МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)

СТАТИСТИЧЕСКИЙ АНАЛИЗ ТЕХПРОЦЕССОВ И КОНСТРУКЦИИ РЭС ПРОГРАММНЫМИ СРЕДСТВАМИ

Рекомендовано редакционно-издательским советом федерального государственного автономного образовательного учреждения высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева» в качестве методических указаний для студентов Самарского университета, обучающихся по основным образовательным программам высшего образования по направлениям подготовки 11.03.03 Конструирование и технология электронных систем и устройств, 12.03.04 Биотехнические системы и технологии

Составитель: *В.А. Зеленский,*
К.И. Сухачев

С А М А Р А
Издательство Самарского университета
2020

УДК 621.38(075)+004.9(075)
ББК 31.2я7+32.81я7

Составители: ***В.А. Зеленский, К.И. Сухачев***

Рецензент канд. техн. наук, доц. К. Е. Воронов

Статистический анализ техпроцессов и конструкции РЭС программными средствами: методические указания / *В.А. Зеленский, К.И. Сухачев;* Министерство науки и высшего образования Российской Федерации, Самарский университет. − Самара: Издательство Самарского университета, 2020. – 1 CD-ROM (1,5 Мб). – Загл. с титул. экрана. – Текст: электронный.

Рассмотрены вопросы решения задач статистического анализа конструкции и техпроцессов РЭС. Приводится методика обработки экспериментальных данных, оценка параметров распределения. Даны варианты исходных данных для самостоятельной работы студентов с примерами её выполнения.

Методические указания могут применяться при проведении лабораторных работ и практических занятий по дисциплинам «Теоретические основы конструирования, технологии и надежности», «Методы и средства обеспечения надежности РЭС», «Программирование на алгоритмических языках».

Предназначены для бакалавров, обучающихся по направлениям подготовки 11.03.03 Конструирование и технология электронных средств, 12.03.04 Биотехнические системы и технологии.

Методические указания разработаны на кафедре «Конструирование и технология электронных систем и устройств».

Минимальные системные требования:
PC, процессор Pentium, 160 МГц; оперативная память 32 Мб;
на винчестере 16 Мб; Microsoft Windows XP/Vista/7;
разрешение экрана 1024x768 с глубиной цвета 16 бит;
DVD-ROM 2-х и выше, мышь; Adobe Acrobat Reader.

УДК 621.38(075)+004.9(075)
ББК 31.2я7+32.81я7

Редактор Т.К. Кретинина

Компьютерная вёрстка А.В. Ярославцевой

Подписано для тиражирования 13.08.2020.

Объем издания 1,5 Мб.

Количество носителей 1 диск.

Тираж 10 дисков.

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С. П. КОРОЛЕВА»
(САМАРСКИЙ УНИВЕРСИТЕТ)
443086, САМАРА, МОСКОВСКОЕ ШОССЕ, 34.

Издательство Самарского университета.
443086, Самара, Московское шоссе, 34.

СОДЕРЖАНИЕ

СОДЕРЖАНИЕ ... 3
1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ ... 5
2. СТАТИСТИЧЕСКАЯ ОЦЕНКА ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ ... 6
2.1. Предварительная обработка статистических данных ... 6
2.2. Точечные оценки параметров распределения ... 8
2.3. Интервальные оценки. Нахождение доверительных интервалов для дисперсии и математического ожидания ... 11
2.4. Практическая часть ... 13
3. СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ ПРОГРАММНЫМИ СРЕДСТВАМИ ... 15
4. КОНТРОЛЬНЫЕ ВОПРОСЫ ... 24
5 ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ ... 25
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ ... 27
ПРИЛОЖЕНИЕ ... 28

1. ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Значения параметров технологического процесса отклоняются от нормы со временем и под воздействием случайных факторов, действующих на РЭС. Поэтому значения параметров выпускаемых изделий являются случайными. При изучении степени отклонения значений какого-либо параметра статистически однородных изделий практически невозможно и экономически нецелесообразно исследовать каждый объект изучаемой совокупности. Поэтому из всей совокупности однородных объектов, называемой **генеральной**, отбирают случайно определенное количество, называемое **выборкой**. Число объектов в выборке называют объемом выборки. Пусть для получения некоторого количественного признака из генеральной совокупности извлечена выборка объема **n**. Значения $x_1, x_2, \ldots, x_n$ признака X в выборке называют **вариантами**, а последовательность вариант, записанных в возрастающем порядке **- вариационным рядом**.

2. СТАТИСТИЧЕСКАЯ ОЦЕНКА ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

2.1. Предварительная обработка статистических данных

Для ускорения расчетов и предупреждения ошибки необходима предварительная обработка данных, полученных в результате измерений. Если варианты выборки представлены дробными числами, то целесообразно умножить их на какую-то постоянную величину, чтобы оперировать далее только с целыми числами. Если варианты являются числами и различаются лишь в нескольких последних знаках, то следует отбросить постоянную часть вариант. После завершения расчетов необходимо провести с результатом обратные операции. Если объем выборки невелик, следует расположить варианты в виде вариационного ряда и пронумеровать их. Выборочные данные иногда могут содержать резко отклоняющиеся результаты, так называемые выскакивающие варианты. Они являются, как правило, следствием грубой ошибки в проведении эксперимента или измерения, оставшейся незамеченной. Здесь рассматривается очень быстрый способ выявления выскакивающих вариант, основанный на оценке различий крайних вариант вариационного ряда, который позволяет с достаточной строгостью решить эту задачу. Пусть имеем выборку объема n, данные которой представлены в виде вариационного ряда $x_1, x_2, \ldots x_i, \ldots x_n$. Для проверки вариант, относительно которых можно предположить, что они являются выскакивающими, следует вычислить отношения:

Для проверки наибольшей варианты: $\frac{x_n - x_{n-1}}{x_n - x_2}$. (1)

Для проверки наименьшей варианты: $\frac{x_2 - x_1}{x_n - x_2}$. (2)

Затем нужно сравнить вычисленные значения с табличными данными (табл. 1) для объема выборки n и уровней значимости α=0,05 или 0,01.

