ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ МЕДИЦИНСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.И. ПИРОГОВА» МИНИСТЕРСТВА ЗДРАВООХРАНЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

(ФГБОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России)

РУКОВОДСТВО К ВЫПОЛНЕНИЮ ЗАДАНИЙ ПО МАТЕМАТИЧЕСКОЙ СТАТИСТИКЕ

Москва, 2017

Занина Е.Н., Парпара А.А., Бойко А.Я.
Руководство к выполнению заданий по математической статистике. – 2017, 69 с.

Задания по математической статистике составлены для статистической обработки экспериментальных данных с целью выяснения закона распределения, числовых характеристик и статических связей исследуемых признаков. Они включают методологические разработки по четырем темам: вариационный ряд и его числовые характеристики, точечные и интервальные оценки, проверка статистических гипотез и корреляционный анализ.

Задания рассчитаны на медиков и биологов, имеющих подготовку по курсу математической статистики. Они с успехом могут быть использованы практическими врачами при статистической обработке экспериментального материала.

© Занина Е.Н., Парпара А.А., Бойко А.Я., 2017.
© Федеральное государственное бюджетное образовательное учреждение высшего образования «Российский национальный исследовательский медицинский университет имени Н.И. Пирогова» Министерства здравоохранения Российской Федерации (ФГБОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России), 2017.

Задание №1

Построение интервального ряда и вычисление основных статистических характеристик

1.1. Предварительные сведения

1.1.1. Понятие таблицы данных

Наиболее распространённая форма, в которой представляются исходные данные для статистических вычислений — это *таблица данных.* При постановке эксперимента форма регистрации значений выбранного признака должна быть наиболее простой и удобной для обработки экспериментальных данных. Выбор формы, в свою очередь, зависит от вида признака (количественного или качественного), числа регистрируемых признаков (переменных) и метода обработки полученных данных. В этом случае выбирается несколько величин признаков *(переменных*), каждая из которых измеряется одинаковое число раз, причём *измерения* разных переменных, как правило, являются *связанными,* то есть относятся к одному и тому же объекту и моменту времени. В статистике набор измеренных значений переменной часто называют *выборкой*, а их количество — *объёмом выборки* по числу измеренных объектов.

Например, пусть в исследовании участвуют 50 пациентов. Для каждого из них фиксируются четыре переменные: схема терапии, самочувствие (плохое, удовлетворительное, хорошее, отличное), число сеансов терапии и температура тела.

В таблице данных переменные откладываются в столбцах, а измеренные значения — в строках. Так, в результате вышеописанного исследования получится таблица из 4 столбцов и 50 строк, причём значения в каждой строке будут связаны, потому что относятся к одному и тому же пациенту.

Нужно заметить, что в нашем примере все переменные имеют различные *типы*. Это важно понимать, поскольку от типа данных зависит способ статистической обработки. Типы статистических данных различаются в зависимости от того, как они могут быть пронумерованы. Так, схемы терапии не имеют никакого естественного способа подсчета: их можно нумеровать в любом порядке любыми числами. Такой тип переменной называют *неупорядоченным фактором.*

Уровень самочувствия отличается тем, что имеет естественный *порядок:* от «плохого» к «отличному». Однако какая-либо определённая *мера* у него отсутствует: нельзя сказать, во сколько раз «хорошее» самочувствие лучше «плохого». Поэтому его уровни можно пронумеровать любыми четырьмя различными числами, нужно только чтобы «плохому» соответствовало наименьшее, «удовлетворительному» - следующее по величине, а «отличному» — наибольшее. Такой тип называют *упорядоченным фактором.*

Число же сеансов терапии имеет свою меру: один сеанс, поэтому никакого произвола в подсчёте не допускает. Но эта величина может принимать лишь ряд отдельных значений, между которыми промежуточные величины невозможны, например, может быть проведено 5 или 6 сеансов, а 5,5 — уже нет. Такого типа переменную называют *дискретной.*

Наконец, температура тела имеет меру (1^{o}C) и может непрерывно изменяться. Такую переменную называют *непрерывной.* Впрочем, нужно заметить, что в статистике различие между дискретными и непрерывными величинами не столь принципиально, как в теории вероятностей. Дело в том, что статистика работает с данными измерений, а любой измерительный прибор позволяет получать результат лишь с точностью до цены деления шкалы. Таким образом, при измерении непрерывная величина неизбежно становится дискретной**.**

1.1.2. Правила приближённых вычислений

Итак, если целочисленная переменная, как правило, известна абсолютно точно, то вещественная — всегда с какой-то степенью точности. Вещественную переменную характеризуют числом десятичных знаков или числом значащих цифр.

Десятичные знаки (д.з.) — точно известные цифры после запятой. Если последние цифры являются нулями, их записывают для обозначения точности, например, 2,3 имеет 1 д.з., а 2,300 — 3 д.з.

Значащие цифры (з.ц.) — это все цифры числа от первой ненулевой до последней точно известной. Например, 210,20 содержит 5 з.ц., а 0,00193 — 3 з.ц.

Если возникает необходимость *округлить* число до n з.ц., поступают следующим образом. Если (n + 1)-я цифра меньше 5, лишние цифры просто отбрасывают, если больше или равна 5 — *n-ю* цифру при этом увеличивают на 1, кроме случая, когда n-я цифра чётная, (n + 1)-я равна 5, а (n + 2)-я — 0. Так, при округлении до 3 з.ц. числа 21,335, получаем 21,3, для 1,34708 - 1,35, для 344,512 - 345, для 7,45501 - 7,46, но для 33,450 - 33,4.

При вычислениях часто получаются неверные д.з., которые нужно округлять. При сложении и вычитании нужно оставлять столько д.з., сколько содержится в наименее точном слагаемом. Например, 2,78 + 3,9313 ≈ 6,71. При умножении и делении оставляют столько з.ц., сколько содержится в наименее точном множителе. Например, 0,0321 • 8,7678 ≈ 0,281. Обратите внимание, что при сложении и вычитании имеет значение число д.з., которое характеризует абсолютную погрешность числа, а при умножении и делении — число з.ц., характеризующее относительную погрешность. При возведении в степень и извлечении корня следует оставлять столько з.ц., сколько имеет основание.

1.1.3. Данные для задания.

Для выполнения задания требуется таблица данных, содержащая две вещественные выборки объёмом n = 50, полученные на пятидесяти объектах.

Будем называть их А и Б. Каждая выборка содержит значения признаков (переменных), измеренных у объектов определенного типа, например, площадь (AREA) и оптическая плотность ядра лимфоцитов (ODR). Соответствующие значения выборок А и В измерены у одного и того же объекта, поэтому их можно рассматривать как связанные переменные.

При выполнении заданий с первого по четвертый используются оба признака (к примеру, AREA, ODR), представленные выборками А и В.

Каждому из студентов в группе выдаются две выборки А и В, представляющие измерения двух признаков (например, AREA, ODR) для одного из двух разных типов объектов (например лимфоцитов или сегментоядерных нейтрофилов).

При выполнении первых двух заданий для построения вариационного ряда и вычисления его основных числовых характеристик, а также для нахождения точечных и интервальных оценок рассматриваемых признаков одного из типов объектов, кроме исходных данных, дополнительных сведений не требуется.

Однако в третьем задании, посвященном проверке статистических гипотез, требуется сравнить одинаковые признаки (например, AREA, ODR), относящиеся к разным типам объектов (например, лимфоцитам и нейтрофилам).

При этом требуется знать оценки математического ожидания (МО)$-\bar{x}$ и дисперсии (σ^2)$-s^2$ распределения вариационного ряда исследуемых признаков, а также значения ошибок репрезентативности (представительности выборки) $\sigma_{\bar{x}}$ и σ_{s^2} представленных выборками А и В для разных типов объектов (например, лимфоцитов и нейтрофилов).

Поэтому студенты, работающие в парах, изучают одни и те же признаки для разных типов. В 3-ем задании они обмениваются полученными точечными оценками для МО - $\bar{x}$, для дисперсии $-$ $s_{\bar{x}}^2$, и ошибками репрезентативности $-$ σx, $\sigma s2$ для одинаковых признаков разных типов объектов.

Кроме того, для проверки статистических гипотез в задании 3 требуются эталонные значения математического ожидания m_0 и дисперсии σ_0^2 по одному для каждого признака, а в задании 4 — эталонное значение коэффициента корреляции r_0 между этими признаками (переменными), которые содержатся в задании или выдаются преподавателем.

1.2. Построение интервального ряда, заполнение статистической таблицы

Первый шаг при статистической обработке выборки — группировка данных. Она нужна для следующих целей.

1. Для построения гистограммы, полигона и кумуляты, нахождения медианы (M_e) и моды (M_o).
2. Для проверки гипотезы о законе распределения с помощью критерия Пирсона χ^2. (см. 3.2.1.)

3. Для составления корреляционной решетки при вычислении коэффициента корреляции.
4. Для облегчения работы с большими выборками, если вычисления производятся без помощи ЭВМ.

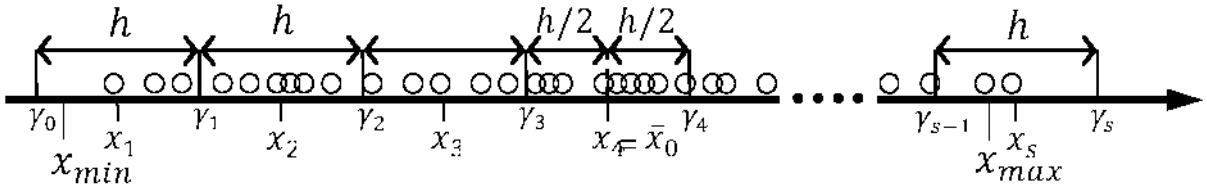


Рисунок 1. Построение интервального ряда. Кружки – выборочные данные, $\boldsymbol{\gamma_0 \div \gamma_s}$ *- границы интервалов,* $\boldsymbol{h}$ *- шаг разбиения,* $\boldsymbol{x_{min}}$ *и* $\mathbf{x_{max}}$ *- минимальное и максимальное выборочные значения,* $\boldsymbol{\bar{x}_0}$ *- начальная точка отсчета,* $\boldsymbol{x_k}$ *– середины интервалов, где k=*$\overline{\mathbf{1}, \boldsymbol{s}}$ *.*

Итак, разделим числовой промежуток, в пределах которого изменяются данные, на равные интервалы-классы (рис. 1) и подсчитаем количество выборочных точек, попавших в каждый интервал. Таким образом, получим возрастающий вариационный интервальный ряд — таблицу из двух столбцов, в первом из которых записываются границы интервалов, а во втором — *выборочная частота*, то есть число выборочных данных, которые попали в интервал.

1.2.1. Работаем с исходной выборкой объёма n.

1.2.2**.** Среди всех выборочных значений найдите минимальное значение $-x_{min}$.

1.2.3. Среди всех выборочных значений найдите максимальное значение $-\mathrm{x_{max}}$.

1.2.4**.** Вычислите размах $\mathrm{R} = \mathrm{x_{max}} - x_{min}$.

1.2.5. Определите число интервалов $\mathrm{s} = 1 + 3{,}3\, lg\, n$.

1.2.6. Определите длину интервала (шаг разбиения) $\mathrm{h} = \mathrm{R/s}$. Значение h округлите до числа десятичных знаков после запятой, содержащихся в выборочных значениях x_k изучаемого признака (см. 1.1.2) так, чтобы на конце была четная цифра.

1.2.7. Выбор начальной точки отсчёта $\bar{x}_0$.

1.2.7.1. Всем значениям выборок А и В приписываем номера от 00 ÷ n-1: для выборки А – значения x_k; для выборки В – значения y_k, где к = 00 ÷ $n - 1$.

1.2.7.2. С помощью *таблицы I случайных чисел приложений* выберите 10 двузначных чисел (номеров) (можно с помощью генератора случайных чисел). Например, для выборки А появились номера 01; 18; 34; …49. Всего 10 номеров.

1.2.7.3. Из исходной выборки выпишите 10 выборочных значений с этими номерами, то есть $x_{01}; x_{18}; x_{34}; \ldots x_{49}$ (продолжение примера). Всего 10 значений x_k.

1.2.7.4. Вычислите

$$\bar{x}_0 = \frac{1}{10}\sum_{k=1}^{10} x_k$$

и округлите до стольких знаков после запятой, сколько имеется в h.

1.2.8. Нахождение границ интервалов γ_k.

1.2.8.1. От $\bar{x}_0$ отложите вправо с (+) и влево с (-) по значению $h/2$ (см. рис. 1).
1.2.8.2. От полученных граничных значений отложите по целому значению h.

