Министерство транспорта Российской Федерации
Федеральное агентство железнодорожного транспорта
ГОУ ВПО "Дальневосточный государственный университет
путей сообщения"

Кафедра "Высшая математика"

Г.П. Кузнецова, О.А. Горкуша

НАХОЖДЕНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН ПО РЕЗУЛЬТАТАМ ЭКСПЕРИМЕНТА.

Методические указания

Авторы — Г. П. Кузнецова,

О. А. Горкуша.

Зав. каф. "Высшая математика" — В. И. Жукова

Председатель МК ЕНФ — С.Ю. Ситникова

Председатель РИК ЕНФ — К.В. Пупатенко

Ответственный по кафедре "Высшая математика" — Л.Г. Гамоля

Хабаровск
Издательство ДВГУПС
2007

Министерство транспорта Российской Федерации
Федеральное агентство железнодорожного транспорта
ГОУ ВПО "Дальневосточный государственный университет
путей сообщения"

Кафедра "Высшая математика"

Г.П. Кузнецова, О.А. Горкуша

НАХОЖДЕНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН ПО РЕЗУЛЬТАТАМ ЭКСПЕРИМЕНТА.

Методические указания

Хабаровск
Издательство ДВГУПС
2007

УДК 519.2(075.8)
ББК 171.3 я 73
К 891

Рецензент:
Заведующий кафедрой "Высшая математика"ДВГУПС,
кандидат физико-математических наук,
доцент *В.И. Жукова*

К 891 Нахождение законов распределения случайных величин по результатам эксперимента: метод. указания / Г.П. Кузнецова, О.А. Горкуша. —Хабаровск: Изд—во ДВГУПС, 2007 - 23 с.: ил. - 5.

Методические указания соответствуют государственному образовательному стандарту дисциплины "Математика".

В данной работе изложены методы получения оценок законов распределения случайных величин и способ проверки гипотез о законах распределения случайных величин, основанный на критерии χ^2. Приведены примеры нахождения законов, даны варианты индивидуальных заданий.

Методические указания рекомендуются студентам всех специальностей технического университета, изучающих соответствующий раздел математической статистики.

УДК 519.2(075.8)
ББК 171.3 я 73

Введение

При проведении случайного эксперимента одним из основных факторов, интересующих исследователя, является закон распределения случайной величины. Закон устанавливает связь между всеми возможными значениями случайной величины и соответствующими вероятностями. Однако, часто в силу различных причин, получить все значения случайной величины и их вероятности нельзя.

В математической статистике изучаются методы обработки опытных данных, полученных в результате наблюдений. В частности, решаются такие задачи:

1) по данному статистическому материалу оценить, хотя бы приближенно, интересующие нас числовые характеристики случайной величины, например, математическое ожидание, дисперсию, построить доверительный интервал для математического ожидания;

2) получить закон распределения (в действительности, оценку закона) случайной величины по результатам опытов. Так как при этом неизбежно присутствуют элементы случайности, то выдвигаемая гипотеза о законе распределения проверяется на правдоподобие.

Настоящие методические указания предлагают студенту побыть в роли экспериментатора и решить указанные задачи, используя предложенные опытные данные, правда, уже частично обработанные.

При выполнении лабораторной работы студенту предстоит ознакомиться с базовыми понятиями математической статистики, рассмотреть предложенные примеры нахождения законов распределения случайной величины. Ее заключительным этапом является выполнение индивидуального задания по по одному из предложенных вариантов. Там же приведены теоретические вопросы, на которые необходимо ответить при защите работы. В конце прилагается список литературы для более полного изучения.

Содержание настоящих методических указаний соответствует требованиям Государственного образовательного стандарта по математике. Пособие предназначено для студентов всех специальностей университета, изучающих данный раздел математики.

1. ПОСТРОЕНИЕ СТАТИСТИЧЕСКОГО РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Одной из основных характеристик случайной величины X (в дальнейшем - с.в.) является закон распределения с.в..

Существует несколько законов распределения. Здесь мы подробно рассмотрим некоторые из них.

Обозначим через X - с.в. Проведем над X n независимых наблюдений:

$$x_1, x_2, \ldots, x_n. \tag{1}$$

Сгруппируем одинаковые значения этой выборки. Тогда (1) перепишется так:

$$\begin{gathered} x_1 \text{ встречается } n_1 \text{ раз,} \\ x_2 \text{ встречается } n_2 \text{ раза,} \\ \cdot\quad\cdot\quad\cdot \\ x_k \text{ встречается } n_k \text{ раз.} \end{gathered}$$

При этом

$$\sum_{i=1}^{k} n_i = n.$$

Число n_i называется частотой x_i, а n_i/n - относительной частотой значения x_i. Легко заметить, что

$$\sum_{i=1}^{k} n_i/n = 1. \tag{2}$$

Определение 1. Статистический ряд распределения с. в. X - это соответствие между значениями x_i и относительными частотами n_i/n.