Таблица 1 – Проверка на выскакивающие варианты

Объем выборки n	Уровень значимости α	
	0,05	0,01
3	1	1
4	0,955	0,991
5	0,807	0,916
6	0,689	0,805
7	0,61	0,74
8	0,554	0,683
9	0,512	0,635
10	0,477	0,597
11	0,45	0,566
12	0,428	0,541
15	0,381	0,486
20	0,334	0,43
24	0,309	0,4
30	0,283	0,369
35	0,257	0,338

В общем случае под уровнем значимости в математической статистике понимают вероятность принятия ошибочного решения. Здесь это вероятность того, что мы ошибочно исключим проверяемую к 1 варианту, хотя в действительности она не является грубой ошибкой эксперимента, т. е. фактически эта варианта характерна для изучаемой генеральной совокупности. Если хотя бы одно из трех вычислительных отношений превышает соответствующее табличное значение, это уже дает право на безоговорочное исключение крайней варианты. Если каждое из трех вычисленных значений меньше соответствующего табличного, то проверяемая крайняя варианта не может быть значимости 0,05 и 0,01. В таком случае нет оснований для безоговорочного вывода об исключении крайней варианты. Можно лишь отметить, что велика вероятность грубой ошибки при получении этой варианты. Выскакивающую варианту необходимо исключить из всех последующих операций по статистической обработке.

2.2. Точечные оценки параметров распределения

При изготовлении каких-либо деталей при конструировании РЭС необходимо знать, удовлетворяют ли их параметры требованиям технологической точности. Для этого и служит оценка, которая характеризует истинное значение параметра в некоторой точке (точечная оценка) либо в интервале (доверительная оценка). Оценкой Θ_n случайной величины X по объему выборки n называется однозначно определенная функция результатов наблюдений над этой случайной величиной, и можно записать: $\Theta_n = f(x_1, x_2, ..., x_n)$.

В практике обработки статистических данных оценивают математическое ожидание и дисперсию случайной величины X. Математическое ожидание - это центр группирования случайной величины и в общем случае определяется выражением

$$\mu = \int_{-\infty}^{+\infty} x\, f(x)\, dx\,, \tag{3}$$

где $f(x)$ - плотность распределения случайной величины.

Дисперсия - это отклонение случайной величины от ее математического ожидания, определяется выражением

$$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\, dx\ . \tag{4}$$

Чтобы приблизить с достаточной точностью значение случайной оценочной функции $f(x_1, x_2, ..., x_n)$ к истинному значению параметра, эта функция должна по возможности обладать следующими свойствами: состоятельностью, несмещенностью и эффективностью.

Оценка Θ_n называется состоятельной, если с увеличением n - объема выборки она приближается (сходится по вероятности) к оцениваемому параметру Θ. Несмещенной называется такая оценка, математическое ожидание от которой равно оцениваемому параметру: M[Θn] = Θ, т. е. она контролирует наличие систематической ошибки. Эффективной оценкой называется такая несмещенная оценка, которая имеет наименьшую дисперсию всех несмещенных оценок параметра, вычисленных по выборкам одного и того же объема.

Несмещенной и состоятельной оценкой математического ожидания μ случайной величины - X является арифметическое среднее $\bar{X}$, вычисленное по n независимым наблюдениям над этой случайной величиной:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \ , \qquad (5)$$

где x_i - результат i-го наблюдения.

Эффективность этой оценки зависит от вида закона распределения случайной величины X . Если случайная величина распределена по нормальному закону с параметрами μ, σ^2, плотность вероятности нормального закона распределения

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ , \qquad (6)$$

то среднее арифметическое $\bar{X}$ имеет минимальную дисперсию, равную $\frac{\sigma^2}{n}$, и является эффективной оценкой математического ожидания μ .

На рис. 1 представлены примеры графиков плотности нормального распределения случайной величины. Стандартным нормальным распределением называется нормальное распределение с математическим ожиданием μ = 0 и стандартным отклонением σ = 1

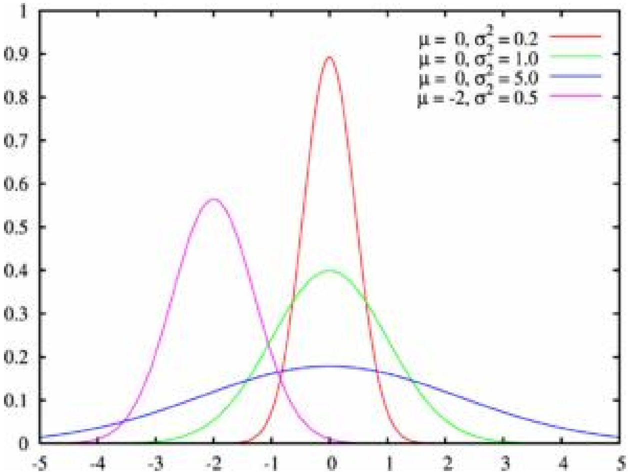


Рисунок 1 – Плотности вероятности нормального закона распределения

Состоятельная оценка дисперсии σ^2 определяется выражением

$$\Delta = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2 \ . \tag{7}$$

Состоятельной и несмещенной оценкой дисперсии является оценка

$$\Delta_* = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2 \ . \tag{8}$$

Сравнивая Δ и Δ_* получаем:

$$\Delta_* = \frac{n}{n-1}\Delta \ .$$

(9)

Величину $\frac{n}{n-1}$ называют поправкой Бесселя, а оценку Δ* исправленной выборочной дисперсией. Введение поправки Бесселя существенно лишь для малого объема выборки.

Оценка Δ* не является эффективной, но с увеличением при нормальном законе распределения отношение ее дисперсии к минимально возможной приближается к единице и ее можно считать "'асимптотически эффективной".

Состоятельной, несмещенной и эффективной оценкой дисперсии является:

$$\Delta_{**} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \ , \tag{10}$$

где μ - математическое ожидание случайной величины X. Из трех рассмотренных оценок дисперсии наиболее употребительна в практических расчетах оценка Δ* (8), так как μ обычно неизвестно, и в этих условиях она предпочтительнее оценки Δ (7).

Среднее квадратическое отклонение σ как характеристика меры рассеяния случайной величины X относительно математического ожидания не менее часто используется на практике, чем дисперсия σ^2. Удобство этой характеристики заключается в том, что ее размерность равна размерности самой случайной величины X.

Несмещенная и состоятельная оценка среднеквадратического отклонения с учетом (8) имеет вид

$$S=\sqrt{\Delta_*}=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{X})^2}\ . \qquad (11)$$

2.3. Интервальные оценки. Нахождение доверительных интервалов для дисперсии и математического ожидания

Вероятность попадания случайной величины в заданный интервал, можно найти как определенный интеграл от плотности вероятности. Если рассмотреть вариант стандартного нормального распределения с плотностью

$$f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\ , \qquad (12)$$

то вероятность попадания случайной величины X в интервал $[AB]$

$$P=\frac{1}{\sqrt{2\pi}}\int_A^B e^{-\frac{x^2}{2}}dx\ . \qquad (13)$$

Рассмотрим функцию Лапласа, которая определяет площадь под кривой плотности вероятности случайной величины, распределенной по стандартному нормальному закону, в интервале от 0 до Z:

$$\Phi(Z)=\frac{1}{\sqrt{2\pi}}\int_0^Z e^{-\frac{z^2}{2}}dz\ . \qquad (14)$$

Тогда вероятность попадания нормально распределенной случайной величины X в заданный интервал $[AB]$ можно найти с помощью таблиц значений функции Лапласа:

$$P(A<X<B)=\Phi(\frac{B-\mu}{\sigma})-\Phi(\frac{A-\mu}{\sigma})\ . \qquad (15)$$

На выражениях (12)-(15) базируется нахождение доверительных интервалов и интервальная оценка. Она особенно необходима при малом объеме выборки, когда точечная оценка в значительной мере случайна. Интервальной называют оценку, которая определяется двумя числами - концами интервала, покрывающими оцениваемый параметр. В качестве границ доверительного интервала принимаются такие случайные величины, при которых вероятность того, что истинное значение параметра Θ окажется внутри этого интервала, близка к единице. Эту вероятность называют доверительной и ее берут равной

0,9…0,99. Величину $\alpha=1\text{-}P$ называют уровнем значимости или вероятностью ошибки.