1.2.8.3. От новых граничных значений ещё раз отложите h и так далее, пока не будут покрыты x_{min} и x_{max}.

1.2.8.4. Полученные значения γ_k – границ интервалов внесите в таблицу № 1 (статистическая таблица).

1.2.8.5. Рассчитайте середины интервалов: $x_1 = \gamma_1 - h/2$ или $x_1 = \gamma_0 + h/2$; $x_k = x_{k-1} + h$, где k: 2, 3, 4, …, s и внесите в таблицу №1.

1.2.8.6. При составлении статистической таблицы №1 к левой границе каждого следующего интервала прибавляется единица последнего знака (ε) правой границы предыдущего интервала. Например, γ_k – правая граница $k-$ го интервала, ε - единица его последнего знака, $\gamma_k + \varepsilon$- левая граница $(k+1)-$ го интервала, а γ_{k+1}- правая его граница. Если $\gamma_k = 2{,}35$, то $\varepsilon = 0{,}01$, откуда левая граница $(k+1)$ интервала $\gamma_k + \varepsilon = 2{,}36$. Таким образом, каждый интервал имеет открытую левую и закрытую правую границы интервала, то есть $(\gamma_k, \gamma_{k+1}]$ (при этом k меняется от 0 до $s-1$, где s– число интервалов).

1.2.9. Подсчет частот.

1.2.9.1. Просмотрите по порядку все выборочные значения от x_{00} до x_{n-1}. Для каждого из них найдите интервал, в который оно попадает, и в соответствующей строке столбца «Кодировка» таблицы 1 поставьте вертикальную черту. Когда накапливаются 4 вертикальные черты, перечеркивайте их горизонтальной, чтобы считать пятерками. Кодировка «пятками»: 1-I; 2-II; 3-III; 4-IIII; 5-~~IIII~~; 6-~~IIII~~ I; 7-~~IIII~~ II ⋯ 12- ~~IIII IIII~~ II …

1.2.9.2. В каждом интервале сосчитайте черты в столбце «Кодировка» и запишите их количество в столбец «Частота» - m_k.

1.2.9.3. Сосчитайте накопленные частоты: $M_1 = m_1$, $M_k = M_{k-1} + m_k$, где k: 2, 3, …, s.

Таблица №1. Статистическая таблица.

№ k	Границы интервала $\gamma_{k-1}+\varepsilon, \gamma_k$	Кодировка 1 2 3 4 5 6 7 ...	Частота m_k	Накопленная частота M_k	Середина интервала x_k
1	$x \le \gamma_1$		m_1	$\rightarrow M_1$	$x_1 = \gamma_1 - h/2$
2	$\gamma_1+\varepsilon, \gamma_2$		$m_2 \swarrow$	$\rightarrow M_2$	$x_2 = x_1 + h$
3	$\gamma_2+\varepsilon, \gamma_3$		$m_3 \swarrow$	$\rightarrow M_3$	$x_3 = x_2 + h$
…	…	…	…	…	…
$s-1$	$\gamma_{s-2}+\varepsilon, \gamma_{s-1}$		$m_{s-1} \swarrow$	$\rightarrow M_{s-1}$	$x_{s-1} = x_{s-2} + h$
s	$\gamma_{s-1}+\varepsilon, \gamma_s$		$m_s \swarrow$	$\rightarrow M_s = n$	$x_s = x_{s-1} + h$

1.3. Графическое представление статистической таблицы**.**

С помощью статистической таблицы постройте гистограмму, полигон и кумуляту. Гистограмма – это график зависимости выборочных частот от значения переменной. Внутри интервала частота считается постоянной (рис. 2.), Полигон отличается от гистограммы тем, что выборочные частоты откладываются в серединах интервалов и соединяются отрезками прямых (рис. 2). Кумулята отличается от гистограммы тем, что на графике откладываются накопленные частоты, то есть сумма выборочных частот во всех интервалах левее данного включительно (рис. 3).

Гистограмма и полигон дают оценку плотности распределения генеральной совокупности (ГС), а кумулята – её функции распределения. Гистограмму обязательно нужно строить перед началом анализа данных, чтобы получить наглядное представление о характере их распределения. На графике гистограммы можно приблизительно оценить основные характеристики выборки (среднее, медиану, моду, стандартное отклонение), оценить симметричность распределения и наличие выбросов.

Выбросами в статистике называют нетипичные наблюдения, которые расположены далеко от основной массы данных. Они возникают обычно в качестве ошибки при проведении эксперимента. Наличие выбросов может сильно исказить результаты статистического анализа, особенно корреляционного. На гистограмме они выглядят как столбики с одной – двумя точками, отделенные от остального графика несколькими пустыми интервалами (рис. 4). Часто выбросы появляются из-за несоблюдения условий эксперимента или случайного включения в выборку объектов, нетипичных для исследуемой ГС (например, в выборку жителей Архангельска попал один уроженец Анголы). В этом случае, возможно, следует исключить такое измерение из статистического анализа.

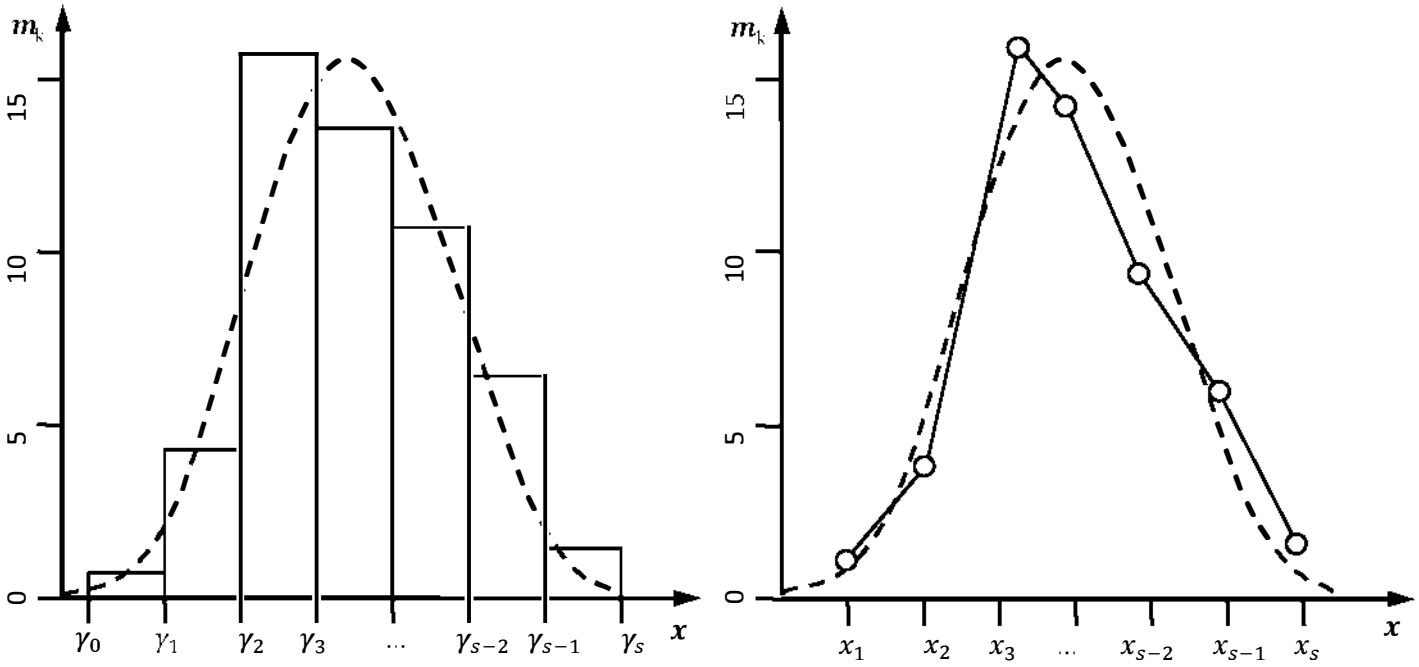


Рисунок 2. Гистограмма (слева) и полигон (справа). Пунктирная линия – плотность нормального распределения с параметрами равными выборочным значениям среднего и дисперсии. По оси $\boldsymbol{x}$ *– значение переменной, по оси* $\boldsymbol{y}$ *– частота* $\boldsymbol{m_k}$*.* $\boldsymbol{\gamma_0 \div \gamma_s}$ *– границы интервалов,* $\boldsymbol{x_1 \div x_s}$ *- середины интервалов.*

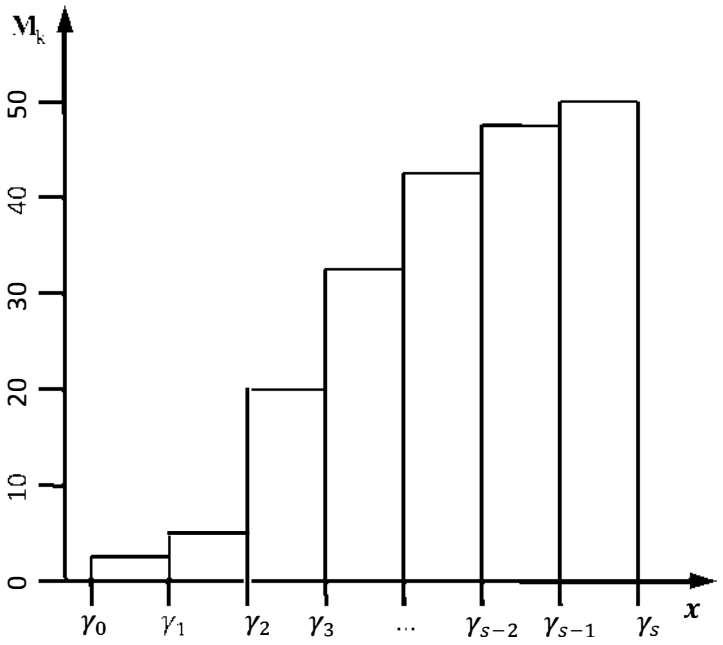


Рисунок 3. Кумулята. По оси x - значение переменной, по оси y - накопленные частоты M_k. $\gamma_0 \div \gamma_s$- границы интервалов.

Примечание: При построении гистограммы, полигона и кумуляты вместо буквенных значений γ_k и x_k, m_k, M_k записываются и отмечаются на графиках расчетные значения из табл.№1.

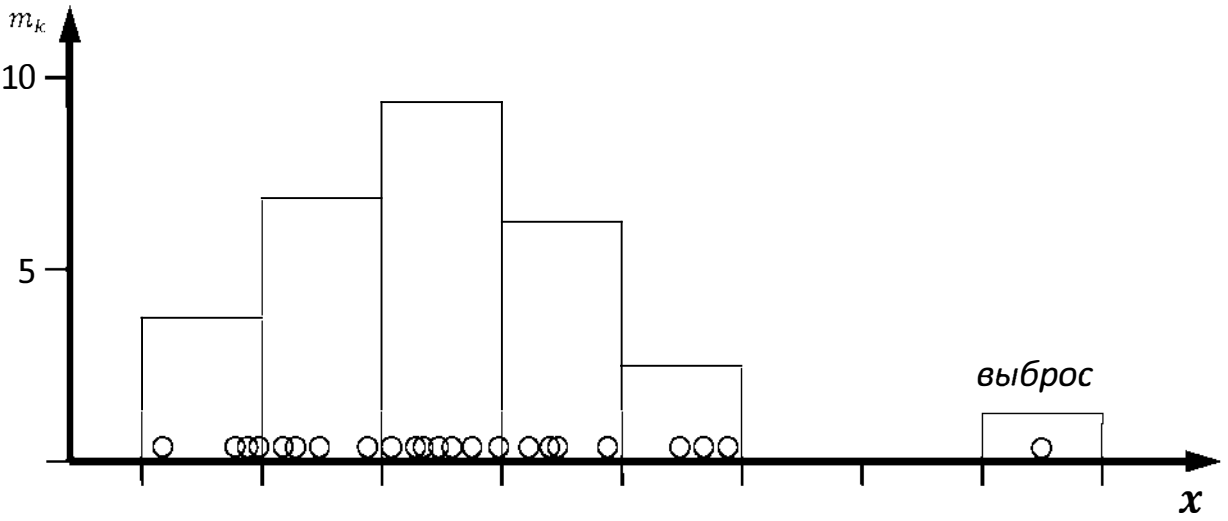


Рисунок 4. Выброс - наблюдение, далекое от основной массы данных.

1.3.1 Нахождение медианы (M_e).

