202	0,690	1,916
7	0,588	2,66	0,616	2,377	0,661	2,035	0,705	1,797
8	0,604	2,44	0,631	2,205	0,675	1,916	0,718	1,711
9	0,618	2,277	0,644	2,076	0,688	1,826	0,729	1,645
10	0,630	2,154	0,656	1,977	0,699	1,755	0,739	1,593
11	0,641	2,056	0,667	1,898	0,708	1,698	0,748	1,550
12	0,651	1,976	0,677	1,833	0,717	1,651	0,755	1,515
13	0,660	1,910	0,685	1,779	0,725	1,611	0,762	1,485
14	0,669	1,854	0,693	1,733	0,732	1,577	0,769	1,460
15	0,676	1,806	0,700	1,694	0,739	1,548	0,775	1,437
16	0,683	1,764	0,707	1,659	0,745	1,522	0,780	1,418
17	0,690	1,727	0,713	1,629	0,750	1,499	0,785	1,400
18	0,696	1,695	0,719	1,602	0,756	1,479	0,790	1,385
19	0,702	1,666	0,725	1,578	0,760	1,460	0,794	1,370
20	0,707	1,640	0,730	1,556	0,765	1,444	0,798	1,358
21	0,712	1,617	0,734	1,536	0,769	1,429	0,802	1,346
22	0,717	1,595	0,739	1,519	0,773	1,416	0,805	1,335
23	0,722	1,576	0,743	1,502	0,777	1,402	0,809	1,326
24	0,726	1,558	0,747	1,487	0,781	1,391	0,812	1,316
25	0,730	1,541	0,751	1,473	0,784	1,380	0,815	1,308
26	0,734	1,526	0,755	1,460	0,788	1,371	0,818	1,300
27	0,737	1,512	0,758	1,448	0,791	1,361	0,820	1,293

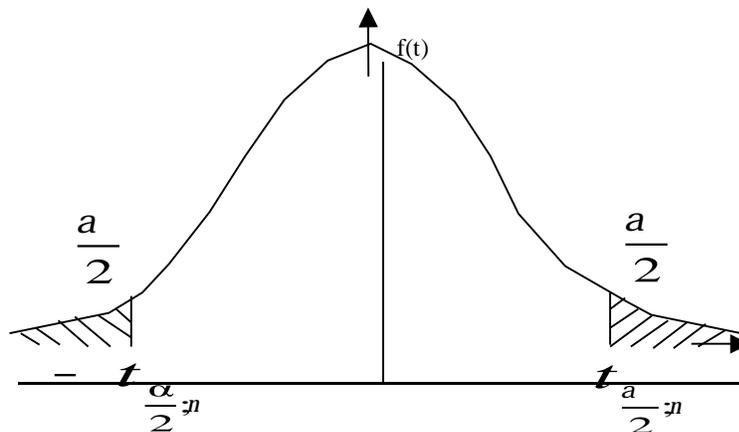
P v=n-1	0.99		0.98		0.95		0.90	
	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2	γ_1	γ_2
28	0,741	1,499	0,762	1,436	0,794	1,352	0,823	1,286
29	0,744	1,487	0,765	1,426	0,796	1,344	0,825	1,279
30	0,748	1,475	0,768	1,417	0,799	1,337	0,828	1,274
40	0,774	1,39	0,792	1,344	0,821	1,279	0,847	1,228
50	0,793	1,336	0,810	1,297	0,837	1,243	0,861	1,199
60	0,808	1,299	0,824	1,265	0,849	1,217	0871	1,179
70	0,820	1,272	0,835	1,241	0,858	1,198	0,879	1,163
80	0,829	1,250	0,844	1,222	0,866	1,183	0,886	1,151
90	0,838	1,233	0,852	1,207	0,873	1,171	0,892	1,141
100	0,845	1,219	0,858	1,195	0,878	1,161	0,897	1,133
200	0,887	1,15	0,897	1,13	0,912	1,11	0,925	1,09



(рис. 7)



(рис. 8)



(рис. 9)

Литература

1. Чистяков В. П. Курс теории вероятности. – М.: Наука, 1987.
2. Крамер Г. Математические методы статистики. – М.: Мир, 1975.
3. Гмурман В. Е. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1998.
4. Коваленко И. Н., Филиппова А. А. Теория вероятностей и математическая статистика. – М.: Высшая школа, 1982.
5. Герасимович А. И. Математическая статистика. – Минск: Высшая школа, 1983.

Оглавление

1. Задача математической статистики.....	3
2. Генеральная и выборочная совокупность.....	3
3. Статистический ряд. Статистический закон распределения случайной величины.....	4
4. Эмпирическая функция распределения.....	5
5. Основные законы распределения случайных величин, используемых в математической статистике.....	7
6. Точечные оценки параметров нормального распределения.....	12
7. Интервальные оценки параметров нормального распределения.....	13
8. Примеры обработки результатов эксперимента.....	16
9. Критерий согласия χ^2	20
10. Курсовая работа.....	21
11. Приложение 1.....	29
12. Приложение 2.....	31