Доверительный интервал – это интервал, в границы которого [Θн; Θв] попадает неизвестное истинное значение параметра с вероятностью ошибки $\alpha=1\text{-}P$.

Функции доверительного интервала случайны, т. е. находятся по выборочным данным. Чем больше объем выборки, тем уже доверительный интервал для той же доверительной вероятности P. Чем больше выбранная вероятность, тем шире для той же самой выборки доверительный интервал.

Таким образом, зная выборочное распределение оценки, можно определить границы доверительного интервала Θн и Θв для выбранной доверительной вероятности P.

В математической статистике наиболее полно получены выборочные распределения оценок математического ожидания и дисперсии для нормально распределенной величины X. При этом выборочное распределение оценки X при известной генеральной дисперсии σ^2 является также нормальным, и поэтому доверительные интервалы здесь находятся с использованием нормального распределения.

Доверительный интервал для математического ожидания при известной дисперсии σ^2 *генеральной совокупности:*

$$\left[\bar{X} - Z_p\frac{\sigma}{\sqrt{n}}; \bar{X} + Z_p\frac{\sigma}{\sqrt{n}}\right], \tag{16}$$

где $\bar{X}$ - оценка математического ожидания; n – объем выборки; σ - среднеквадратичное отклонение генеральной совокупности; Z_p - такое значение аргумента функции Лапласа, при котором:

$$\Phi(Z_p) = \frac{1}{2}P\,. \tag{17}$$

Доверительный интервал для математического ожидания при неизвестной дисперсии σ^2 *генеральной совокупности равен:*

$$\left[\bar{X} - t_{k,\frac{a}{2}}\sqrt{\frac{\Delta_*}{\sqrt{n}}}\right], \left[\bar{X} + t_{k,\frac{a}{2}}\sqrt{\frac{\Delta_*}{\sqrt{n}}}\right], \tag{18}$$

где Δ_* - несмещенная состоятельная оценка дисперсии (8); $t_{k,\frac{a}{2}}$ - аргумент функции плотности распределения Стьюдента, находится по таблицам для заданного уровня значимости a и числа степеней свободы $k = n - 1$.

Доверительный интервал для дисперсии при известном математическом ожидании μ случайной величины X определяется выражением

$$\frac{n\Delta_{**}}{X_2^{\,2}}, \frac{n\Delta_{**}}{X_1^{\,2}}, \qquad (19)$$

где n - объем выборки; Δ_{**} - состоятельная несмещенная и эффективная оценка дисперсии; $X_2^{\,2}, X_1^{\,2}$ - аргументы функции плотности распределения Пирсона. Площадь под кривой плотности, заключенная между точками $X_2^{\,2}, X_1^{\,2}$ равна выбранному значению доверительной вероятности Р.

Значения $X_2^{\,2}, X_1^{\,2}$ находятся по таблице для выбранной доверительной вероятности Р и числа степеней свободы $K{=}n$.

Доверительный интервал для дисперсии при неизвестном математическом ожидании случайной величины X определяется выражением:

$$\left[\frac{(n-1)\Delta_{*}}{X_2^{\,2}}, \frac{(n-1)\Delta_{*}}{X_1^{\,2}}\right], \qquad (20)$$

где n - объем выборки; Δ_{*} - состоятельная и несмещенная оценка дисперсии; $X_2^{\,2}, X_1^{\,2}$ - критические точки распределения Пирсона, находятся по таблице для числа степеней свободы k=n-1.

В большинстве практических случаев определяют верхний односторонний доверительный интервал для дисперсии, тем самым фиксируя верхнюю границу степени разброса исследуемого параметра с уровнем значимости а.

2.4. Практическая часть

Задача 1

Найдите исправленную дисперсию Δ_{*} для выборки объема $n{=}10$, если выборочная дисперсия равна Δ = 180.

Задача 2

Найдите несмещенную оценку математического ожидания измерений некоторой случайной величины одним прибором (без систематических ошибок), результаты измерения которой: 4,5,8,9,11.

Задача 3

Распределение случайной величины описывается плотностью вероятности $f(x)=4xe^{-2x}$, какова вероятность нахождения случайной величины в интервале от 0 до 1?

Задача 4

Найти двусторонний доверительный интервал для математического ожидания сопротивления резисторов по выборке: 102,3; 101,5; 92,8; 81,8; 71,4; 101; 104,2; 100,7; 102,8; 89,4; 109,7 Ом. Известна дисперсия сопротивления резисторов $\sigma^2=28\,Ом^2$. Принять доверительную вероятность P=0,98. Предварительно проверить выборку на выскакивающие варианты по уровню значимости a = 0,01.

Задача 5

Для 12 реле, случайным образом выбранных из партии, измерено напряжение срабатывания и вычислены среднее значение X = 27,4 В и состоятельная оценка дисперсии $\Delta=7B^2$. Какова вероятность того, что выборочное среднее оценивает математическое ожидание напряжения срабатывания для всей партии с точностью ±4 В?

Задача 6

Определить односторонний доверительный интервал для дисперсии выходного напряжения микросхем по выборке: 2,41; 2,38; 2,39; 2,37; 2.43; 2.37; 2.24; 2.27; 2.32; 2.30; 2.37; 2.31 В. Уровень значимости принять равным $a=0.05$.

3. СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ ПРОГРАММНЫМИ СРЕДСТВАМИ

Мощным инструментальным средством при выполнении статистических исследований является использование компьютерной техники. В этой связи широкое распространение получили специальные пакеты прикладных программ. Они позволяют обеспечить весьма впечатляющую быстроту статистических расчетов, высокую надежность и достоверность результатов, возможность легко представлять данные в аналитической, графической или табличной формах. Среди подобных программ большой известностью пользуется приложение Microsoft Excel, которое включает в себя программную надстройку "Пакет анализа" и богатую библиотеку из большого числа статистических функций, кроме того, используя режим конструктора и средства VBA, легко создавать свои узко специализированные макросы.