Медиана – это значение переменной, которое делит распределение пополам, то есть вероятность того, что половина значений случайной величины окажется ≥ медианы, равна вероятности того, что половина окажется < нее и равна ½.

Для оценки значения медианы по гистограмме в таблице №1 выберете первую строку, в которой накопленная частота $M_k \geq n/2$.

1.3.2. Рассчитайте оценку медианы по формуле

$$\widehat{M}_{\text{e}} = \gamma_{k-1} + h\left(\frac{n/2 - M_{k-1}}{m_k}\right)$$

1.3.3. Нахождение моды (M_o).

Модой называется наиболее вероятное значение случайной величины, то есть то значение, при котором достигается максимальная вероятность или плотность вероятности.

Для оценки значения моды по гистограмме в таблице №1 выберите строку с максимальной частотой m_k.

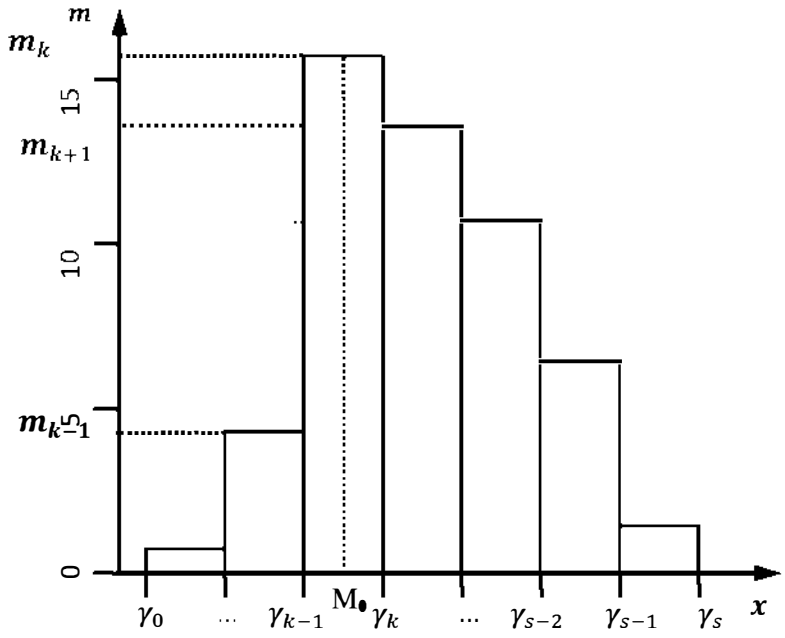
1.3.4. Вычислите оценку моды по формуле

$$\widehat{M}_{\text{o}} = \gamma_{k-1} + h\left(\frac{m_k - m_{k-1}}{2m_k - m_{k-1} - m_{k+1}}\right)$$

1.3.5. Если $N \geq 2$ соседних интервалов от k_1–го до k_2– го имеют одинаковую максимальную частоту, возьмите среднее арифметическое середин этих интервалов

$$\frac{1}{N}\sum_{k=k_1}^{k_2} x_k$$

1.3.6. Если одинаковую максимальную частоту имеют два не соседних интервала, нужно вычислить соответствующее количество мод. Возможно, имеется бимодальное распределение.



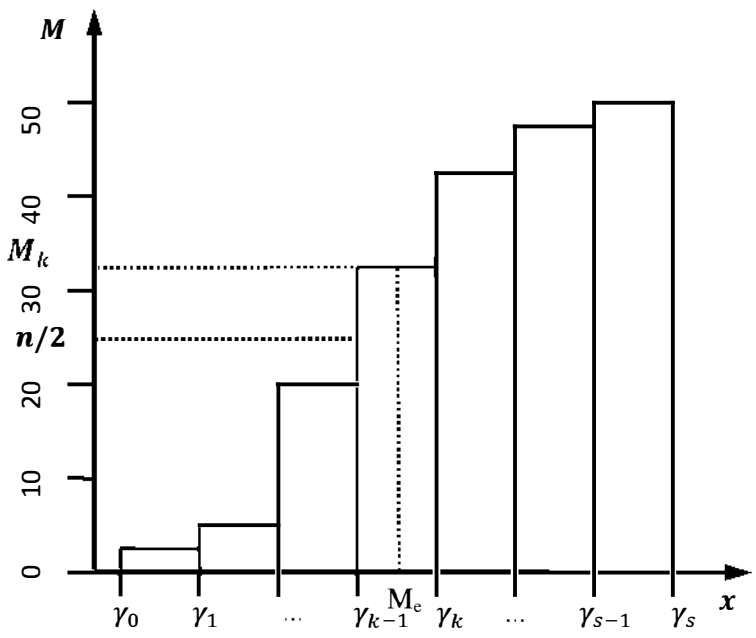


Рисунок 5. Мода и медиана. Слева. Нахождение моды: K –номер модального интервала; γ_{k-1} и γ_k – его левая и правая границы; m_{k-1}, m_k, m_{k+1} – соответствующие частоты предмодального (k-1) , модального (k) и постмодального (k+1) интервалов , M_0 *– мода . Справа. Нахождение медианы: K –номер интервала, содержащий медиану; γ_{k-1} и γ_k – его левая и правая границы; M_k – его накопленная частота, M_e – медиана. Показано, что $M_k \geq n/2$, а $M_{k-1} < n/2$.*

Таблица № 2. Метод произведений

x_k	m_k	y_k	$m_k y_k$	$m_k y_k^2$	$m_k y_k^3$	$m_k y_k^4$	y_k+1	$m_k(y_k+1)^4$
x_1	m_1	y_1	$m_1 y_1$	$m_1 y_1^2$	$m_1 y_1^3$	$m_1 y_1^4$	y_1+1	$m_1(y_1+1)^4$
x_2	m_2	y_2	$m_2 y_2$	$m_2 y_2^2$	$m_2 y_2^3$	$m_2 y_2^4$	y_2+1	$m_2(y_2+1)^4$
…	…	…	…	…	…	…	…	…
$a= x_i$	$max\ m_i$	●	●	●	●	●	1	m_i
…	…	…	…	…	…	…	…	…
x_s	m_s	y_s	$m_s y_s$	$m_s y_s^2$	$m_s y_s^3$	$m_s y_s^4$	$y_s + 1$	$m_s(y_s + 1)^4$
	Σ●		Σ_1	Σ_2	Σ_3	Σ_4		$\Sigma_{контр.}$

1.4. Метод произведений для вычисления характеристик вариационного ряда.

Характеристиками распределения вариационного ряда изучаемого признака являются оценки неизвестных параметров генеральной совокупности, из которой была сделана выборка. Точечными оценками МО, M_o, M_e являются соответственно среднее значение $\overline{x}$, выборочные значения моды $\widehat{M_o}$ и медианы $\widehat{M_e}$; оценками дисперсии σ^2 являются смещенная $\hat{\sigma}^2$ и несмещенная оценка s^2; оценками начальных и центральных моментов α_ℓ, μ_ℓ являются соответственно выборочные моменты $\hat{\alpha}_\ell$ и $\hat{\mu}_\ell$, где $\ell = 1,2,3,4$. Выборочные характеристики $\widehat{A}_s$ и $\widehat{E}_x$ являются точечными оценками соответственно коэффициентов асимметрии A_s и эксцесса E_x в генеральной совокупности.

Для расчета оценок МО, дисперсии (σ^2), асимметрии (A_S) и эксцесса (E_X) понадобятся выборочные начальные и центральные моменты с первого по четвертый порядок, то есть $\hat{\alpha}_1 \div \hat{\alpha}_4$ и $\hat{\mu}_2 \div \hat{\mu}_4$. Если объем выборки велик, их вычисление без помощи ЭВМ становится слишком трудоемким. Поэтому можно применить приближенный метод произведения. Уменьшение объема вычислений достигается путем группировки данных и перехода к целым числам с помощью изменения масштаба и переноса начала отсчета (метод «ложного нуля»).

1.4.1. Работаем со статистической таблицей №1.

Статистическая таблица №1 используется для составления таблицы №2 (метод произведения) и таблицы №3 (метод сумм), которые позволяют с помощью замены переменной простейшим способом рассчитать выборочные начальные моменты $\hat{\alpha}_\ell$, где $\ell = 1,2,3,4$.

1.4.2. Выберите в качестве «ложного нуля» $a=x_i$ середину интервала с максимальной частотой m_i, где i=1÷ s. Если таких несколько, возьмите тот, который ближе к середине интервального ряда.

1.4.3. Вычислите масштабированные значения $y_k=(x_k-a)/h$, где k=1÷ s. Если интервальный ряд построен правильно, значения y_k должны получиться целыми, и соседние y_k должны отличатся на 1.

1.4.4. Вычислите столбцы $m_k y_k \div m_k y_k^4$ слева направо, каждый раз числа, стоящие в предыдущем столбце, умножая на y_k, где k=1÷ s.

1.4.5. Сформируйте столбцы контрольных значений y_k+1 и $(y_k+1)^4 m_k$

1.4.6. Найдите суммы по столбцам $\Sigma_0 \div \Sigma_4$ и $\Sigma_{контр}$.

1.4.7. Проверьте правильность вычислений. Они произведены правильно, если выполняется равенство $\Sigma_{контр} = \Sigma_{●} + 4\,\Sigma_1 + 6\,\Sigma_2 + 4\Sigma_3 + \Sigma_4$.

1.4.8. Рассчитайте начальные моменты для масштабированной переменной *y:*

$$\hat{\alpha}_{l,y} = \frac{1}{n}\sum_{k=1}^{s} m_k\, y_k^l, \text{ где } l = 1,2,3,4.$$

1.4.9. Рассчитайте центральные моменты для *y* по формулам.

$$\hat{\mu}_{2,y} = \hat{\sigma}_y^2 = \hat{\alpha}_{2,y} - \hat{\alpha}_{1,y}^2;\ \ \hat{\sigma}_y \ \ \sqrt[2]{\hat{\sigma}_y^2}$$

$$\hat{\mu}_{3,y} = \hat{\alpha}_{3,y} - 3\hat{\alpha}_{2,y}\hat{\alpha}_{1,y} + 2\hat{\alpha}_{1,y}^3$$

$$\hat{\mu}_{4,y} = \hat{\alpha}_{4,y} - 4\hat{\alpha}_{3,y}\hat{\alpha}_{1,y} + 6\hat{\alpha}_{2,y}\hat{\alpha}_{1,y}^2 - 3\hat{\alpha}_{1,y}^4$$

1.4.10. Вычислите числовые характеристики исходной выборки по формулам.

$$\bar{x} = h\hat{\alpha}_{1,y} + a; \qquad \hat{A}_s = \hat{\mu}_{3,y}/\hat{\sigma}_y^3$$

$$\hat{\sigma}_x^2 = h^2\hat{\mu}_{2,y} \approx S^2; \qquad \hat{E}_x = \frac{\hat{\mu}_{4,y}}{\hat{\sigma}_y^4} - 3$$

1.5. Метод сумм

Метод сумм является контрольным для метода произведений при вычислении начальных моментов $\hat{\alpha}_{ly}$ (l = 1, 2, 3, 4) для нахождения числовых характеристик распределения частот вариационного ряда без применения сложной вычислительной техники. Может применяться самостоятельно без метода произведений.

1.5.1**.** Работаем с таблицей №2.

1.5.2. Перепишем три первых столбца в таблицу № 3.

1.5.3. Столбцы (1); (2); (3); (4); (5) – столбцы накопленных частот m_k.

1.5.4. Строка, в которой $y_k = 0$, называется нулевой →“0” – й строкой, относительно которой происходит накопление частот m_k

сверху - вниз выше “0” – й строки

снизу вверх ниже “0” – й строки.

1.5.6. В столбце (1) накопление частот m_k из 3-го столбца таблицы №3 происходит вплоть до “0” – й строки по правилу:

$$m_1=m_1$$

$$m_k=m_{k-1}+m_k,\ k\text{: } 2, 3, \ldots \text{ до “0” – ой строки и}$$

$$m_s=m_s$$

$$m_{s-1}=m_s+m_{s-1},\ s\text{: } 8, 7, \ldots \text{ снизу вверх до “0” – ой строки}$$

(см. табл. № 3)

1.5.7. В каждом следующем столбце (2)÷(5) накопление происходит на строчку меньше выше и ниже “0” – й строки пока в строчках не останется частот.

1.5.8. В каждом из накопленных столбцов подсчитываем сумму частот под и над нулевой строкой.