Следуя этому определению, построим статистический ряд распределения и представим его двумя способами. Первый способ применим для *дискретной с.в.* - табличный способ:

Таблица 1

x_i	x_1	x_2	$\ldots$	x_k
n_i/n	n_1/n	n_2/n	$\ldots$	n_k/n

Опишем второй способ - построение статистического ряда распределения для *непрерывной с.в.*.

Разобьем весь диапазон наблюдаемых значений с.в. X на k интервалов (x_i, x_{i+1}) одинаковой длины $h = (x_{max} - x_{min})/k$. Тогда статистический ряд в этом случае имеет вид:

Таблица 2

(x_i, x_{i+1})	(x_1, x_2)	(x_2, x_3)	$\dots$	(x_k, x_{k+1})
n_i/n	n_1/n	n_2/n	$\dots$	n_k/n

.

Здесь $n_i/n-$ относительная частота значений $X \in (x_i, x_{i+1})$. И для n_i/n выполняется (2).

Полученный по выборке статистический ряд является аналогом закона распределения с.в. в теории вероятностей.

Еще одно представление статистического ряда - построение графика. Для дискретной с.в. график представлен на рис. 1 в виде полигона частот, для непрерывной с.в. - на рис. 2 (гистограмма).

Рис.1. Полигон

Рис.2. Гистограмма

Полигон относительных частот строят для дискретной с.в. X. Для этого на оси абсцисс откладывают значения x_i, а на оси ординат - соответствующие им относительные частоты n_i/n. Ломаная, отрезки которой соединяют точки $(x_1, n_1/n)$, $(x_2, n_2/n), \cdots, (x_k, n_k/n)$ есть полигон частот (рис.1).

Если $X-$ непрерывная с.в., то по ряду таблицы 2 строим гистограмму. Это ступенчатая фигура, состоящая из прямоугольников, площадь каждого из которых равна относительной частоте. Для этого на оси абсцисс откладываются частичные интервалы и на каждом из них строят прямоугольник высотой, равной относительной частоте, деленной на длину интервала h. При этом площадь S_i прямоугольника равна n_i/n, а площадь всей гистограммы - сумме относительных частот, то есть единице (рис.2).

2.ВЫРАВНИВАНИЕ СТАТИСТИЧЕСКИХ РАСПРЕДЕЛЕНИЙ

Графическое представление статистического ряда помогает решить, какой закон распределения имеет с.в.. Подобный выбор закона называется выравниванием статистических рядов. В этом случае график заменяется плавной кривой (функцией плотности распределения $f(x)$), которая хорошо вписывается в анализируемую фигуру и имеет простое аналитическое выражение.

Мы подробно остановимся на двух распределениях:

нормальном (рис.3)

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mathbf{x}-\mathbf{m})^2}{2\sigma^2}} = N(m, \sigma^2);\ m = M(X); \sigma^2 = D(X) \tag{3}$$

и показательном (рис.4)

$$f(x) = \lambda \cdot e^{-\lambda \cdot x}, \lambda = \frac{1}{M(X)}, \lambda > 0, x \geq 0. \tag{4}$$

Рис.3. Нормальный закон Рис.4. Показательный закон

Отметим, что законы распределения с.в. содержат параметры, характеризующие эту величину. Значения этих параметров выбирают из оценок с.в., полученных из опытных данных: $\overline{X}, S^2$ - для нормального закона, и $\lambda = 1/\overline{X}$ - для показательного закона. Тогда полученный закон будет на самом деле оценкой закона распределения.

3. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ И КРИТЕРИЙ χ^2 ПИРСОНА

Далее, выдвинутая гипотеза нуждается в проверке- согласуется ли она с опытными данными или противоречит им. Для этого используют *статистический критерий.* Это правило, указывающее, когда статистическую гипотезу следует принять, когда отвергнуть. Одним из таких критериев является критерий χ^2 Пирсона.

Изложим его идею. Разобьем множество всех наблюдаемых значений с.в.

X на k разрядов (k интервалов или k различных значений). Пусть по результатам наблюдений с.в. составлен статистический ряд (таблица 1, таблица 2).

В случае, когда $X-$ непрерывная с.в., вместо интервалов в верхней строке таблицы 2 берем середины этих интервалов.

Согласно выдвинутой гипотезе, найдем теоретические вероятности того, что X попадает в соответствующий разряд:

$$p_i = p(X = x_i) - \text{ для таблицы 1,}$$

$$p_i = \int_{x_i}^{x_{i+1}} f(x)dx - \text{ для таблицы 2.}$$

Если гипотеза верна, то при большом числе наблюдений

$$\frac{n_i}{n} \approx p_i. \tag{5}$$

Причем, возможные расхождения объясняются случайными причинами. Если гипотеза не верна, то (5) не выполняется. В качестве меры расхождения между теоретическими и статистическими распределениями берется следующая величина - с.в. χ^2 :

$$\sum_{i=1}^{k} \frac{n}{p_i}\left(\frac{n_i}{n} - p_i\right)^2 = \sum_{i=1}^{k} \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i} = \chi^2. \tag{6}$$

Пирсон показал, что χ^2 имеет распределение, которое не зависит от закона распределения с.в. X, а зависит от параметра $r-$ числа степеней свободы

$$r = k - l, \tag{7}$$

где $k-$ число разрядов, $l-$ число связей, наложенных на частоты n_i в каждом конкретном законе для X. Эти связи и условия имеют вид:

$$\sum_{i=1}^{k} n_i = n, \ \frac{\sum_{i=1}^{k} x_i \cdot n_i}{n} = \overline{X}, \frac{\sum_{i=1}^{k} (x_i - X)^2 \cdot n_i}{n-1} = S^2. \tag{8}$$

Тогда для нормального распределения $l = 3$, для показательного закона и закона Пуассона $l = 2$.