Рассмотрим примеры решения задач, возникающих при статистическом анализе. Пусть в результате экспериментальных исследований получена совокупность значений исследуемого параметра, приведенная в табл. 2. При применении для анализа электронных таблиц excel исследуемый одномерный параметр целесообразно записывать в столбец или строку для простоты организации перебора значений. Таблица 2 имеет такой вид для удобства записи в печатном формате.

Таблица 2 – Таблица значений исследуемого параметра

№	Изм.	№	Изм.	№	Изм.	№	Изм.	№	Изм.
1	54	**22**	56	**43**	52	**64**	60	**85**	58
2	52	**23**	58	**44**	55	**65**	62	**86**	59
3	58	**24**	55	**45**	53	**66**	65	**87**	56
4	55	**25**	54	**46**	55	**67**	58	**88**	54
5	55	**26**	51	**47**	55	**68**	54	**89**	45
6	54	**27**	52	**48**	56	**69**	71	**90**	49
7	55	**28**	53	**49**	56	**70**	52	**91**	48
8	55	**29**	57	**50**	54	**71**	51	**92**	49
9	57	**30**	53	**51**	58	**72**	53	**93**	49
10	48	**31**	5	**52**	57	**73**	52	**94**	48

11	48	**32**	54	**53**	59	**74**	54	**95**	49
12	55	**33**	55	**54**	58	**75**	58	**96**	54
13	45	**34**	62	**55**	58	**76**	59	**97**	55
14	55	**35**	62	**56**	44	**77**	56	**98**	58
15	59	**36**	63	**57**	49	**78**	57	**99**	51
16	55	**37**	69	**58**	49	**79**	55	**100**	52
17	58	**38**	55	**59**	49	**80**	51	**101**	53
18	58	**39**	59	**60**	63	**81**	52	**102**	52
19	56	**40**	59	**61**	62	**82**	53	**103**	«конец»
20	59	**41**	59	**62**	55	**83**	56	**104**	
21	58	**42**	54	**63**	59	**84**	54	**105**	

Для создания собственных макросов необходимо сохранить документ в формате, поддерживающем макросы, и включить в панели инструментов вкладку разработчика, в которой появится возможность вставить элементы управления формой. Самым простым и быстрым способом задания сценария работы макроса является кнопка (рис. 2), при нажатии которой выполняется программа, написанная в среде VBA. Для этого нужно указать имя макроса и нажать кнопку "создать".

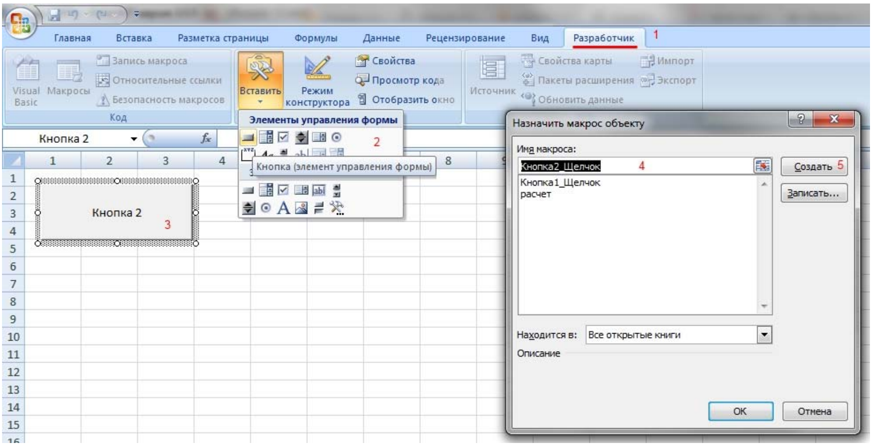


Рисунок 2 – Назначение макроса событию по нажатию кнопки

После создания макроса откроется окно (рис. 3), где необходимо написать исполняемую программу между заголовком Sub... и командой End Sub.

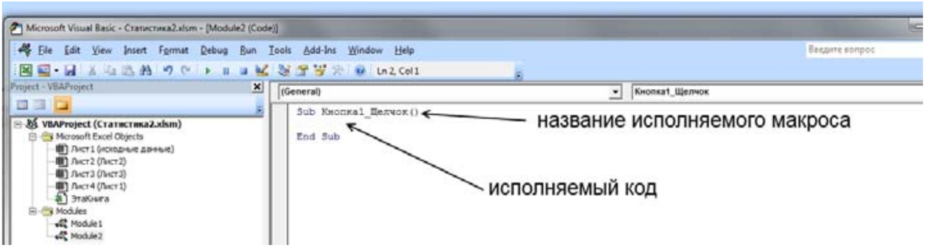


Рисунок 3 – Окно исполняемого кода макроса

Excel позволяет обращаться к ячейкам таблицы, для этого необходимо указать номер листа и координаты ячейки: *x=Лист2.Cells(3, 1),* переменной x присваивается значение ячейки первого столбца третьей строки, тип переменной определяется автоматически.

Рассмотрим синтаксис основных команд VBA, которые понадобятся для выполнения задания.

1. Условие

If
условие
Then
действие, если условие выполнено
Else
действие, если условие не выполнено
End If

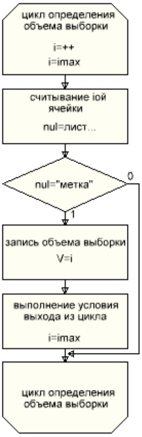
2. Цикл

For *начало счета = 0* To *конец счета*
тело цикла
Next *счет + 1*

3. Оператор перехода

GoTo *метка*
другой код
метка:

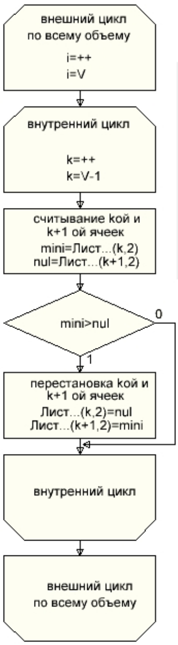
При большом количестве испытаний сначала необходимо определить объем совокупности (кол-во испытаний). Для этого можно использовать программу и соответствующий ей алгоритм, представленные ниже.




```
V = 0        ' объем совокупности (выборки)
nul = 0      ' текущее значение из таблицы
i = 0        ' счетчик цикла i
k = 0        ' счетчик цикла k

Лист2.Cells(3, 1) = "исходн.выб"    ' запись в ячейку 3.1 на втором листе строки: "исходн.выб"
For i = 0 To 10000                  ' начало цикла
    nul = Лист1.Cells(4 + i, 1)
    Лист2.Cells(4 + i, 1) = nul
    Лист2.Cells(4 + i, 2) = nul
    If nul = "конец" Then           ' проверка условия становки цикла
        V = i - 1
        i = 10000                   ' выход из цикла, достижением конечного значения
    End If
Next i                 ' конец цикла i = 10000
Лист2.Cells(1, 1) = "N" ' запись в ячейку 1.1 на втором листе строки: "N"
Лист2.Cells(1, 2) = V   ' запись в ячейку 1.2 на втором листе строки: "V"
' оббъем выборки найден
```