1.5.9. Рассчитываем

$$d_l = \sum_{\text{под}} - \sum_{\text{над}}$$

$$S_l = \sum_{\text{под}} + \sum_{\text{над}}, \text{ где } l\text{: } 1,2,3,4(5).$$

1.5.9.1. Значения d_l и s_l подставляем в формулы для нахождения начальных моментов $\hat{\alpha}_{l,y}$, где l: 1,2,3,4.

$$\hat{\alpha}_{1,y} = \frac{d_1}{n}; \qquad \hat{\alpha}_{3,y} = \frac{d_1 + 6d_2 + 6d_3}{n};$$

$$\hat{\alpha}_{2,y} = \frac{s_1 + 2s_2}{n}; \qquad \hat{\alpha}_{4,y} = \frac{s_1 + 14s_2 + 36s_3 + 24s_4}{n};$$

1.5.9.2. Сравниваем полученные значения $\hat{\alpha}_{l,y}$ с соответствующими значениями, полученными методом произведений (***совпадение абсолютное!***).

Таблица № 3.

x_k	y_k	m_k	(1)	(2)	(3)	(4)	(5)
x_1	-3	m_1	m_1	m_1	m_1	-	-
x_2	-2	m_2	$m_1 + m_2$	$2m_1 + m_2$	-	-	-
x_3	-1	m_3	$m_1 + m_2 + m_3$	-	-	-	-
x_4	0	m_4	Нулевая	Строка	-	-	-
x_5	1	m_5	$m_5 + m_6 + m_7 + m_8$	-	-	-	-
x_6	2	m_6	$m_6 + m_7 + m_8$	$m_6 + 2m_7 + 3m_8$	-	-	-
x_7	3	m_7	$m_7 + m_8$	$m_7 + 2m_8$	$m_7 + 3m_8$	-	-
x_8	4	m_8	m_8	m_8	m_8	m_8	-
$\sum_{\text{под}}$			…	…	$m_7 + 4m_8$	m_8	
$\sum_{\text{над}}$			…	…	m_1	0	
$S_l = \sum_{\text{под}} + \sum_{\text{над}}$			…	…	$m_7 + 4m_8 + m_1$	m_8	
$d_l = \sum_{\text{под}} - \sum_{\text{над}}$			…	…	$m_7 + 4m_8 - m_1$	m_8	

1.6. Результаты выполненного задания представить в следующем виде:

1) Рисунки 1,2,3 (отметить характеристики распределения $\bar{x}$, $\widehat{M}_0$, $\widehat{M}_e$, $\hat{\sigma}_x$, $\hat{A}_s$, $\hat{E}_x$ на рис. 1, 3)

2) На основе полученных результатов выдвинуть гипотезу о законе распределения изучаемого признака.

3) Таблицы с №1, 2, 3, 4.

Таблица № 4

№ зад	n	R	h	s	a	$\bar{x}$	$\hat{\sigma}_x^2$	$\hat{A}_x$	$\hat{E}_x$	$\widehat{M}_0$	$\widehat{M}_e$	$\widehat{\mu}_{4y}$

Задание №2

Оценки параметров распределения

2.1. Общие сведения

В данном задании мы научимся вычислять точечные и интервальные оценки МО и дисперсии генеральной совокупности. Эти оценки имеют большое значение в статистике. Первая дает представление о характерной величине изучаемого признака, а вторая – о его разбросе.

Для того чтобы научиться вычислять значения оценок по точным формулам и увидеть влияние объема выборки на точность оценивания, будем проводить вычисления для выборки двух объемов: исходной $n = 50$ и сокращенной $n = 10$ (малой выборки). В качестве второй возьмем первые 10 чисел x_k с номерами k=00÷ 09 исходной выборки объемом n=50 . (Не путать с 10 числами, которые использовались в задании №1 для выбора точки отсчета интервального ряда $\bar{x}_0$!).

Интервальные оценки вычислим двумя способами: с помощью точных распределений Стьюдента и χ^2 , а также с помощью асимптотической нормальной аппроксимации. Это нужно, чтобы иметь представление о точности асимптотической нормальной аппроксимации при разных объемах выборки. Для асимптотической интервальной оценки дисперсии необходима также оценка четвертого центрального момента.

Обратите внимание на правила приближенных вычислений: оценка МО должна иметь столько цифр после запятой, сколько имеют выборочные данные. Оценки дисперсии и 4-го момента должны иметь столько значащих цифр, сколько имеет размах выборки R.

2.2 Точечные оценки параметров.

Таблица №5. Ошибки репрезентативности.

n	$\bar{x}$	s^2	$\sigma_{\bar{x}} = s/\sqrt{n}$	$\sigma_{S^2} = \sqrt{(\hat{\mu}_4 - s^4)/n}$
50				
10				

2.2.1. Построим таблицу №5. Заполняем первые три столбца верхней строчки для $n = 50$ значениями, найденными в разделе 1.4 с помощью метода произведений.

2.2.2. Точечную оценку $\bar{x}$ МО вносим в таблицу №5 без изменений.

2.2.3. В качестве оценки дисперсии σ^2 принимаем несмещенную оценку $S^2 = \frac{n}{n-1}\hat{\sigma}^2$;

2.2.4. Оценка четвертого центрального момента $\hat{\mu}_{4,50,x} = h^4\hat{\mu}_{4,y}$

2.2.5. Чтобы заполнить вторую строчку таб. 5 для выборки n=10 надо рассчитать точечные оценки $\bar{x}_{10}$, s^2_{10}, $\hat{\mu}_{4,10}$, а также ошибки репрезентативности $\sigma_{\bar{x}}$, и σ_{s^2} для данного объема. Точечные оценки МО, дисперсии и 4-го момента для $n = 10$ вычислим с помощью точных формул, которые содержатся в нижней строке втором, четвертом и пятом столбцах таб.6 .Для удобства вычислений строим таблицу № 6.

Таблица №6. Вычисление моментов для выборки объемом n = 10.

k	x_k	$x_k - \bar{x}_{10}$	$(x_k - \bar{x}_{10})^2$	$(x_k - \bar{x}_{10})^4$
00	x_{00}	$x_{00} - \bar{x}_{10}$	$(x_{00} - \bar{x}_{10})^2$	$(x_{00} - \bar{x}_{10})^4$
…	…	…	…	…
09	x_{09}	$x_{09} - \bar{x}_{10}$	$(x_{09} - \bar{x}_{10})^2$	$(x_{09} - \bar{x}_{10})^4$
	$\frac{1}{10}\sum_{k=00}^{09} x_k = \bar{x}_{10}$		$\frac{1}{9}\sum_{k=00}^{09}(x_k - \bar{x}_{10})^2 = S^2_{10}$	$\frac{1}{10}\sum_{k=00}^{09}(x_k - \bar{x}_{10})^4 = \hat{\mu}_{4,10}$

2.2.6. В первый столбец запишем номера вариант от 00÷ 09. Во второй столбец заносим первые 10 чисел выборки с номерами k: 00÷09, складываем их, делим на 10 и получаем $\bar{x}_{10}$.

2.2.7. Вычитаем $\bar{x}_{10}$ из каждого числа второго столбца и вносим в третий столбец.

2.2.8. Возводим каждое число третьего столбца в квадрат и вносим в четвертый столбец. Складываем числа в четвертом столбце, делим на 9, получаем S^2_{10} (Примечание: при малых значениях $(x_k - \bar{x}_{10})^2$ выделяем мантиссу и порядок)

2.2.9. Возводим каждое число четвертого столбца в квадрат и вносим в пятый столбец. Складываем числа в пятом столбце, делим на 10, получаем $\hat{\mu}_{4,10}$.

Полученные значения точечных оценок $\bar{x}_{10}$ и s_{10}^2 для n=10 вносим во второй и третий столбцы второй строки таб.5.

2.3. Асимптотические нормальные интервальные оценки.

2.3.0. Вычисляем ошибки репрезентативности: $\sigma_{\bar{x}}$ - ошибка среднего значения; σ_{S^2} – ошибка несмещенной оценки дисперсии по формулам из таблицы №5 соответственно для n=50 и n=10, используя данные второго и третьего столбцов.

2.3.1**.** Заполняем оставшиеся два столбца таблицы №5 следующим образом:

2.3.1.1. Ошибка среднего значения $\bar{x}$: $\sigma_{\bar{x}} = \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}}$

2.3.1.2. Ошибка несмещенной оценки дисперсии S^2: $\sigma_{S^2} - \sqrt{\frac{\hat{\mu}_4 - S^4}{n}}$

2.3.2. Строим две таблицы №7 для выборок $n = 50$ и $n = 10$. Заполняем каждую из них для трех уровней значимости α: 0,1, 0,05 и 0,01 следующим образом:

2.3.2.1.Используя *таблицы II значений функции* $F(x)$ *приложений*, найдем значения статистик $t_{1-\alpha/2}$ для трех уровней значимости α: 0,1;0,05 и 0,01.

2.3.2.2. Вычисляем границы интервалов для МО: $\hat{m}_1 = \bar{x} - t_{1-\alpha/2}\sigma_{\bar{x}}$, $\hat{m}_2 = \bar{x} + t_{1-\alpha/2}\sigma_{\bar{x}}$

2.3.2.3. Вычисляем границы интервалов для дисперсии σ^2:

$\hat{\sigma}_1^2 = s^2 - t_{1-\alpha/2}\sigma_{S^2}$, $\hat{\sigma}_2^2 = s^2 + t_{1-\alpha/2}\sigma_{S^2}$ (значения $\sigma_{\bar{x}}$ и σ_{S^2} берем из таблицы №5 для соответствующего объема n и исследуемого признака)

Таблица №7 (n=50) Интервальные оценки.

n	Пара-метр	Оценка	α		
			0,1	0,05	0,01
50	m	Асимптотическая	$\boldsymbol{t_{0,95}} =$	$\boldsymbol{t_{0,975}} =$	$\boldsymbol{t_{0,995}} =$
			$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$
		Точная	$\boldsymbol{t_{49,\,0,95}} =$	$\boldsymbol{t_{49,\,0,975}} =$	$\boldsymbol{t_{49,\,0,995}} =$
			$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$
	σ^2	Асимптотическая	$\boldsymbol{t_{0,95}} =$	$\boldsymbol{t_{0,975}} =$	$\boldsymbol{t_{0,995}} =$
			$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$
		Точная	$\boldsymbol{\chi^2_{49,\,0,05}, \chi^2_{49,\,0,95}}$	$\boldsymbol{\chi^2_{49,\,0,025}, \chi^2_{49,\,0,975}}$	$\boldsymbol{\chi^2_{49,\,0,005}, \chi^2_{49,\,0,995}}$
			$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$

Таблица №7 (n=10).

10	m	Асимптотическая	$\boldsymbol{t_{0,95}} =$	$\boldsymbol{t_{0,975}} =$	$\boldsymbol{t_{0,995}} =$
			$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$
		Точная	$\boldsymbol{t_{9,\,0,95}} =$	$\boldsymbol{t_{9,\,0,975}} =$	$\boldsymbol{t_{9,\,0,995}} =$
			$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$	$(\hat{m}_1, \hat{m}_2)$
	σ^2	Асимптотическая	$\boldsymbol{t_{0,95}} =$	$\boldsymbol{t_{0,975}} =$	$\boldsymbol{t_{0,995}} =$
			$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$
		Точная	$\boldsymbol{\chi^2_{9,\,0,05}, \chi^2_{9,\,0,95}}$	$\boldsymbol{\chi^2_{9,\,0,025}, \chi^2_{9,\,0,975}}$	$\boldsymbol{\chi^2_{9,\,0,005}, \chi^2_{9,\,0,995}}$
			$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$	$(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$

При заполнении таблицы № 7 обратите внимание, что первой записывается левая граница интервала (меньшее число), второй – правая (большее число).

2.4. Точные интервальные оценки параметров нормального закона.

2.4.1. Продолжаем заполнять таблицу №7.

2.4.2. Точные интервальные оценки для МО:

2.4.2.1. Найдем число степеней свободы $\nu = n - 1$

2.4.2.2. Используя *таблицы III приложений (Критерий Стьюдента , 2Q),* находим табличные значения статистики $t_{\nu,1-\alpha/2}$ для трех уровней значимости α:0,1;0,05 и 0,01 .