Функция плотности $\Psi_r(\chi^2)$ для с.в. χ^2 имеет сложное аналитическое выражение. Вид графика представлен на рис 5.

Рис.5. Закон распределения χ^2

Поэтому мы воспользуемся специальными таблицами (см. Таблицу 1 в приложении), в которых для полученного значения r и заданной вероятности β указаны такие значения χ^2_β, что

$$p(\chi^2 \geq \chi^2_\beta) = \int_{\chi^2_\beta}^{\infty} \Psi_r(u)du = \beta. \qquad (9)$$

На рис. 5 заштрихованная площадь равна β. Вероятность $p(\chi^2 \geq \chi^2_\beta)$ можно понимать, как вероятность того, что в силу случайных факторов мера расхождения χ^2 между гипотезой и экспериментом больше или равна значению χ^2_β.

Рассмотрим схему применения критерия χ^2.

Предположим, что гипотеза верна. Выберем вероятность β настолько малой, что событие, имеющее эту вероятность можно считать практически невозможным. По заданным β и r из таблицы 1 приложения найдем χ^2_β. Это максимально возможная мера расхождения между опытными данными и гипотезой, когда гипотезу еще можно принять. Вычислим по формуле (6) значение χ^2 и сравним его с χ^2_β.

Если $\chi^2 > \chi^2_\beta$, то выдвинутую гипотезу следует отвергнуть, так как расхождения между n_i/n и p_i оказались слишком велики и случайными их признать нельзя.

Вывод: гипотеза противоречит опытным данным.

Если $\chi^2 < \chi^2_\beta$, то значения χ^2 не являются слишком большими и расхождения могли произойти из-за случайности выборки.

Вывод: гипотеза не противоречит опытным данным.

Во всех этих случаях мы ничего не можем сказать о верности гипотезы. Так, в первом случае ($\chi^2 > \chi^2_\beta$) с вероятностью, равной β мы отвергаем верную гипотезу, а во втором говорим о не противоречивости гипотезы.

4. ПРИМЕРЫ ВЫПОЛНЕНИЯ ЛАБОРАТОРНЫХ РАБОТ

Пример 1. Произведено $n = 100$ измерений некоторой непрерывной случайной величины. Результаты представлены в таблице 4.1 в виде статистического ряда.

Таблица 4.1

Результаты измерений случайной величины

Интервалы значений	(0,2)	(2,4)	(4,6)	(6,8)	(8,10)	(10,12)
Число попаданий	3	16	28	32	14	7

Построить гистограмму, выдвинуть гипотезу о законе распределения с.в. и проверить, насколько она согласуется с экспериментом.

Составим статистический ряд, аналогичный ряду, представленному в таблице 2:

Таблица 4.2

Статистический ряд

Середина (x_i, x_{i+1})	1 (0,2)	3 (2,4)	5 (4,6)	7 (6,8)	9 (8,10)	11 (10,12)
n_i/n	0.03	0.16	0.28	0.32	0.14	0.07
p_i	0.18		0.29	0.30	0.16	0.05

Для построения гистограммы над каждым интервалом строим прямоугольник высотой $H_i = n_i/(2n)$, где 2− длина интервала (рис. 6)

$$H_1 = \frac{0.03}{2} = 0.015; H_2 = \frac{0.16}{2} = 0.08; H_3 = \frac{0.28}{2} = 0.14; \cdots$$

По виду ступенчатой фигуры можно предположить, что с.в. имеет нормальный закон распределения (3) $N(m, \sigma^2)$. Значения m и σ возьмем равными соответствующим оценкам. В качестве "представителя"каждого интервала возьмем его середину:

$$m \approx \overline{X} = \frac{\sum_{i=1}^{6} x_i n_i}{n} = \frac{1 \cdot 3 + 3 \cdot 16 + 5 \cdot 28 + 7 \cdot 32 + 9 \cdot 14 + 1 \cdot 17}{100} = 6.18;$$

$$\sigma^2 \approx S^2 = \frac{\sum_{i=1}^{6} (x_i - \overline{X}^2)}{n-1} =$$

$$= \frac{(1-6.18)^2 \cdot 3 + (3-6.18)^2 \cdot 16 + \cdots + (11-6.18)^2 \cdot 7}{99} =$$
$$= 3.83; \ S = 2.41.$$

Итак, выдвигаем гипотезу о том, что изучается с.в. с функцией плотности $f(x)$ (рис. 3):

$$f(x) = \frac{1}{\sqrt{2\pi}2,41} \cdot e^{-\frac{(\mathbf{X}-\mathbf{6,18})^{\mathbf{2}}}{\mathbf{2 \cdot 5,83}}} = N(6,18; 5,83). \qquad (10)$$