Почти всегда для статистического анализа необходимо построить вариационный ряд (провести ранжирование элементов совокупности по возрастанию). Для этого можно воспользоваться следующим кодом.




```
Лист2.Cells(3, 2) = "вар.ряд"   ' обозначение заголовка нового столбца
For i = 0 To V  ' внешний цикл, где V - найденный ранее объем выборки
    For k = 0 To V - 1
        mini = Лист2.Cells(4 + k, 2)     ' минимальное значение из двух
        nul = Лист2.Cells(5 + k, 2)      ' текущее значение параметра
        If mini > nul Then               ' при выполнении условия параметры меняются местами
            Лист2.Cells(4 + k, 2) = nul
            Лист2.Cells(5 + k, 2) = mini
        End If
    Next k
Next i
' вариационный ряд построен
```

Следующим этапом, как правило, идет проверка на «выскакивающие варианты» и нахождение точечных оценок по формулам (1)-(2) и (5)-(10) соответственно. Данный участок программы линейный и не представляет интереса.

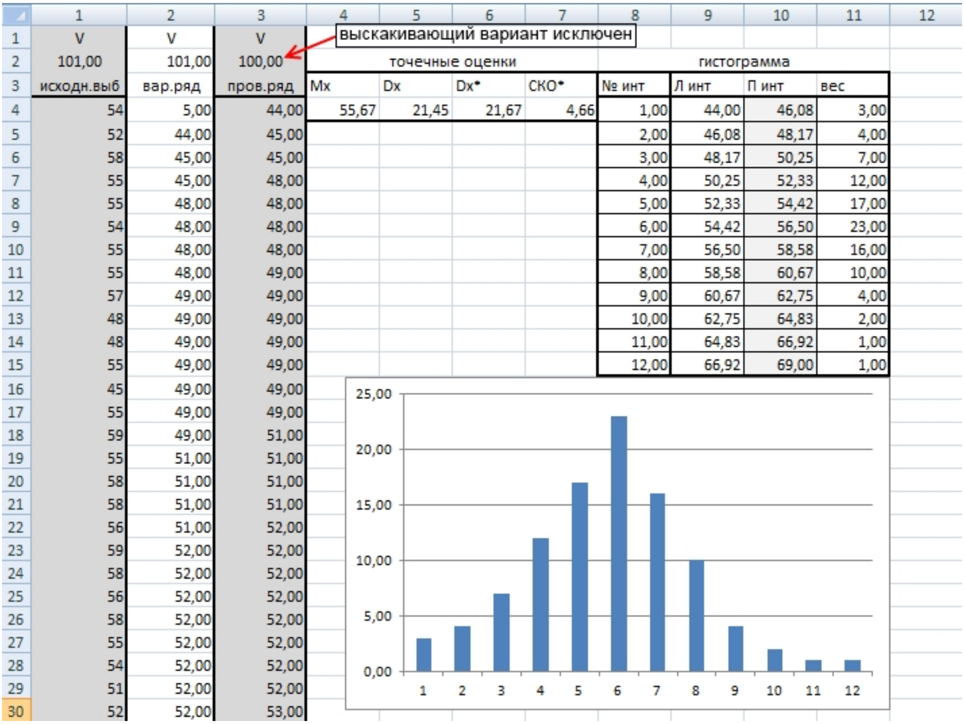
Наглядное представление о характере распределения случайной величины дает гистограмма, далее представлен участок программы, отвечающий за её выполнение на рис. 3 представлен результат обработки данных из табл. 2.

```
' !!! построение гистограммы !!!
nul = 0
detmax = 50     'максимальное кол-во интервалов (произвольно но влияет на скорость)
det = detmax    'кол-во интервалов (меняется при итерациях)
    For g = 0 To detmax 'перебор по текущему значению интервалов
        l = Лист2.Cells(3 + V, 3) - Лист2.Cells(4, 3)   ' вычисление всего диапазона значений
        dn = l / det                                     ' вычисление ширины интервала
        Лист2.Cells(4, 9) = Лист2.Cells(4, 3)            ' запись нижней границы
        For i = 0 To det - 1                ' цикл для записи и вычисления границ интервалов
            Лист2.Cells(4 + i, 8) = i + 1
            Лист2.Cells(4 + i, 10) = Лист2.Cells(4 + i, 9) + dn
            Лист2.Cells(5 + i, 9) = Лист2.Cells(4 + i, 9) + dn
        Next i

        Лист2.Cells(det + 4, 9) = ""

        c = 0               ' переменная для записи попаданий значений в текущий интервал
        For i = 0 To det    ' перебор интервалов
            lima = Лист2.Cells(4 + i, 9)        ' определение границ текущего интервала
            limb = Лист2.Cells(4 + i, 10)
                For k = 0 To V                  ' перебор вариант по всему объему
                    A = Лист2.Cells(4 + k, 3)   ' текущее значение
                    If (A >= lima) Then         ' проверка на условие попадания в границы
                        If (A < limb) Then
                           c = c + 1            ' счет попаданий в интервал
                           p = p + 1            ' общая сумма (в итоге должна быть равна V)
                        End If
                    End If
                Next k
            Лист2.Cells(4 + i, 11) = c          ' запись попаданий в интервал
            c = 0                               ' сброс переменной
        Next i

        For i = 0 To det                ' перебор всех интервалов для поиска нулевой частоты
            A = Лист2.Cells(4 + i, 11)  ' текущее значение частоты
            If A = 0 Then
                nul = nul + 1           ' если обнаружен нулевой интервал
            End If
        Next i
            If nul >= 1 Then
                det = det - 1
                nul = 0
            End If
    ' если нулевых интервалов больше чем 1, то кол-во интервалов уменьшается на 1
    Next g
' далее идут элементы оформления
For i = 0 To detmax
    Лист2.Cells(det + 4 + i, 8) = ""
    Лист2.Cells(det + 4 + i, 9) = ""
    Лист2.Cells(det + 4 + i, 10) = ""
    Лист2.Cells(det + 4 + i, 11) = ""
    Лист2.Cells(3, 8) = "№ инт"
    Лист2.Cells(3, 9) = "Л инт"
    Лист2.Cells(3, 10) = "П инт"
    Лист2.Cells(3, 11) = "вес"
Next i
' !!! построение гистограммы завершено !!!
```
** алгоритм для данного участка программы студентам предлагается составить самостоятельно*