2.4.2.3. Вычисляем границы интервалов для МО:

$\hat{m}_1 = \bar{x} - t_{\nu,1-\alpha/2}\sigma_{\bar{x}}$, $\hat{m}_2 = \bar{x} + t_{\nu,1-\alpha/2}\sigma_{\bar{x}}$

2.4.3. Точные интервальные оценки для дисперсии:

2.4.3.1. Находим табличные значения $\chi^2_{\nu,\alpha/2}$ и $\chi^2_{\nu,1-\alpha/2}$ для трех уровней значимости α: 0,1;0,05;0,01; (*Таблица XI приложений* χ^2 *для* $n \leq 10$ + графическая линейная интерполяция для неизвестных значений $\chi^2_{\nu,\alpha/2}$).(Рис.6)

2.4.3.2. Значения $\chi^2_{\nu,1-\alpha/2}$ и $\chi^2_{\nu,2,1-\alpha/2}$ для объема выборки n=50 находятся по формуле:

$$\chi^2_{\nu,1,2} = \frac{(\sqrt{2\nu - 1} \pm t_{1-\alpha/_2})^2}{2}$$

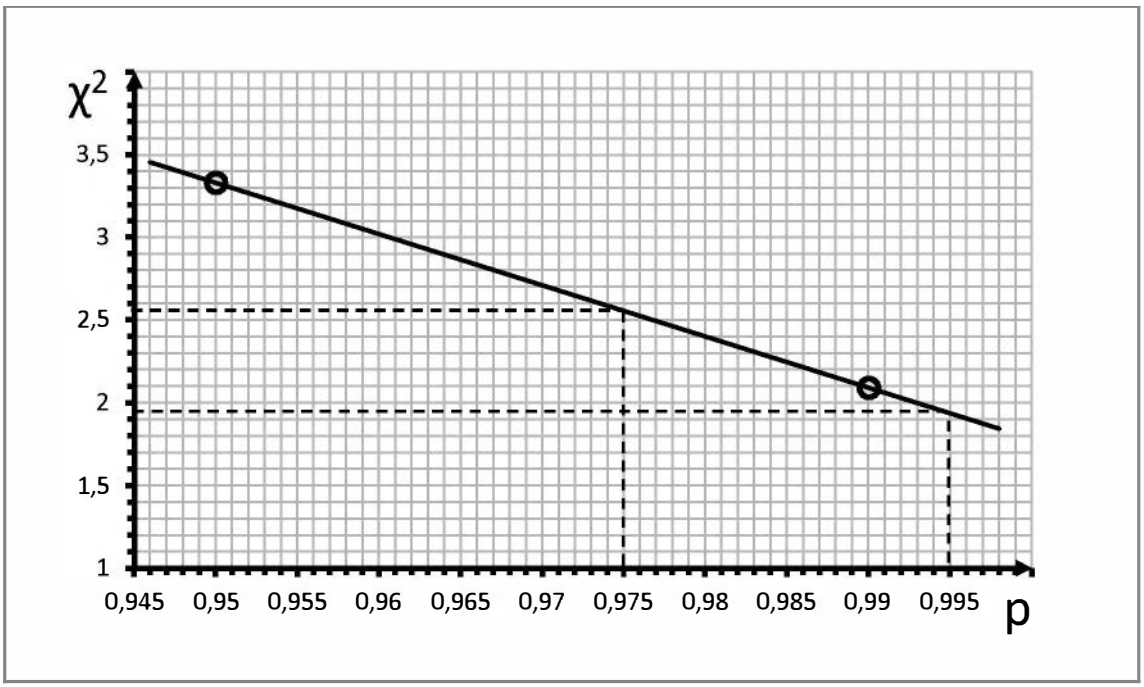


Рисунок 6. Графическая линейная интерполяция.

2.4.3.3. Вычислите границы интервалов для дисперсии:

$$\hat{\sigma}_1^2 = \frac{\nu\, S^2}{\chi^2_{\nu,1-\alpha/2}}, \quad \hat{\sigma}_2^2 = \frac{\nu\, S^2}{\chi^2_{\nu,\,\alpha/2}}$$

2.4.4. Таблица №7 заполняется для объемов выборки $n = 50$ и $n = 10$ исследуемых признаков.

Задание №3

Проверка статистических гипотез

3.1. Общие сведения

В данном задании мы научимся проверять гипотезы о математическом ожидании (МО) и дисперсии генеральной совокупности (ГС), из которой была извлечена выборка. Во-первых, проверим гипотезы о равенстве МО и дисперсии ГС заданным значениям (одновыборочные гипотезы). Проверка этих гипотез может решить следующие практические задачи:

1. Извлечена ли данная выборка случайным образом из ГС с некоторым эталонным значением МО (m_0) или дисперсией (σ_0^2)? Эта задача требует проверки двусторонней гипотезы.
2. Превышает ли МО или дисперсия ГС, из которой извлечена данная выборка, некоторое критическое значение? Здесь требуется проверить одностороннюю гипотезу.
3. Укладывается ли МО или дисперсия ГС, из которой извлечена данная выборка, в заданный интервал? Эта задача сводится к проверке двух односторонних гипотез.

Во-вторых, будем проверять двухвыборочные гипотезы о МО и дисперсии. Для этого потребуется вторая выборка с тем же признаком, но относящаяся к другому типу объектов (см. 1.1.3). Сравнение выборок позволяет установить:

1. Отличаются ли ГС, из которых были извлечены выборки, по уровню величины МО?
2. Отличаются ли эти ГС по разбросу величин?

Утверждение, которое требуется проверить, закладывается в нулевую гипотезу и обозначается H_0. В противовес содержанию нулевой гипотезы выдвигается альтернативная гипотеза и обозначается через H_1. Для проверки каждой гипотезы в зависимости от её содержания существует свое правило. Это правило называется статистическим критерием значимости. Каждому критерию значимости соответствует статистика, закон распределения которой известен из теории вероятностей. Для принятия решения в пользу той или иной гипотезы строится решающее правило, согласно которому если значение статистики выбранного критерия превышает критическое значение, то справедливо

утверждение, что найдены статистически значимые различия (H_1). Если значение статистики выбранного критерия меньше критического, то статистически значимые различия отсутствуют (H_0).

Статистические критерии делятся на две большие группы: параметрические и непараметрические. При использовании параметрических критериев вид распределения ГС предполагается известным. Чаще всего предполагается, что это нормальное распределение. И если это действительно так, то параметрические критерии имеют наибольшую мощность. Но если в действительности распределение ГС существенно отличается от нормального распределения, то используя параметрические критерии, созданные для нормального распределения, можно сделать совершенно необоснованные выводы.

Поэтому перед применением параметрических критериев требуется производить проверку гипотезы о законе распределения. Если гипотеза о нормальности распределения отклоняется или данных для её проверки недостаточно (объем выборки $n < 30$), необходимо пользоваться непараметрическими критериями. За счет некоторой потери мощности эти критерии позволяют не делать предположений о форме распределения ГС.

Часто свойства критериев значимости описывают в несколько иных терминах: не вероятностью правильного результата, а вероятностью ошибки.

Каждую гипотезу будем проверять при трех уровнях значимости $\alpha = 0{,}1;\ 0{,}05$ и $0{,}01$. Максимально приемлемая вероятность ошибки I рода называется уровнем значимости и обозначается α, то есть вероятность принятия H_1 при справедливости H_0 (α=P(H_1/H_0)). Соответственно, если H_0 отклоняется при уровне значимости α, она отклоняется и при всех больших уровнях значимости (то есть меньших α).

Ошибка первого рода p оказывается меньше минимального уровня значимости, при котором отклоняется H_0 , поэтому, когда хотят охарактеризовать эту ошибку, пишут «$p < \alpha$», например $p < 0{,}05$.

На практике при проверке гипотез для заданных трех уровней значимости α могут встретиться результаты принятия гипотез, которые описываются следующим образом:

1. H_0 принимается при всех α, тогда записывается $p > 0{,}1$: нет оснований предполагать справедливость H_1.
2. H_0 отклоняется при $\alpha = 0{,}1$ и принимается при $\alpha = 0{,}05$ и $0{,}01$, записывается $p < 0{,}1$: H_1 можно предполагать с низкой степенью достоверности.
3. H_0 отклоняется при α=0,1 и 0,05 и принимается при α=0,01,→ $p < 0{,}05$: H_1 принимается со средней степенью достоверности.
4. H_0 отклоняется при всех фиксированных α, → p < 0,01: гипотеза H_1 принимается с высокой степенью достоверности.

При формулировке вывода следует переписать соответствующий пункт приведенного выше списка, заменяя H_0 и H_1 их словесными формулировками и

указав значение p. Например, если H_0 – «МО двух ГС равны», а H_1 – «МО первой ГС больше, чем второй», возможны следующие выводы.

1. Нет оснований предполагать, что МО первой ГС больше, чем МО второй ($p > 0{,}1$).
2. Можно предполагать с низкой степенью достоверности, что МО первой ГС больше, чем второй ($p < 0{,}1$), то есть при 10% уровне значимости.
3. МО первой ГС больше, чем МО второй со средней степенью достоверности ($p < 0{,}05$), то есть при 5% уровне значимости.
4. МО первой ГС больше, чем МО второй с высокой степенью достоверности ($p < 0{,}01$), то есть при 1% уровне значимости.

Как правило, рассматриваются два вида альтернатив: двусторонние и односторонние. В первом случае критическими являются как очень большие, так и очень маленькие значения статистики, поэтому с обеих сторон её распределения отсекаются интервалы с вероятностью $\alpha/2$. То есть H_1 принимается, если $t \le t_{\alpha/2}$ или $t \ge t_{1-\alpha/2}$, где t – значение статистики, а t_p – квантиль её распределения порядка p.

Если альтернатива односторонняя, то критическими являются только очень большие (правая односторонняя альтернатива) или только очень маленькие (левая) значения статистики. В случае правой альтернативы критическим будет интервал $t \ge t_{1-\alpha}$, а в случае левой → $t \le t_{\alpha}$.

Таблица №8. Критерий Пирсона.

γ_k	m_k	$t_k = \frac{\gamma_k - \bar{x}}{\hat{\sigma}}$	$\Phi_k = F_N(t_k)$	p_k	np_k	$\frac{(m_k - np_k)^2}{np_k}$
γ_1				Φ_1		
γ_2				$\Phi_2 - \Phi_1$		
…	…	…	…	…	…	…
γ_{s-1}				$\Phi_{s-1} - \Phi_{s-2}$		
$\gamma_s\ (\infty)$			1	$1 - \Phi_{s-1}$		

$$\cdot\cdot \chi_v^2 = \sum \frac{(m_k - np_k)^2}{np_k}$$

3.2. Проверка гипотез о законе распределения

В этом разделе мы рассмотрим два критерия проверки сложной гипотезы о нормальности.

H_0: f(x) $= N(m, \sigma^2) = N(\bar{x}, \hat{\sigma}^2)$ - Распределение ГС является нормальным с неизвестными заранее параметрами ГС математическим ожиданием m и дисперсией σ^2.

H_1: f(x) $\neq N(\mathrm{m}, \sigma^2) \neq N(\bar{x}, \hat{\sigma}^2)$ - Распределение ГС отлично от нормального.

(Проверяем наши предположения о нормальном законе распределения, сделанные в первом задании на основе аппроксимации гистограммы плотностью распределения нормального закона в ГС).

3.2.1. Критерий согласия Пирсона (χ^2)

Критерий χ^2−это универсальный критерий для проверки гипотез о законе распределения, которым можно пользоваться для распределения любого вида. Кроме того, критерий Пирсона можно применять для проверки сложных гипотез, когда параметры теоретического распределения оцениваются по выборке.

Недостатком этого критерия является необходимость группировки данных. В каждом интервале должно быть не меньше 5 значений. Это конечно снижает мощность критерия, особенно при небольшом объеме выборки. Однако если вычисления производятся вручную, это свойство имеет положительную сторону: группировка данных снижает объем вычислений.

3.2.1.1. Заполняем табл.№8. Из статистической табл. № 1 перепишите правые границы интервалов γ_k и соответственно их частоты m_k.

3.2.1.2. Вычисляем нормированные границы интервалов $t_k = \frac{\gamma_k - \bar{x}_{50}}{\hat{\sigma}_{50}}$ (используя значения $\bar{x}$ и $\hat{\sigma}^2$ из табл. №4 первого задания).

3.2.1.3. По *табл.II Значений функции F(x)* стандартного нормального распределения приложений найдем *значения* Φ_k , соответствующие t_k так что $\Phi_k = F_N(t_k)$. Учтите, что в таблице приведены значения $F_N(t)$ только для $t \geq 0$. Величины для $t < 0$ можно получить по формуле $F_N(-t) = 1 - F_N(t)$. $F_N(\infty) = 1$.

3.2.1.4. Вычисляем теоретические вероятности p_k: $p_1 = \Phi_1$, $p_k = \Phi_k - \Phi_{k-1}$ для к≥ 2.

3.2.1.5. Умножаем теоретические вероятности p_k на объем выборки n.