Рис.6. Гистограмма и функция плотности

Эта кривая представлена на рис. 6. При ее построении достаточно найти значение $f(x)$ в точке максимума $x = \overline{X} = 6.18$ и в точках перегиба $x = \overline{X} \pm S = 6.18 \pm 2.41$. Затем следует эти точки соединить плавной кривой, учитывая форму кривой нормального распределения. Для построения $f(x)$ вместо непосредственного вычисления по формуле (10) удобно пользоваться функцией

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{-\mathbf{t}^{\mathbf{2}}}{\mathbf{2}}} \text{ существуют специальные таблицы [1].}$$

Тогда

$$f(x) = \frac{1}{2.41}\varphi\left(\frac{x-6.18}{2.41}\right).$$

В частности,

$$f(6.18) = \frac{\varphi(0)}{2.41} = \frac{0.3989}{2.41} = 0.166; f(6.18 \pm 2.41) = \frac{\varphi(1)}{2.41} = \frac{0.2420}{2.41} = 0,1.$$

Зададим уровень значимости β. Эта величина принимает любое значение из табл. 1 приложения. Пусть $\beta = 0,05$. Для улучшения надежности объединим первый интервал, содержащий мало значений, со вторым. Тогда $k = 5$ и число степеней свободы (7) для χ^2 равно $k - 3 = 2$. По табл. 1 приложения по

$r = 2$ и β находим $\chi^2_\beta = 5.99$. Критической областью в данном случае будет интервал $[5.99; \infty)$.

Вычислим χ^2 по формуле (6). Для этого потребуется вероятность попадания с.в. X на каждый интервал, согласно выдвинутой гипотезе:

$$P_1 = P(0 < X < 4) = \int_0^4 f(x)dx = \Phi(\frac{4-6.18}{2.41}) - \Phi(\frac{0-6.18}{2.41}) =$$

$$= \Phi(-0.9) - \Phi(2.56) = 0.1787;$$

$$P_2 = P(4 < X < 6) = \Phi(\frac{6-6.18}{2.41}) - \Phi(-0.9) =$$

$$-\Phi(0.07) + \Phi(0.9) = -0.0279 + 0.3159 = 0.288; \cdots$$

Здесь функция $\Phi(x)-$ функция Лапласа. Ее значения приведены в таблице 2 приложения. При этом следует учесть, что

$$\Phi(-x) = -\Phi(x).$$

Для сравнения результаты вычислений внесены третьей строкой в таблице 4.2. Вычисления χ^2 рекомендуется оформить в виде таблицы 4.3.

Таблица 4.3

Вычисление функции χ^2.

n_i	p_i	$n \cdot p_i$	$n_i - n \cdot p_i$	$(n_i - n \cdot p_i)^2$	$(n_i - n \cdot p_i)^2/n \cdot p_i$
19	0.179	17.9	1.1	1.21	0.068
28	0.288	28.8	0.8	0.64	0.022
32	0.304	30.4	1.6	2.56	0.084
14	0.165	16.5	-2.5	6.25	0.379
7	0.051	5.1	1.9	3.61	0.708
$n = 100$					$\chi^2 = 1.261$

Величина χ^2 равна сумме значений последнего столбца таблицы 4.3. Значение $\chi^2 = 1.261$ не входит в критическую область: $1,261 < 5,99$.

Вывод: Гипотеза не противоречит опытным данным. И вероятность того, что мы при этом ошибаемся, меньше 0,05.

Замечание. Если изменить уровень значимости β, то изменится критическая область, а значит, и ошибка прогноза. И может оказаться, что в этих условиях значение χ^2 попадает в критическую область.

Пример 2. Для исследования надежности некоторого прибора произвели подсчет числа выходов его из строя в течении 100 дней. Результаты представлены в виде статистического ряда в таблице 4.4.

Таблица 4.4

Результаты исследования

Число поломок в день x_i	0	1	2	3	4 и более
Число дней с данным числом поломок	54	27	14	5	0

Построим полигон относительных частот (рис. 7), предварительно составив ряд в виде таблицы 1. Пунктирная линия - это многоугольник распределения, сплошная линия - полигон частот.

Рис.7. Полигон и многоугольник распределения

По виду полигона частот можно выдвинуть предположение, что изучаемая случайная величина имеет Пуассоновский закон распределения:

$$P_k = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Оценим значение λ. Известно, что в законе распределения Пуассона значение параметра λ равно математическому ожиданию случайной величины, а оценкой математического ожидания является

$$\overline{X} = \frac{\sum_i x_i \cdot n_i}{n} = \frac{0 \cdot 54 + 1 \cdot 27 + 2 \cdot 14 + 3 \cdot 5 + 4 \cdot 0}{100} = 0,7.$$

Значит, и оценкой λ будет значение 0,7. Выдвигаем гипотезу, что изучаемая с.в. имеет закон распределения:

$$P(X = k) = P_k = \frac{(0,7)^k}{k!} e^{-0,7},$$

где $k = 0, 1, 2, \cdots$

Зададим уровень значимости β, например, $\beta = 0,20$. Последние два разряда, имеющие мало наблюдений, можно объединить. Таким образом, мы имеем четыре разряда, и число степеней свободы в этом случае равно $r = 4 - 2$ (так как на n_i наложены две связи $\sum n_i = n$ и $\overline{X} = 0,7$). Из таблицы 1 приложения для заданного $\beta = 0,20$ и полученного $r = 2$ находим, что критическая область для χ^2 представляет интервал $[3,219;\infty)$.