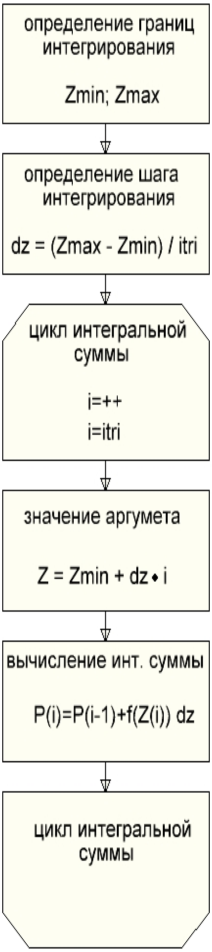


	1	2	3	4	5	6	7	8	9	10	11
1	V	V	V	выскакивающий вариант исключен							
2	101,00	101,00	100,00	точечные оценки				гистограмма			
3	исходн.выб	вар.ряд	пров.ряд	Mx	Dx	Dx*	СКО*	№ инт	Л инт	П инт	вес
4	54	5,00	44,00	55,67	21,45	21,67	4,66	1,00	44,00	46,08	3,00
5	52	44,00	45,00					2,00	46,08	48,17	4,00
6	58	45,00	45,00					3,00	48,17	50,25	7,00
7	55	45,00	48,00					4,00	50,25	52,33	12,00
8	55	48,00	48,00					5,00	52,33	54,42	17,00
9	54	48,00	48,00					6,00	54,42	56,50	23,00
10	55	48,00	48,00					7,00	56,50	58,58	16,00
11	55	48,00	49,00					8,00	58,58	60,67	10,00
12	57	49,00	49,00					9,00	60,67	62,75	4,00
13	48	49,00	49,00					10,00	62,75	64,83	2,00
14	48	49,00	49,00					11,00	64,83	66,92	1,00
15	55	49,00	49,00					12,00	66,92	69,00	1,00
16	45	49,00	49,00								
17	55	49,00	49,00								
18	59	49,00	51,00								
19	55	51,00	51,00								
20	58	51,00	51,00								
21	58	51,00	51,00								
22	56	51,00	52,00								
23	59	52,00	52,00								
24	58	52,00	52,00								
25	56	52,00	52,00								
26	58	52,00	52,00								
27	55	52,00	52,00								
28	54	52,00	52,00								
29	51	52,00	52,00								
30	52	52,00	53,00								

Рисунок 4 – пример результатов обработки массива данных (табл.2)

Для анализа довольно часто необходимо вычислять определенные интегралы. Для примера рассмотрим вычисление функции Лапласа $\Phi(Z)=\frac{1}{\sqrt{2\pi}}\int_{0}^{Z}e^{-\frac{z^2}{2}}dz$ численным методом. Пусть известен уровень значимости и необходимо найти аргумент Z. Тогда, разбив интервал доступных значений Z на множество небольших отрезков и вычисляя значение функции на каждом из них, можно примерно найти площадь фигуры, ограниченной границами интегрирования и самой функцией. Так как в данном примере необходимо найти не сам интеграл, а значение аргумента, то на каждом шаге необходимо сравнивать интегральную сумму с известной доверительной вероятностью. Как только условие будет выполнено (с заданной точностью), расчет необходимо остановить, а принять искомое значение аргумента равное текущему. Ниже представлена программа, позволяющая вычислять значение аргумента функции Лапласа по известной вероятности и упрощенный алгоритм. Алгоритм описывает принцип нахождения значения определенных интегралов методом прямоугольников.

Студентам предлагается дополнить данный алгоритм, чтобы он соответствовал коду программы самостоятельно. Предложить долее точный способ вычисления интегралов (трапеции, сплаин).




```
' !!! вычисление функции Лапласа!!!
Zmin = 0     ' нижний предел значения аргумента
Zmax = 5     ' верхний предел
itri = 10000     ' кол-во разбиений интервала при первой итерации
itrk = 1         ' кол-во итераций +1

For k = 0 To itrk    ' цикл по итерациям
    Pold = 0         ' переменные для мгновенных значений функции
    Pnew = 0
    i = 0

    dz = (Zmax - Zmin) / itri    ' шаг
    For i = 0 To itri            ' цикл интегральной суммы
        c = c + 1
        Z = Zmin + dz * i        ' значение аргумета
        ' далее вычисление интегральной суммы
        Pnew = Pold + (1 / (Sqr(2 * 3.14159265358979))) * Exp(-(Z * Z) / 2) * dz
        ' далее проверка на равенство интегральной суммы и доверительной вероятности
        If Pnew >= Pmx Then
            i = itri
            Zmax = Z + dz * 100
            itri = 30000
        Else
            Pold = Pnew
        End If
    Next i
Next k
' далее оформление и запись результатов
Лист2.Cells(6, 4) = "Z"
Лист2.Cells(6, 5) = "P"
Лист2.Cells(7, 4) = Z
Лист2.Cells(7, 5) = Pnew
' !!! вычисление функции Лапласа завершено!!!
```

Вычисляя значение аргумента функции Лапласа в ходе выполнения программы, появляется возможность определять доверительные интервалы для точечных оценок или решать обратные задачи - например определить необходимый объем выборки из генеральной совокупности для дальнейшего анализа с заданной точностью. На рис. 5 представлен более полный анализ данных из табл. 2.

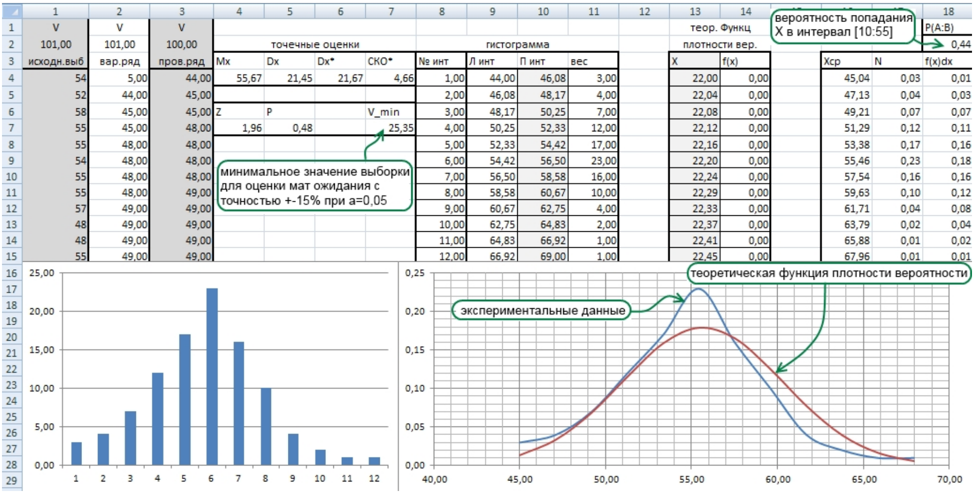


Рисунок 5 – Пример программной обработки массива данных (табл. 2)

4. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Дайте определение понятию «вариационный ряд».
2. Какие точечные оценки параметров распределения вам известны?
3. Какие законы распределения случайной величины вы знаете, какими особенностями они обладают?
4. Что такое нормальное распределение случайной величины?
5. Что такое интервальная оценка?
6. Какой физический смысл функции Лапласа?
7. Дайте определение понятию «доверительный интервал».
8. Дайте ответы на вопросы в сносках в третьем разделе.