3.2.1.6. Полученные значения np_k называются ожидаемыми частотами. Часто оказывается, что в крайних интервалах вариационного ряда ожидаемых частот очень мало, поэтому интервалы объединяются так, чтобы минимальная частота np_k (соответственно и m_k) была по возможности ≥ 5, но при этом получилось не меньше 4 интервалов. Количество интервалов после объединения обозначается символом s'.

3.2.1.7. Вносим в последний столбец слагаемые статистики χ^2, равные

$$\frac{(m_k - np_k)^2}{np_k}$$

3.2.1.8. Вычисляем статистику χ^2_v, сложив числа в последнем столбце таблицы №8.

3.2.1.9. Сосчитаем число степеней свободы $v = s' - r - 1$, где r- число параметров, оцениваемых по выборке, равное в данном случае 2 (m и σ^2); s'- число новых интервалов после объединения.

3.2.1.10. В *таблице XI квантилей распределения* χ^2 приложений найдем квантили $\chi^2_{\nu,1-\alpha}$ для $\alpha = 0{,}1$, 0,05 и 0,01 для данного числа степеней свободы. Гипотеза H_1 принимается, если $\chi^2_\nu \geq \chi^2_{\nu,1-\alpha}$.
3.2.1.11. Сформулируйте вывод о законе распределения и представьте в виде таблицы №9:

χ^2 / α	0,1	0,05	0,01
$\chi^2_{\nu,1-\alpha}$			
χ^2_ν			
H:			

3.2.2. Критерий согласия Колмогорова-Смирнова-Лильефорса

Классический - критерий Колмогорова-Смирнова рассчитан на проверку простой гипотезы о виде распределения, когда все его параметры известны априори. В этом случае распределение статистики не зависит от распределения ГС. Когда параметры оцениваются по выборке, теоретическое распределение фактически подгоняется к данным, поэтому невязка, естественно, будет меньше при всех прочих равных условиях. Статистика χ^2 при этом не меняет вид распределения, только уменьшается число степеней свободы. А вот распределение – статистики становится другим. Более того, оно начинает зависеть от вида распределения ГС.

С помощью данного критерия проверяем наши предположения, сделанные в первом задании на основе аппроксимации кумуляты функцией распределения нормального закона в ГС, то есть следующие гипотезы:

H_0: F(x) =F_0(x); H_1: F(x) ≠ F_0(x);

Критерий Колмогорова – Смирнова для проверки сложной гипотезы о нормальности называется критерием Лильефорса.

Статистика вычисляется так же, как в классическом случае, но критические значения будут другими. Статистика Лильефорса L считается по не группированным данным следующим образом:

1. Расположите выборочные значения x_k в порядке возрастания.
2. Вычислите нормированные значения t_k выборочных значений x_k (вычесть среднее $\bar{x}_{50}$ и поделить на оценку среднеквадратического отклонения $\hat{\sigma}_{50}$, как выше описано в критерии Пирсона).

$$t_k = \frac{x_k - \bar{x}_{50}}{\hat{\sigma}_{50}}$$

3. Вычислите для каждого значения t_k значение функции стандартного нормального распределения $\Phi_k = F_N(t_k)$ *по табл.II* (см. приложения), умноженное на объем выборки $nF_N(t_k)$.

4. Запишите значения эмпирической функции распределения, умноженные на n. Поскольку эта функция ступенчатая, каждому значению t_k будут соответствовать два значения Φ_k, справа и слева: k и k-1.
5. Найти максимальный модуль разности между теоретической и эмпирической функциями распределения $L = \max_k(nF_N(t_k) - k, nF_N(t_k) - k - 1)/n$. Так как для удобства вычислений мы использовали значения функций, умноженные на n, теперь мы должны поделить на n.

Таблица №10. Критерий Колмогорова-Смирнова-Лильефорса.

γ_k	$t_k = \frac{\gamma_k - \bar{x}}{\sigma_x}$	$n\Phi_k = nF_N(t_k)$	M_k	$\Delta_k = \lvert n\Phi_k - M_k \rvert$
γ_1				
γ_2				
…	…	…	…	…
γ_s				

Но поскольку без помощи ЭВМ довольно сложно отсортировать 50 чисел, в учебных целях будем считать критерий по группированным данным. Надо подчеркнуть, что при группировке значение статистики уменьшается, и мощность критерия заметно падает. Поэтому на практике следует проводить вычисления по не группированным данным.

3.2.2.1. Заполните таблицу №10. Из статистической таблицы №1 перепишите правые границы интервалов γ_k. Объединять интервалы не нужно, поскольку, чем больше интервалов, тем больше мощность критерия.

3.2.2.2. Вычислите нормированные границы интервалов t_k (см. таб.8).

3.2.2.3. По *таблице II приложений* функции стандартного нормального распределения $F(x)$ найдите значения $\Phi_k = F_N(t_k)$, умножьте их на объем выборки n.

3.2.2.4. Запишите накопленные частоты M_k для каждого интервала из таблицы №1.

3.2.2.5. Вычислите модули разностей между $n\Phi_k$ и M_k (т. е. см. табл. №10 $\Delta_k = \lvert n\Phi_k - M_k \rvert$), найдите максимальный из них $L = \Delta_{max}$. В силу отсутствия таблиц для критерия Лильефорса можно воспользоваться критерием Колмогорова-Смирнова, который проверяет ту же гипотезу, что и критерий Лильефорса . Для этого достаточно вычислить статистику $\lambda = \frac{L}{\sqrt{n}}$.

3.2.2.6. Можно было самостоятельно вычислить статистику λ Колмогорова-Смирнова по правилу $\lambda = \sqrt{n} \max\left|\Phi_k - \frac{M_k}{n}\right|$.

3.2.2.7. В *Таблице VIII Критерий Колмогорова* приложений найдите критические значения статистики Колмогорова λ для $\alpha = 0{,}1$, $0{,}05$ и $0{,}01$. Гипотеза H_1 принимается, если $\lambda \geq \lambda_{1-\alpha}$

3.2.2.8. Сформулируйте вывод о законе распределения и представьте в виде таблицы:

Таблица № 11

λ ---------- α	0,1	0,05	0,01
$\lambda_{1-\alpha}$			
λ			
H:			

3.3. Непараметрические критерии для проверки гипотез о совпадении двух распределений.

Проверяется гипотеза о совпадении двух распределений:

$H_0: p_x = p_y$

$H_1: p_x \neq p_y$

При проверке данной гипотезы применяется непараметрические методы, которые используют знаки, ранги, инверсии и не требуют знаний о законах распределения исследуемых признаков X и Y в генеральных совокупностях.

Непараметрические критерии работают на малых выборках ($n < 30$), на связанных (зависимых) и несвязанных (независимых) совокупностях.

Зависимые совокупности $\{x_k\}$ и $\{y_k\}$ признаков X и Y связываются, как правило, через объект, причем объемы совокупностей должны быть равными, то есть $n_x = n_y$. Например, исследуется число сердечных сокращений (ЧСС) до нагрузки $-$признак X и после нагрузки – Y на одних и тех же объектах. Для независимых совокупностей признак X и Y (ЧСС) исследуется на разных объектах (например, $X -$ контрольная совокупность, $Y -$ экспериментальная), при этом объемы выборок n_x и n_y не обязательно одинаковые, то есть n_x может быть больше n_y или $n_x \neq n_y$. Итак, для связанных (зависимых) совокупностей используется критерий знаков и парный критерий Уилкоксона.

При проверке гипотезы о совпадении двух распределений требуются данные партнера, исследующего тот же признак только другого типа.

Таблица № 12.

$X: x_k$	x_{00}	x_{01}	…	x_{09}
$Y: y_k$	y_{00}	y_{01}	…	y_{09}
$x_k - y_k$	$x_{00} - y_{00}$	$x_{01} - y_{01}$	…	$x_{09} - y_{09}$
$R_{\lvert x_k - y_k \rvert}$				

3.3.1. Критерий знаков.

3.3.1.1. Формируем две связанные совокупности X:x_k и $Y: y_k$ с номерами k = 00 ÷ 09 из исходных выборок одного и того же признака разных типов

(например, площадь лимфоцитов и нейтрофилов, или ЧСС до эксперимента и ЧСС после эксперимента) одинакового объема $n_x = n_y = 10$

3.3.1.2. Вычитаем $x_k - y_k$, $k: 00 \div 09$ (см. табл. №12)

3.3.1.3. Статистика, соответствующая критерию знаков, равна $n = \min(n_+, n_-)$, где n_+ – сумма « + » при положительных разностях; n_- – сумма « − » при отрицательных разностях.

3.3.1.4. Проверочная формула $n_+ + n_- = n$

3.3.1.5.По *табл. XIII приложений находим критическое значение* n_α *критерия знаков.*

3.3.1.6. Условие принятия гипотез: $n \leq n_\alpha \rightarrow \overline{H}_0(H_1) \rightarrow$ отклоняется H_0(принимается H_1)

3.3.1.7. Полученный результат принятия гипотез представляем в виде таблицы:

Таблица № 13.

n \ α	0,05	0,01
n_α		
n		
H:		

3.3.2. Парный критерий Уилкоксона (Вилкоксона).

3.3.2.1.В таблице №12 отбрасываем знаки у разностей и расставляем порядковые номера (ранги) по возрастанию модуля разностей от меньшего модуля к большему.

3.3.2.2. Статистика парного критерия Уилкоксона $T = \min(T_+, T_-)$, где T_+ −сумма рангов при положитеьных разностях и T_-— сумма рангов при отрицательных разностях.

3.3.2.3. Проверочная формула $T_+ + T_- = \frac{n(n+1)}{2}$

3.3.2.4. *По табл.XIV приложений находим критическое значение* T_α *парного Т-критерия Уилкоксона.*

3.3.2.5. Условие принятия гипотез: $T \leq T_\alpha \rightarrow \overline{H}_0(H_1)$, то есть отклоняется гипотеза H_0, принимается H_1.

3.3.2.6. Полученные результаты принятия гипотез представляем в виде таблицы:

Таблица № 14.

T \ α	0,05	0,01
T_α		
T		
H:		

Для несвязанных (независимых) совокупностей мы проверяем гипотезы о совпадении двух распределений, используя критерий Уилкоксона и критерий инверсий.

3.3.3. Критерий Уилкоксона.

3.3.3.1. Составляем общий вариационный ряд признака X своего задания и такого же признака У из задания партнера, относящегося к другому типу объектов по возрастанию признака, оставляя значения признака в своей строке.

Таблица № 15.

$X: x_k$	x_{00}		x_{01}	x_{02}	x_{03}			…	x_{10}	
$Y: y_k$		y_{00}				y_{01}	y_{02}	…		y_{10}
R	1	2	3	4	5	6	7	…	19	20

3.3.3.2. Присваиваем ранги по возрастанию признака (табл.№15).

3.3.3.3. По критерию Уилкоксона подсчитываем статистику $T = \min(T_x, T_y)$, где $T_x -$ сумма рангов при значениях x_k, $T_y -$ сумма рангов при значениях y_k.

3.3.3.4. *По таблицам XXI приложений находим теоретические значения* T_α *для независимых (несовместных) совокупностей критерия Уилкоксона.*

3.3.3.5.Гипотеза H_1 принимается, если $T \le T_{\alpha\ \text{несв.с.}}$, $\left(T \le T_\alpha \to H_1(\overline{H_0})\right)$.

3.3.3.6. Проверочная формула: $T_x + T_y = \frac{(n_x+n_y)(n_x+n_y+1)}{2}$.

3.3.3.7. Результаты принятия гипотез собираем в таблицу:

Таблица № 16.

T ---------- α	0,05	0,01
T_α		
T		
H:		

3.3.4. Критерий инверсий

3.3.4.1.Выбираем статистику $U = \min(U_x, U_y)$, где $U_x -$ число инверсий по x (изменение порядка x по отношению y), и $U_y -$ число инверсий по y (изменение порядка следования y по отношению x).

3.3.4.2. Для этого выписывается общий ряд следования

x x y Y y y x x x … x x y

3.3.4.3.Подсчитываем количество х-в перед каждым у-ом от начала ряда слева направо. Полученные значения суммируем и обозначаем U_x. И, наоборот, подсчитываем количество у-в перед каждым х-ом от начала ряда слева направо. Полученные значения суммируем и обозначают U_y.

3.3.4.4. Проверочная формула: $U_x + U_y = n_x \cdot n_y$

3.3.4.5. *По табл.XII приложений U –критерия Манна-Уитни находят критическое значение* U_α .