Для вычисления χ^2 по формуле (6) сначала определим

$$P_0 = \frac{(0,7)^0}{0!}e^{-0,7} = 0,5; P_1 = \frac{(0,7)^1}{1!}e^{-0,7} = 0,35; P_2 = \frac{(0,7)^2}{2!}e^{-0,7} = 0,12;$$

$$P_3 + P_4 + \cdots = 1 - 0,5 - 0,35 - 1,12 = 0,03.$$

Многоугольник распределения, соответствующий эти вероятностям построен на рис. 7 пунктирной линией. Вычисления χ^2 производим, фиксируя промежуточные результаты в таблице 4.5.

Таблица 4.5

Вычисление величины χ^2.

n_i	p_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$(n_i - np_i)^2/np_i$
54	0.50	50	4	16	0.32
27	0.35	35	-8	64	1.83
14	0.12	12	2	4	0.33
5	0.03	3	2	4	1.33

Величина χ^2 равна сумме результатов вычислений, стоящих в последнем столбце: $\chi^2 = 3,81$.

Вычисленное значение 3,81 входит в критическую область, поэтому делаем вывод: выдвинутая гипотеза противоречит опытным данным.

Пример 3. Требуется по уровню значимости $\beta = 0.05$ проверить гипотезу о том, что время работы элементов распределено по показательному закону, если в результате испытания 200 элементов получен статистический ряд, представленный в таблице 4.6.

Таблица 4.6

Результаты испытаний.

Интервалы времени (x_i, x_{i+1})	2,5 (0;5)	7,5 (5;10)	12,5 (10;15)	17,5 (15;20)	22,5 (20;25)	27.5 (25;30)
Частоты n_i (число элементов)	133	45	15	4	2	1

Найдем оценку параметра показательного распределения, используя формулы (3) и (4). В качестве "представителя"интервала примем середину интервала:

$$X = \frac{133 \cdot 2,5 + 45 \cdot 7,5 + 15 \cdot 12,5 + 4 \cdot 17,5 + 2 \cdot 22,5 + 1 \cdot 27,5}{200} = 5.$$

Таким образом, оценка параметра предполагаемого показательного распределения следующая: $= 1/\overline{X} = 0,2$.

Плотность распределения вероятности при таком предположении имеет вид:

$$f(X) = 0,2e^{-0,2X}; (X > 0).$$

Найдем вероятности попадания в каждый интервал.

$$P_i = P(x_i < X < x_{i+1}) = \int_{x_i}^{x_{i+1}} \lambda e^{-\lambda x} dx = e^{-\lambda x_i} - e^{-\lambda x_{i+1}}.$$

Так, для первого интервала

$$P_1 = P(0 < X < 5) = \int_0^5 0,2e^{-0,2x} dx = e^{-0,2 \cdot 0} - e^{-0,2 \cdot 5} = 0,6321.$$

Заполним таблицу 4.7, предварительно объединив малочисленные частоты $(4 + 2 + 1 = 7)$.

По таблице 1 приложения при уровне значимости $\beta = 0,05$ и числе степеней свободы $k = 4 - 2$ находим $\chi^2_\beta = 5,991$. Так как $\chi^2 < \chi^2_\beta$, то гипотеза опытным данным не противоречит.

Таблица 4.7

Вычисление величины χ^2.

n_i	p_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$(n_i - np_i)^2/np_i$
133	0.6321	126.42	6.58	43.2964	0.3425
45	0.2326	46.52	-1.52	2.3104	0.9497
15	0.0855	17.10	-2.10	4.4100	0.2579
7	0.0474	9.48	-2.48	6.1504	0.6488
					$\chi^2 = 1.2969 \approx 1.30$

5. ИНДИВИДУАЛЬНЫЕ ЗАДАНИЯ ДЛЯ ЛАБОРАТОРНОЙ РАБОТЕ

5.1 Теоретические вопросы.

1. Точечные оценки для $M(X), D(X)$. Понятие о доверительном интервале для $M(X)$ с надежностью γ. Зависимость длины интервала от γ.

2. Понятие построения полигона частот и гистограммы.

3. Выравнивание статистических распределений. Гипотеза о законе распределения. Можно ли по статистическим данным получить точный закон распределения?

4. Оценка расхождения между гипотетическим и статистическим распределением. Критерий χ^2. Схема применения критерия.

5.2 Задания к выполнению лабораторной работы.

1. Найти оценки для $M(X)$ и $D(X)$. Построить доверительный интервал для $M(X)$ с надежностью $\gamma_1 = 0,9; \gamma_2 = 0,95$. (Для выполнения этого пункта задания рекомендуется воспользоваться пособием [7].)