5. ИНДИВИДУАЛЬНОЕ ЗАДАНИЕ

На испытание поставлена партия деталей в количестве 2000 шт. были проведены идентичные измерения некоторого параметра. Необходимо провести полный статистический анализ генеральной совокупности (2000шт.), после чего определить необходимый объем *Nв* выборки, оценивающей математическое ожидание с точностью *±10%* при вероятности принятия ошибки *a=0,05,* что позволит сократить количество испытаний для следующих партий деталей от данного производителя. Для проверки полученных результатов провести выборку объема *Nв* из генеральной совокупности и сравнить их статистические параметры. Сделать выводы о возможности проведения выборочных испытаний из партий.

Порядок выполнения индивидуального задания

С помощью программы genreatorPSV сгенерировать совокупность данных с параметрами, указанными преподавателем.

Разработать макрос в среде excel для проведения необходимых статистических исследований.

Провести статистический анализ.

Написать отчет о проделанной работе и о результатах исследований.

Отчет должен содержать:

- титульный лист,
- исходное задание,
- формулировку задачи,
- этапы выполнения индивидуального задания,
- заключение.

Этапы выполнения индивидуального задания:

- составление легенды по полученной выборке;
- составление вариационного ряда;
- предварительная обработка статистических данных (проверка на выскакивающие варианты);

- разбить диапазон значений случайной величины на интервалы и построить гистограмму распределения случайной величины и экспериментальный график плотности вероятности (частоты);
- произвести точечные оценки параметров распределения (оценка мат.ожидания и дисперсий);
- по полученным значениям построить теоретический график нормального распределения и сравнить с графиком для частоты, полученным ранее;
- вычислить, используя полученную теоретическую функцию плотности, вероятность попадания случайной величины в интервал *(0;Mx);*
- Рассчитать необходимый объем выборки для оценки мат. ожидания генеральной совокупности по ней с заданной точностью;
- провести сравнительный анализ выборки рассчитанного объема и генеральной совокупности;
- сделать выводы из анализа.

Отчет необходимо сдать в электронном и бумажном видах, к электронной версии приложить файл, содержащий разработанную программу.

Бумажную версию отчета оформить согласно СТО Самарского университета.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. М.: Наука, 1976.
2. Ашмарин И.П. Быстрые методы статистической обработки и планирования эксперимента. ЛГУ, 1975.
3. Смирнов В.Н., Бунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. М.: Наука, 1975.
4. Гарбер Г.Р. Основы программирования на Visual Basic и VBA в Excel2007. М.: Солонн-Пресс, 2008.
5. Михеев Р. VBA и программирование в MS OFFICE. Специальный курс по информационным технологиям. СПб.: ASKI, 2010.
6. Базар В.Р., Пегашкин В.Ф. Использование MS EXCEL для анализа статистических данных. 2-е изд. доп. Нижний Тагил, 2014.
7. Зеленский В.А. Основы конструкторско-технологического проектирования радиоэлектронных средств: учебное пособие. Самара: Изд-во СГАУ, 2016. 80 с.

ПРИЛОЖЕНИЕ

Аргумент функции распределения плотности Пирсона

k /α	0,01	0,025	0,05	k /α	0,01	0,025	0,05
1	6,6349	5,02389	3,84146	26	45,64168	41,92317	38,88514
2	9,21034	7,37776	5,99146	27	46,96294	43,19451	40,11327
3	11,34487	9,3484	7,81473	28	48,27824	44,46079	41,33714
4	13,2767	11,14329	9,48773	29	49,58788	45,72229	42,55697
5	15,08627	12,8325	11,0705	30	50,89218	46,97924	43,77297
6	16,81189	14,44938	12,59159	31	52,19139	48,23189	44,98534
7	18,47531	16,01276	14,06714	32	53,48577	49,48044	46,19426
8	20,09024	17,53455	15,50731	33	54,77554	50,72508	47,39988
9	21,66599	19,02277	16,91898	34	56,06091	51,966	48,60237
10	23,20925	20,48318	18,30704	35	57,34207	53,20335	49,80185
11	24,72497	21,92005	19,67514	36	58,61921	54,43729	50,99846
12	26,21697	23,33666	21,02607	37	59,8925	55,66797	52,19232
13	27,68825	24,7356	22,36203	38	61,16209	56,89552	53,38354
14	29,14124	26,11895	23,68479	39	62,42812	58,12006	54,57223
15	30,57791	27,48839	24,99579	40	63,69074	59,34171	55,75848
16	31,99993	28,84535	26,29623	41	64,95007	60,56057	56,94239
17	33,40866	30,19101	27,58711	42	66,20624	61,77676	58,12404
18	34,80531	31,52638	28,8693	43	67,45935	62,99036	59,30351
19	36,19087	32,85233	30,14353	44	68,70951	64,20146	60,48089
20	37,56623	34,16961	31,41043	45	69,95683	65,41016	61,65623
21	38,93217	35,47888	32,67057	46	71,2014	66,61653	62,82962
22	40,28936	36,78071	33,92444	47	72,44331	67,82065	64,00111
23	41,6384	38,07563	35,17246	48	73,68264	69,02259	65,17077
24	42,97982	39,36408	36,41503	49	74,91947	70,22241	66,33865
25	44,3141	40,64647	37,65248	50	76,15389	71,4202	67,50481