3.3.4.6. Гипотеза H_1 принимается, если $U \leq U_\alpha$. Иначе, если $U \leq U_\alpha \rightarrow H_1(\bar{H}_0)$

3.3.4.7. Результаты принятия собираем в таблицу:

Таблица № 17.

U ---------- α	0,05	0,01
U_α		
U		
H:		

3.4 Проверка гипотез о параметрах распределения

3.4.1. Критерий Стьюдента для одной выборки.

Критерий Стьюдента применяется для проверки гипотез о математическом ожидании (МО).

Этот критерий позволяет сравнивать МО m с некоторым фиксированным значением m_0. К критерию Стьюдента для одной выборки сводится так же критерий для двух выборок, если ГС связанные. Альтернатива может быть как двусторонней (2Q) , так и односторонней (Q).

3.4.1.1. Поверяем гипотезу H_0 о совпадении со значением m_0 против правой односторонней альтернативы $H_1: m > m_0$. Впредь, гипотезы будем записывать в следующем виде:

$H_0: m = m_0$

$H_1: m > m_0$ (Q)

m_0 указано в задании, расчеты выполняются и для $n = 50$, и для $n = 10$.

3.4.1.2. Вычислите статистику $t_\nu = \frac{\bar{x} - m_0}{\sigma_{\bar{x}}}$, где значения n, $\bar{x}$ и $\sigma_{\bar{x}}$ берутся из табл.№5, $\nu = n - 1$ – число степеней свободы.

3.4.1.3. *Найдите в таблице III приложений критерий Стьюдента для Q* критические значения $t_{\nu,1-\alpha}$ для $\nu = n - 1$ и трех $\alpha = 0{,}1$, 0,05 и 0,01.

3.4.1.4. Сравните t_ν и $t_{\nu,1-\alpha}$: гипотеза H_1 принимается, если $t_\nu \geq t_{\nu,1-\alpha}$

3.4.1.5. Результаты оформите в виде двух таблиц №18: для n=50 и n=10

3.4.1.6. Запишите выводы.

3.4.1.7. Сравните выводы для $n = 50$ и $n = 10$.

Таблица № 18 (n=50)

t ---------- α	0,1	0,05	0,01
$t_{\nu,1-\alpha}$			
t_ν			
H:			

Таблица №18 (n=10)

t / α	0,1	0,05	0,01
$t_{\nu,1-\alpha}$			
t_{ν}			
H:			

3.4.1.8. Теперь проверим эту же гипотезу H_0 против двусторонней альтернативы:

H_0: m = m_0

H_1: m ≠ m_0 (2Q)

3.4.1.9 . *Найдите в таблице III приложений критерий Стьюдента для 2Q* критические значения $t_{\nu,1-\frac{\alpha}{2}}$ для ν = n-1 и трех α = 0,1;0,05 и 0.01.

3.4.1.10. Сравните t_{ν} и $t_{\nu,1-\frac{\alpha}{2}}$: гипотеза H_1 принимается, если $|t_{\nu}| \geq t_{\nu,1-\frac{\alpha}{2}}$

3.4.1.11. Результаты оформите в виде двух таблиц №19 для n=50 и n=10

Таблица №19 (n−50).

t / α	0,1	0,05	0,01
$t_{\nu,1-\frac{\alpha}{2}}$			
t_{ν}			
H:			

Таблица №19 (n=10).

t / α	0,1	0,05	0,01
$t_{\nu,1-\frac{\alpha}{2}}$			
t_{ν}			
H:			

3.4.1.12. Запишите выводы.

3.4.1.13. Сравните выводы для n = 50 и n = 10.

3.4.2. Критерий χ^2

Этот критерий позволяет проверить аналогичную гипотезу для дисперсии (ГС) σ^2, сравнивать её с некоторым фиксированным значением σ_0^2. Правая односторонняя альтернатива используется при контроле качества: если дисперсия показаний прибора достоверно превышает паспортное значение, возможна его неисправность.

3.4.2.1. Проверяем гипотезу H_0 о совпадении с фиксированным значением дисперсии σ_0^2 против двусторонней альтернативы H_1:

H_0: $\sigma^2 = \sigma_0^2$
H_1: $\sigma^2 \neq \sigma_0^2$ (2Q)
σ_0^2 указано в задании. Расчеты выполняются и для $n = 50$, и для $n = 10$.
3.4.2.2. Вычисляем статистику $\chi_\nu^2 = \frac{(n-1)S^2}{\sigma_0^2}$, где $\nu = n - 1$; n и S^2 берутся из табл.№5.
3.4.2.3. *Найдите в таблице XI приложений распределения* χ^2 *критические значения* $\chi^2_{\nu,\alpha/2}$ *и* $\chi^2_{\nu,1-\alpha/2}$
для $\nu = n - 1$ и трех $\alpha = 0{,}1$, 0,05 и 0,01.
3.4.2.4. Сравните χ_ν^2 с критическими значениями: гипотеза H_1 принимается, если $\chi_\nu^2 \leq \chi^2_{\nu,\alpha/2}$ или $\chi_\nu^2 \geq \chi^2_{\nu,1-\frac{\alpha}{2}}$. Либо принимается гипотеза H_0 , если

$$\chi^2_{\nu,\frac{\alpha}{2}} < \chi_\nu^2 < \chi^2_{\nu,1-\frac{\alpha}{2}}$$

3.4.2.5. Результаты оформите в виде двух таблиц №20 для n=50 и n=10.
3.4.2.6. Запишите выводы.
3.4.2.7. Сравните выводы для $n = 50$ и $n = 10$.

Таблица № 20 (n=50).

χ^2 / α	0,1	0,05	0,01
$\chi^2_{\nu,\frac{\alpha}{2}}$; $\chi^2_{\nu,1-\alpha/2}$			
χ_ν^2			
H:			

Таблица №20 (n=10)

χ^2 / α	0,1	0,05	0,01
$\chi^2_{\nu,\frac{\alpha}{2}}$; $\chi^2_{\nu,1-\alpha/2}$			
χ_ν^2			
H:			

3.4.3. Критерий Фишера
Этот критерий позволяет сравнивать дисперсии двух генеральных совокупностей σ_x^2 и σ_y^2 между собой для одноименной переменной, используя полученные несмещенные оценки S_x^2 и S_y^2 в таблицах №5 для разных типов объектов. S_x^2 берется из табл.№5 своего задания, а S_y^2 – из табл.№5 партнера для одноименной переменной, но другого типа объектов.
3.4.3.1. Проверяем гипотезу H_0 о совпадении двух дисперсий против двусторонней альтернативы H_1
H_0: $\sigma_x^2 = \sigma_y^2$
H_1: $\sigma_x^2 \neq \sigma_y^2$ (2Q)
Расчеты выполняются и для $n = 50$, и для $n = 10$.

3.4.3.2. Вычислите статистику $F_{\nu_x,\nu_y} = \frac{S_x^2}{S_y^2}$, если $S_x^2 > S_y^2$, иначе $F_{\nu_x,\nu_y} = \frac{S_y^2}{S_x^2}$, если $S_y^2 > S_x^2$. То есть в числитель дроби ставите большую дисперсию из двух дисперсий $S_x{}^2$ и $S_y{}^2$, где $\nu_x = \nu_y = n - 1$.

3.4.3.3. *Найдите в таблицах IV, V и VII критерия Фишера F критические значения* $F_{\nu_x,\nu_y,1-\alpha/2}$

для $\nu_x = n_x - 1 = n - 1$, $\nu_y = n_y - 1 = n - 1$ и трех α = 0,1, 0,05 и 0,01.

3.4.3.4. Сравните F_{ν_x,ν_y} и $F_{\nu_x,\nu_y,1-\alpha/2}$: гипотеза H_1 принимается,

если $F_{\nu_x,\nu_y} \geq F_{\nu_x,\nu_y,1-\alpha/2}$.

3.4.3.5. Результаты оформите в виде двух таблиц № 21 для n=50 и n=10 .

3.4.3.6. Запишите выводы.

3.4.3.7. Сравните выводы для $n = 50$ и $n = 10$.

Таблица № 21(n=50).

F ---------- α	0,1	0,05	0,01
$F_{\nu_x,\nu_y,1-\alpha/2}$			
F_{ν_x,ν_y}			
H:			

Таблица № 21 (n–10)

F ---------- α	0,1	0,05	0,01
$F_{\nu_x,\nu_y,1-\alpha/2}$			
F_{ν_x,ν_y}			
H:			

3.4.4. Критерий Стьюдента для двух выборок.

Этот критерий позволяет сравнивать математические ожидания между собой и доказывает достоверность различий двух ГС по уровню данного признака. Критерий Стьюдента для двух выборок можно применять только при отсутствии достоверных различий по дисперсии.

3.4.4.1. Используем критерий Стьюдента для проверки гипотезы H_0 о совпадении математических ожиданий m_x и m_y для одного и того же признака, относящихся к разным типам объектов против двусторонней альтернативы H_1 о неравенстве МО. Значения $\bar{x}$ и $\sigma_{\bar{x}}$ берутся из таблицы №5 своего задания, а значения $\bar{y}$ и $\sigma_{\bar{y}}$ того же признака, но относящегося к другому типу объектов, из таблицы №5 партнера.

H_0: $m_x = m_y$

H_1: $m_x \neq m_y$ (2Q)

Расчеты выполняются и для $n = 50$, и для $n = 10$.

3.4.4.2. Вычислите статистику $t_\nu = \frac{\bar{x}-\bar{y}}{\sqrt{\sigma_{\bar{x}}^2+\sigma_{\bar{y}}^2}}$, где $n_x = n_y$.

3.4.4.3. *Найдите в таблице III приложений критерия Стьюдента критические значения* $t_{\nu,1-\alpha/2}$ *для 2Q* и
для $\nu = n_x + n_y - 2 = 2n - 2$ и трех $\alpha = 0{,}1$; 0,05 и 0,01.
3.4.4.4. Сравните t_ν и $t_{\nu,1-\alpha/2}$: гипотеза H_1 принимается, если $|t_\nu| \geq t_{\nu,1-\alpha/2}$.
3.4.4.5. Результаты оформите в виде двух таблиц № 22 для n=50 и n=10.
3.4.4.6. Запишите выводы.
3.4.4.7. Сравните выводы для $n = 50$ и $n = 10$.

Таблица № 22 (n=50).

t / α	0,1	0,05	0,01
$t_{\nu,1-\alpha/2}$			
t_ν			
H:			

Таблица № 22 (n=10)

t / α	0,1	0,05	0,01
$t_{\nu,1-\alpha/2}$			
t_ν			
H:			

Задание №4

Элементы корреляционного анализа

4.1. Общие сведения

В данном задании мы будем анализировать зависимость между двумя признаками (переменными) X и Y. Во-первых, исследуем линейную зависимость между признаками X и Y в генеральной совокупности. Для этого вычислим точечные $\hat{r}_{x,y}$ и интервальные оценки $(\hat{r}_1, \hat{r}_2]$ неизвестного параметрического коэффициента корреляции Пирсона $r_{x,y}$ в генеральной совокупности, если она подчинятся нормальному закону распределения. Проверим гипотезы о значимости вычисленных коэффициентов корреляции (КК) и о равенстве КК Пирсона $r_{x,y}$ заданному значению r_0. Во-вторых, исследуем нелинейную зависимость между признаками X и Y в генеральной совокупности с помощью непараметрических коэффициентов корреляции ρ Спирмена и τ- Кенделла.

Если КК оказывается, достоверно отличен от 0, это говорит о наличии линейной зависимости между переменными. Условия применения параметрического и непараметрических КК аналогичны условиям применения параметрических и непараметрических критериев при проверке статистических гипотез о законах распределения и параметрах распределения: непараметрические КК применяются при отличии распределения от нормального, при $n < 30$ или если данные не являются числовыми.

4.2. Построение корреляционной решетки

Корреляционная решетка – это двумерная статистическая таблица, с помощью которой можно оценить вид совместного распределения, наличие и вид зависимости между переменными.

Для выполнения задания используем исходную двумерную выборку, где первый столбец принимаем за X, второй за Y. Таким образом, получаем n пар (в данной работе $n = 50$) значений (x_k, y_k), где k: $(\overline{00, 49})$ и рассматриваем, как точку с двумя координатами.

4.2.1. Работаем с исходной выборкой объемом $n_x = n_y = 50$.

4.2.2. Для выборок X и Y границы интервалов берем из первого задания.

4.2.3. Строим корреляционную решетку, как на рис. 6, располагая границы интервалов γ_{x_k} для X по оси абсцисс, а γ_{y_k} – по оси ординат для Y. Обратите внимание, что γ_{x_k} возрастает слева направо, а γ_{y_k} – снизу вверх!