2. По заданному статистическому ряду построить гистограмму или полигон относительных частот для дискретной с.в.

3. По виду гистограммы (полигона частот) выдвинуть гипотезу о виде закона распределения с.в.

4. Построить в той же системе координат функцию плотности вероятности (многоугольник распределения) в соответствии с выдвинутой гипотезой.

5. С помощью критерия χ^2 проверить, насколько гипотеза согласуется с опытными данными. Уровень значимости выбрать самостоятельно.

5.3 Варианты заданий.

1. Партию из 100 механизмов исследовали на длительность безотказной работы (до первой поломки).

Длительность времени безотказной работы (мес)	от 0 до 1	от 1 до 2	от 2 до 3	от 3 до 4	от 4 до 5	от 5 до 6	от 6, до 7
Число механизмов	44	20	12	9	7	5	3

2. Из деталей, изготовленных станком, взяли для исследования 100 штук и измерили отклонения их размеров от стандарта.

Границы отклонений (м)	(-25;-15)	(-15;-5)	(-5;5)	(5;15)	(15;25)
Число деталей	8	20	44	25	3

3. Для изучения транспортного потока на шоссе проведен подсчет числа автомобилей за 100 минут.

Число автомобилей в минуту	0	1	2	3	4	5	6
Число минут с данным числом автомобилей	13	34	22	15	7	5	4

4. Произведено контрольное измерение отклонений размера от стандарта у 200 деталей.

Границы отклонений, мк	(-20,-10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	18	39	90	43	10

5. Даны результаты исследования нити на разрыв.

Крепость нити, гр	от 12 до 14	от 14 до 16	от 16 до 18	от 18 до 20	от 20 до 22	от 22 до 24	от 24 до 26	от 26 до 28
Число нитей	1	4	10	14	12	6	2	1

6. Даны результаты наблюдений за среднесуточной температурой воздуха.

Температура	(-40,-20)	(-20,0)	(0,20)	(20,40)
Число дней	14	65	137	54

7. Были изменены отклонения от стандарта у 200 деталей.

Границы отклонений, мк	(-30,-20)	(-20;10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	8	39	62	57	24	10

8. Для изучения транспортного потока через населенный пункт произвели подсчет проходящих автомобилей за 100 мин.

Число автомобилей, проходящих в мин, мк	0	1	2	3	4	5	6
Число минут с данным числом автомобилей.	14	33	21	14	9	7	2

9. Произведено измерение отклонений деталей от стандарта.

Границы отклонений, мк	(-30,-20)	(-20,-10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	1	12	37	32	14	4

10. Партию из 100 механизмов исследовали на длительность безотказной работы (до первой поломки).

Длительность времени безотказной работы (мес)	от 0 до 1	от 1 до 2	от 2 до 3	от 3 до 4	от 4 до 5	от 5 до 6	от 6 до 7	от 7, до 8
Число механизмов	44	20	12	8	7	3	3	3

11. Результаты измерения отклонения от стандарта размера деталей, изготовленных на некотором станке.

Границы отклонений, мк	(-30,-20)	(-20,-10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	7	26	76	67	24	3

12. Для исследования надежности некоторого прибора произвели подсчет числа выходов прибора из строя в течение 100 дней.

Число поломок в день	0	1	2	3	4	5 и более
Число дней с данным числом поломок	35	45	10	5	5	0

13. Проведено 100 измерений некоторой с.в..

Границы интервалов	(-6,-2)	(-2,2)	(2,6)	(6,10)	(10,14)
Число значений в интервалах	7	27	33	22	11

14. Были измерены отклонения размера деталей от стандарта.

Границы отклонений, мк	(-25,-15)	(-15,-5)	(-5,5)	(5,15)	(15,25)
Число деталей	5	20	40	26	9

15. Для исследования потока посетителей предприятия массового обслуживания измерили интервалы времени между последовательно приходящими посетителями.

Интервалы времени между двумя последовательно приходящими посетителями, мин	(0,1)	(1,2)	(2,3)	(3,4)	(4,5)
Число интервалов данной длительности	69	20	6	2	2

16. Произведены измерения отклонения размера деталей от стандарта.

Интервалы отклонений	(-15,-5)	(-5,5)	(5,15)	(15,25)	(25,35)
Число деталей	15	75	100	50	20

17. Результаты наблюдения некоторой с.в.

Границы интервалов	(0,2)	(2,4)	(4,6)	(6,8)	(8,10)	(10,12)
Число значений в интервалах	1	16	33	32	14	4

18. Для исследования надежности автоматических станков произвели подсчет числа поломок станка 100 рабочих дней.

Число разладок станков в день	0	1	2	3	4	5 и более
Число дней с данным числом разладок	35	43	13	5	4	0

19.Результаты измерения отклонения изготовленных на станке деталей от стандарта.

Границы интервалов, мк	(-25,-15)	(-15,-5)	(-5,5)	(5,15)	(15,25)
Число деталей	5	26	39	20	10

20. Для исследования потока заявок в некоторой системе массового обслуживания подсчитали число заявок за 100 минут.