Таблица распределения Стьюдента

k	a=0,05	a=0,01	a=0,001	k	a=0,05	a=0,01	a=0,001
1	12,7	63,65	636,61	21	2,08	2,831	3,819
2	4,303	9,925	31,602	22	2,074	2,819	3,792
3	3,182	5,841	12,923	23	2,069	2,807	3,768
4	2,776	4,604	8,610	24	2,064	2,797	3,745
5	2,571	4,032	6,869	25	2,06	2,787	3,725
6	2,447	3,707	5,959	26	2,056	2,779	3,707
7	2,365	3,499	5,408	27	2,052	2,771	3,690
8	2,306	3,355	5,041	28	2,049	2,763	3,674
9	2,262	3,250	4,781	29	2,045	2,756	3,659
10	2,228	3,169	4,587	30	2,042	2,750	3,646
11	2,201	3,106	4,437	31	2,04	2,744	3,633
12	2,179	3,055	4,318	32	2,037	2,738	3,622
13	2,16	3,012	4,221	33	2,035	2,733	3,611
14	2,145	2,977	4,140	34	2,032	2,728	3,601
15	2,131	2,947	4,073	35	2,03	2,724	3,591
16	2,12	2,921	4,015	36	2,028	2,719	3,582
17	2,11	2,898	3,965	37	2,026	2,715	3,574
18	2,101	2,878	3,922	38	2,024	2,712	3,566
19	2,093	2,861	3,883	39	2,023	2,708	3,558
20	2,086	2,845	3,850	40	2,021	2,704	3,551

Таблицы Лапласа

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0	0	0,5	0,1915	1	0,3413	1,5	0,4332	2	0,4772	3	0,49865
0,01	0,004	0,51	0,195	1,01	0,3438	1,51	0,4345	2,02	0,4783	3,2	0,49931
0,02	0,008	0,52	0,1985	1,02	0,3461	1,52	0,4357	2,04	0,4793	3,4	0,49966
0,03	0,012	0,53	0,2019	1,03	0,3485	1,53	0,437	2,06	0,4803	3,6	0,49984
0,04	0,016	0,54	0,2054	1,04	0,3508	1,54	0,4382	2,08	0,4812	3,8	0,49993
0,05	0,0199	0,55	0,2088	1,05	0,3531	1,55	0,4394	2,1	0,4821	4	0,49997
0,06	0,0239	0,56	0,2123	1,06	0,3554	1,56	0,4406	2,12	0,483	4,5	0,5
0,07	0,0279	0,57	0,2157	1,07	0,3577	1,57	0,4418	2,14	0,4838	5	0,5
0,08	0,0319	0,58	0,219	1,08	0,3599	1,58	0,4429	2,16	0,4846		
0,09	0,0359	0,59	0,2224	1,09	0,3621	1,59	0,4441	2,18	0,4854		
0,1	0,0398	0,6	0,2257	1,1	0,3643	1,6	0,4452	2,2	0,4861		
0,11	0,0438	0,61	0,2291	1,11	0,3665	1,61	0,4463	2,22	0,4868		
0,12	0,0478	0,62	0,2324	1,12	0,3686	1,62	0,4474	2,24	0,4875		
0,13	0,0517	0,63	0,2357	1,13	0,3708	1,63	0,4484	2,26	0,4881		
0,14	0,0557	0,64	0,2389	1,14	0,3729	1,64	0,4495	2,28	0,4887		
0,15	0,0596	0,65	0,2422	1,15	0,3749	1,65	0,4505	2,3	0,4893		
0,16	0,0636	0,66	0,2454	1,16	0,377	1,66	0,4515	2,32	0,4898		
0,17	0,0675	0,67	0,2486	1,17	0,379	1,67	0,4525	2,34	0,4904		
0,18	0,0714	0,68	0,2517	1,18	0,381	1,68	0,4535	2,36	0,4909		
0,19	0,0753	0,69	0,2549	1,19	0,383	1,69	0,4545	2,38	0,4913		
0,2	0,0793	0,7	0,258	1,2	0,3849	1,7	0,4554	2,4	0,4918		
0,21	0,0832	0,71	0,2611	1,21	0,3869	1,71	0,4564	2,42	0,4922		
0,22	0,0871	0,72	0,2642	1,22	0,3883	1,72	0,4573	2,44	0,4927		
0,23	0,091	0,73	0,2673	1,23	0,3907	1,73	0,4582	2,46	0,4931		
0,24	0,0948	0,74	0,2703	1,24	0,3925	1,74	0,4591	2,48	0,4934		
0,25	0,0987	0,75	0,2734	1,25	0,3944	1,75	0,4599	2,5	0,4938		
0,26	0,1026	0,76	0,2764	1,26	0,3962	1,76	0,4608	2,52	0,4941		
0,27	0,1064	0,77	0,2794	1,27	0,398	1,77	0,4616	2,54	0,4945		
0,28	0,1103	0,78	0,2823	1,28	0,3997	1,78	0,4625	2,56	0,4948		
0,29	0,1141	0,79	0,2852	1,29	0,4015	1,79	0,4633	2,58	0,4951		
0,3	0,1179	0,8	0,2881	1,3	0,4032	1,8	0,4641	2,6	0,4953		
0,31	0,1217	0,81	0,291	1,31	0,4049	1,81	0,4649	2,62	0,4956		
0,32	0,1255	0,82	0,2939	1,32	0,4066	1,82	0,4656	2,64	0,4959		
0,33	0,1293	0,83	0,2967	1,33	0,4082	1,83	0,4664	2,66	0,4961		
0,34	0,1331	0,84	0,2995	1,34	0,4099	1,84	0,4671	2,68	0,4963		
0,35	0,1368	0,85	0,3023	1,35	0,4115	1,85	0,4678	2,7	0,4965		
0,36	0,1406	0,86	0,3051	1,36	0,4131	1,86	0,4686	2,72	0,4967		
0,37	0,1443	0,87	0,3078	1,37	0,4147	1,87	0,4693	2,74	0,4969		
0,38	0,148	0,88	0,3106	1,38	0,4162	1,88	0,4699	2,76	0,4971		
0,39	0,1517	0,89	0,3133	1,39	0,4177	1,89	0,4706	2,78	0,4973		
0,4	0,1554	0,9	0,3159	1,4	0,4192	1,9	0,4713	2,8	0,4974		
0,41	0,1591	0,91	0,3186	1,41	0,4207	1,91	0,4719	2,82	0,4976		
0,42	0,1628	0,92	0,3212	1,42	0,4222	1,92	0,4726	2,84	0,4977		
0,43	0,1664	0,93	0,3238	1,43	0,4236	1,93	0,4732	2,86	0,4979		
0,44	0,17	0,94	0,3264	1,44	0,4251	1,94	0,4738	2,88	0,498		
0,45	0,1736	0,95	0,3289	1,45	0,4265	1,95	0,4744	2,9	0,4981		
0,46	0,1772	0,96	0,3315	1,46	0,4279	1,96	0,475	2,92	0,4982		
0,47	0,1808	0,97	0,334	1,47	0,4292	1,97	0,4756	2,94	0,4984		
0,48	0,1844	0,98	0,3365	1,48	0,4306	1,98	0,4761	2,96	0,4985		
0,49	0,1879	0,99	0,3389	1,49	0,4319	1,99	0,4767	2,98	0,4986		