Число заявок за 1 минуту	0	1	2	3	4	5	6	7 и более
Число минут с данным числом заявок	19	21	31	12	8	5	4	0

21. Приведены данные о выполнении плана рабочими.

Интервалы (проценты)	80-90	90-100	100-110	110-120	120-130	130-140
Число рабочих	9	15	34	20	14	8

22. Были измерены отклонения от стандарта у 100 деталей.

Границы отклонений, мк	(-10,-5)	(-5,0)	(0,5)	(5,10)	(10,15)	(20,30)
Число деталей	1	17	33	24	10	

23. Результаты наблюдения некоторой с.в..

Интервалы	(0,2)	(2,4)	(4,6)	(6,8)	(8,10)	(10,12)	(12,14)
Число значений в интервалах	28	35	46	29	13	6	3

24. Для исследования потока заявок на производимую продукцию на предприятии измеряли интервалы между 100 последовательно поступавшими заявками.

Длительность интервала между поступательно поступившими заявками, сут.	(0,1)	(1,2)	(2,3)	(3,4)	(4,5)
Число значений в интервалах	70	19	6	4	1

25. Партию из 100 механизмов исследовали на длительность безотказной работы (до первой поломки)

Длительность времени безотказной работы (тыс. ч.)	(0,1)	(1,2)	(2,3)	(3,4)	(4,5)
Число механизмов	71	17	6	3	3

26. Получены результаты 200 измерений с.в..

Границы интервалов	(-12,-8)	(-8,-4)	(-4,0)	(0,4)	(4,8)	(8,12)
Число значений в интервале	3	33	66	64	27	7

27. Измерено отклонение от стандарта у 100 деталей.

Границы отклонений, мк	(-30,-20)	(-20,-10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	2	10	37	34	13	4

28. Было проведено 200 измерений одной и той же с.в..

Границы интервалов	(-12,-8)	(-8,-4)	(-4,0)	(0,4)	(4,8)	(8,12)
Число значений в интервале	2	32	66	9	28	8

29. Результаты обследования однотипных деталей, изготовленных одним станком

Границы отклонений от стандарта, мк	(-30,-20)	(-20,-10)	(-10,0)	(0,10)	(10,20)	(20,30)
Число деталей	1	12	37	32	14	4

30. Проведено 100 измерений некоторой с.в..

Границы отклонений, мк	(0,2)	(2,4)	(4,6)	(6,8)	(8,10)	(10,12)
Число деталей	1	16	33	32	14	4

ПРИЛОЖЕНИЕ

Таблица 1

Значения χ^2_β в зависимости от числа степеней свободы r и уровня значимости β

$r \backslash \beta$	0.20	0.10	0.05	0.02	0.01
1	1.642	2.706	3.841	5.412	6.635
2	3.219	4.605	5.991	7.824	9.210
3	4.642	6.251	7.815	9.937	11.345
4	6.980	7.779	9.488	11.688	13.277
5	7.289	9.236	11.070	13.388	15.086
6	8.556	10.645	12.592	15.033	16.812
7	9.803	12.017	14.067	16.622	18.475
8	11.030	13.362	15.507	18.168	20.090
9	12.242	14.684	16.919	19.679	21.666
10	13.442	15.987	18.307	21.161	23.209

Таблица 2

Значения функции $\Phi(X) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	0.31	0.1217	0.62	0.2324	0.93	0.3238	1.24	0.3925
0.01	0.0040	0.32	0.1255	0.63	0.2357	0.94	0.3264	1.25	0.3944
0.02	0.0060	0.33	0.1293	0.64	0.2389	0.95	0.3289	1.26	0.3962
0.03	0.0120	0.34	0.1331	0.65	0.2422	0.96	0.3315	1.27	0.3980
0.04	0.0160	0.35	0.1368	0.66	0.2454	0.97	0.3340	1.28	0.3997
0.05	0.0199	0.36	0.1406	0.67	0.2486	0.98	0.3365	1.29	0.4015
0.06	0.0239	0.37	0.1443	0.68	0.2517	0.99	0.3389	1.30	0.4032
0.07	0.0279	0.38	0.1480	0.69	0.2549	1.00	0.3413	1.31	0.4049
0.08	0.0319	0.39	0.1517	0.70	0.2580	1.01	0.3438	1.32	0.4086
0.09	0.0359	0.40	0.1554	0.71	0.2611	1.02	0.3461	1.33	0.4082
0.10	0.0398	0.41	0.1591	0.72	0.2642	1.03	0.3486	1.34	0.4099
0.11	0.0438	0.42	0.1628	0.73	0.2673	1.04	0.3508	1.35	0.4115
0.12	0.0478	0.43	0.1664	0.74	0.2703	1.05	0.3531	1.36	0.4131
0.13	0.0517	0.44	0.1700	0.75	0.2734	1.06	0.3554	1.37	0.4147
0.14	0.0557	0.45	0.1736	0.76	0.2764	1.07	0.3577	1.38	0.4162
0.15	0.0696	0.46	0.1772	0.77	0.2794	1.08	0.3599	1.39	0.4177
0.16	0.0636	0.47	0.1808	0.78	0.2823	1.09	0.3621	1.40	0.4192
0.17	0.0675	0.48	0.1844	0.79	0.2862	1.10	0.3643	1.41	0.4207
0.18	0.0714	0.49	0.1879	0.80	0.2881	1.11	0.3665	1.42	0.4222
0.19	0.0753	0.50	0.1915	0.81	0.2910	1.12	0.3686	1.43	0.4236
0.20	0.0793	0.51	0.1960	0.82	0.2939	1.13	0.3708	1.44	0.4251
0.21	0.0832	0.52	0.1985	0.83	0.2967	1.14	0.3729	1.45	0.4265
0.22	0.0871	0.53	0.2019	0.84	0.2998	1.15	0.3749	1.46	0.4279
0.23	0.0910	0.54	0.2064	0.85	0.3023	1.16	0.3770	1.47	0.4292
0.24	0.0918	0.55	0.2088	0.86	0.3051	1.17	0.3790	1.48	0.4306
0.25	0.0987	0.56	0.2123	0.87	0.3078	1.18	0.3810	1.49	0.4319
0.26	0.1026	0.57	0.2157	0.88	0.3106	1.19	0.3830	1.50	0.4332
0.27	0.1064	0.58	0.2190	0.89	0.3133	1.20	0.3849	1.51	0.4345
0.28	0.1103	0.59	0.2224	0.90	0.3159	1.21	0.3869	1.52	0.4357
0.29	0.1141	0.60	0.2257	0.91	0.3186	1.22	0.3883	1.53	0.4370
0.30	0.1179	0.61	0.2291	0.92	0.3212	1.23	0.3907	1.54	0.4382

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1.55	0.4394	1.76	0.4603	1.97	0.4756	2.36	0.4909	2.78	0.4973
1.56	0.4406	1.77	0.4610	1.98	0.4761	2.38	0.4913	2.80	0.4974
1.57	0.4418	1.78	0.4625	1.99	0.4767	2.40	0.4918	2.82	0.4976
1.58	0.4429	1.79	0.4633	2.00	0.4772	2.42	0.4922	2.84	0.4977
1.59	0.4441	1.80	0.4641	2.02	0.4783	2.44	0.4927	2.86	0.4979
1.60	0.4452	1.81	0.4649	2.04	0.4793	2.46	0.4931	2.88	0.4980
1.61	0.4463	1.82	0.4656	2.06	0.4803	2.48	0.4934	2.90	0.4981
1.62	0.4474	1.83	0.4664	2.08	0.4812	2.50	0.4938	2.92	0.4982
1.63	0.4484	1.84	0.4671	2.10	0.4821	2.52	0.4941	2.94	0.4984
1.64	0.4495	1.85	0.4678	2.12	0.4830	2.54	0.4945	2.96	0.4985
1.65	0.4505	1.86	0.4686	2.14	0.4838	2.56	0.4948	2.98	0.4986
1.66	0.4515	1.87	0.4693	2.16	0.4840	2.58	0.4951	3.00	0.49865
1.67	0.4525	1.88	0.4699	2.18	0.4854	2.60	0.4953	3.20	0.49931
1.68	0.4535	1.89	0.4706	2.20	0.4861	2.62	0.4956	3.40	0.49965
1.69	0.4540	1.90	0.4703	2.22	0.4868	2.64	0.4959	3.60	0.499841
1.70	0.4554	1.91	0.4719	2.24	0.4875	2.66	0.4961	3.80	0.499928
1.71	0.4564	1.92	0.4726	2.26	0.4881	2.68	0.4963	4.00	0.499968
1.72	0.4573	1.93	0.4732	2.28	0.4887	2.70	0.4965	4.50	0.499997
1.73	0.4582	1.94	0.4738	2.30	0.4893	2.72	0.4967	5.00	0.499999
1.74	0.4591	1.95	0.4744	2.32	0.4898	2.74	0.4969		
1.75	0.4599	1.96	0.4750	2.34	0.4904	2.76	0.4971		

СПИСОК ЛИТЕРАТУРЫ

1. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения [текст] /Е.С.Вентцель, Л.А.Овчаров. - М.: Наука, 1988.

2. Ван дер Варден. Математическая статистика[текст]/Ван дер Варден. - М.: ИЛ, 1960.

3. Вентцель, Е.С. Теория вероятностей[текст]/Е.С. Вентцель - М.: Высшая школа, 2003.

4. Коваленко, И.Н. Теория вероятностей и математическая статистика[текст] /И.Н. Коваленко, А.А. Филиппова - М.: Высшая школа, 1982.

5. Смирнов, Н.В. Курс теории вероятностей и математической статистики [текст]/Н.В. Смирнов, И.В. Дунин-Барковский. - М.: Физматгиз, 1969.

6. Гмурман, В.Е. Теория вероятностей и математическая статистика[текст] /В.Е. Гмурман - М.: Высшая школа, 2001.

7. Кононенко, Э.Д. Интервальные оценки. Методическая разработка [текст] /Э.Д. Кононенко, В.И. Жукова.-ДВГАПС, 1994.