4.2.4. Просмотрите последовательно пары выборочных значений. Определяйте, в какую клетку попадает пара, используя кодировку пятерками, как в задании 1.

4.2.5. Оцените наличие зависимости между переменными и её характер (прямая или обратная).

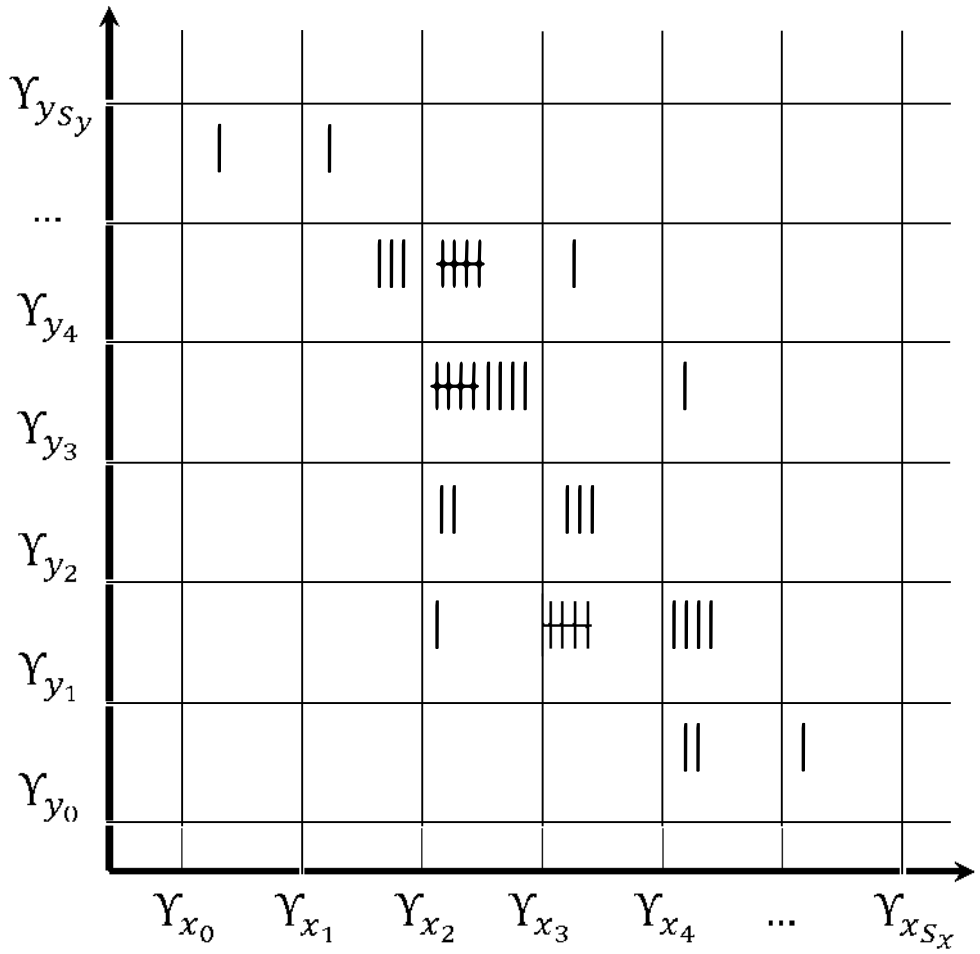


Рисунок №6. Корреляционная решетка.

4.3. Метод произведений для вычисления коэффициента корреляции

4.3.1. Заполняем табл. №23. Занесите в первую строчку слева направо середины интервалов x_k, а в первый столбец снизу верх - y_k (из статистической таблицы первого задания).

4.3.2. Введем новые переменные u_k и v_k.

Запишите значения $u_k = \frac{x_k - a_x}{h_x}$; аналогично $v_k = \frac{y_k - a_y}{h_y}$.

4.3.3. В середину каждой клетки запишите частоту совместного появления $m_{xy}=m_{uv}$, перенося значения из корреляционной решетки.

4.3.4. Выберите любую клетку поближе к середине таблицы с большой частотой $m_{xy} = m_{uv}$. В качестве a_x, ложного нуля по x, возьмите x_k, лежащее в верхней строке над выбранной клеткой с максимальной частотой, а в качестве ложного нуля по y→ a_y выбираем значение y_k в соответствующем левом столбце (см. таб.23).

4.3.5. Заполните строку $m_{u\cdot}=\sum_v m_{uv}$. То есть $m_{u\cdot}$ равно сумме всех m_{uv} в данном столбце. Найденные $m_{u\cdot}$ должны в точности совпадать с частотами m_k в статистической таблице №1 первого задания для признака X. Аналогично заполните столбец $m_{\cdot v}=\sum_u m_{uv}$, складывая все m_{uv} в данной строке. Найденные значения $m_{\cdot v}$ должны полностью совпадать с частотами m_k в таблице №1 первого задания для признака Y.

4.3.6. Сформируйте строчки значений $m_u u$ и $m_u u^2$. Аналогично столбцы $m_v v$ и $m_v v^2$.

4.3.7. Найдите значения $m_{uv}uv$:

4.3.7.1. В нижнем левом углу ячейки запишите значение произведения uv.

4.3.7.2. В верхнем правом углу запишите произведение m_{uv} на uv.

Таблица № 23.

	x_1	x_2	…	x_k a_x	…	x_{s_k}	m.v	v	$m_v v$	$m_v v^2$	$m_{uv}uv$
y_{s_k}											
…											
y_k-a_y				$m_{uv}uv$ / m_{uv} / uv							
…											
y_2											
y_1											
$m_{u\cdot}$							n		Σ_v	Σ_{v^2}	Σ_{uv}
u											
$m_u u$							Σ_u				
$m_u u^2$							Σ_{u^2}				

4.3.7.3. Заполните столбец $m_{uv}uv$, суммируя значения, полученные в каждой строке в верхних углах ячеек.

4.3.8. Найдите суммы Σ_v, Σ_{v^2}, Σ_{uv}, Σ_u и Σ_{u^2}.

4.3.9. Вычислите оценку КК

$$\hat{r}_{x,y} = \frac{\Sigma_{uv} - \frac{1}{n}\Sigma_u\Sigma_v}{\sqrt{(\Sigma_{u^2} - \frac{1}{n}(\Sigma_u)^2)(\Sigma_{v^2} - \frac{1}{n}(\Sigma_v)^2)}}$$

4.3.10. Найдите значения $\bar{x}$, s_x^2, $\bar{y}$, s_y^2 по формулам:

$$\bar{x} = h_x\frac{1}{n}\Sigma_u + a_x; \qquad \hat{\sigma}_x^2 = h_x^2\left[\frac{1}{n}\sum u^2 - \left(\frac{1}{n}\sum u\right)^2\right]$$

$$\bar{y} = h_y\frac{1}{n}\Sigma_v + a_y; \qquad \hat{\sigma}_y^2 = h_y^2\left[\frac{1}{n}\sum v^2 - \left(\frac{1}{n}\sum v\right)^2\right];$$

$\bar{x}$ и $\hat{\sigma}_x^2$; $\bar{y}$ и $\hat{\sigma}_y^2$ должны точно совпадать с соответствующими значениями, найденными в задании 1.

4.4. Вычисление доверительного интервала для $r_{x,y}$.

Для вычисления доверительных интервалов и проверки гипотезы о значении КК используют преобразование Фишера

$$z = arcth\, r = \frac{1}{2}ln\frac{1+r}{1-r} = \varphi(r), \quad \text{где } \varphi(-r) = -\varphi(r)$$

Обратное преобразование Фишера

$$r = th\, z = \frac{e^z - e^{-z}}{e^z + e^{-z}} = \psi(z), \quad \text{где} \quad \psi(-z) = -\psi(z).$$

Если выборка была сделана из двумерного нормального распределения, то z имеет распределение, близкое к нормальному распределению с математическим ожиданием $m = arcth\, r + \frac{r}{2(n-1)}$ и дисперсией $\sigma^2 = \frac{1}{(n-3)}$. Поэтому, используя данное преобразование, можно вычислять интервальные оценки для r с помощью квантилей нормального распределения.

4.4.1. Найдите преобразование Фишера от вычисленной оценки КК:

$\hat{z} = arcth\, \hat{r}$ (значение $\hat{z}$ для оценки $\hat{r}$ находится по формуле $\hat{z} = ½ \ln\frac{1+\hat{r}}{1-\hat{r}}$)

(*таблица X приложений*).

4.4.2. Вычислите ошибку $\hat{z}$:

$$\sigma_{\hat{z}} = \frac{1}{\sqrt{n-3}}.$$

4.4.3. Для трех уровней значимости α найдите границы доверительных интервалов для z:

$$\hat{z}_1 = \hat{z} - t_{1-\alpha/2}\sigma_{\hat{z}}$$
$$\hat{z}_2 = \hat{z} + t_{1-\alpha/2}\sigma_{\hat{z}}$$

4.4.4. С помощью обратного преобразования Фишера перейдите к интервалам для ***r:***

$$\hat{r}_1 = \text{th}\,\hat{z}_1,\ \hat{r}_2 = \text{th}\,\hat{z}_2.\ \text{(находятся по табл. №1X)}$$

(*таблица IX приложений*)

4.4.5. Занесите найденные значения в табл. № 24

Таблица №24. Интервальные оценки коэффициента корреляции.

α	0,1	0,05	0,01
$t_{1-\frac{\alpha}{2}}$	$t_{0,95} =$	$t_{0,975} =$	$t_{0,995} =$
n=50	$(\hat{z}_1, \hat{z}_2)$	$(\hat{z}_1, \hat{z}_2)$	$(\hat{z}_1, \hat{z}_2)$
	$(\hat{r}_1, \hat{r}_2)$	$(\hat{r}_1, \hat{r}_2)$	$(\hat{r}_1, \hat{r}_2)$

4.5. Проверка гипотезы о значении коэффициента корреляции

Проверка этой гипотезы позволяет сравнить КК с заданным значением, например, измеренным на очень большой выборке. Если же требуется сравнить два КК $\hat{r}_1$ и $\hat{r}_2$, измеренные на выборках сравнимого объема, нужно действовать аналогично, только статистика будет равна

$$t_\nu = \frac{\hat{z}_1 - \hat{z}_2}{\sqrt{\sigma^2_{\hat{z}_1} + \sigma^2_{\hat{z}_2}}}.$$

где $\nu = n_x + n_y - 2$.

Проверяем гипотезу H_0 против двусторонней альтернативы H_1

$H_0: r = r_0$

$H_1: r \neq r_0$ *(2Q)*

r_0 указано в задании.

4.5.1. Вычислите статистику $t = (\hat{z} - z_0)/\sigma_{\hat{z}}$, где $z_0 = arcth\ r_0$.

4.5.2. Найдите в таблице критические значения $t_{1-\alpha/2}$ для трех α.

4.5.3. Сравните t и $t_{1-\alpha/2}$: H_1 принимается, если $|t| \geq t_{1-\alpha/2}$. Запишите вывод.

Таблица № 25 (n=50).

t / α	0,1	0,05	0,01
$t_{1-\alpha/2}$			
t			
H:			

4.6. Проверка гипотезы о равенстве коэффициента корреляции нулю

Сравнивая КК с нулем мы можем доказать наличие зависимости между двумя характеристиками объекта: если КК достоверно отличен от 0, то зависимость есть. Однако обратное несправедливо: КК может быть равен нулю при наличии зависимости, если она не является монотонной (линейной).

Проверяем гипотезу H_0 против двусторонней альтернативы H_1:

H_0: r = 0

$H_1: r \neq 0$. *(2Q)*

4.6.1. Вычислите статистику $t_\nu = \frac{\hat{r}\sqrt{n-2}}{\sqrt{1-\hat{r}^2}}$.

4.6.2. Найдите в таблице критические значения $t_{\nu,1-\alpha/2}$ для $\nu = n-2$ и трех α.

4.6.3. Сравните t_ν и $t_{\nu,1-\alpha/2}$: H_1 принимается, если $|t_\nu| \geq t_{\nu,1-\alpha/2}$. Запишите вывод.

Таблица № 26 (n=50).

t / α	0,1	0,05	0,01
$t_{\nu,1-\alpha/2}$			
t_ν			
H:			

4.7. Прямой способ вычисления коэффициента корреляции

4.7.1. Работаем с первыми 10 парами выборочных значений исходной таблицы (x_k, y_k), где k: ($\overline{00,09}$)

4.7.2. Вычислите

$$\hat{\alpha}_{xy} = \frac{1}{n}\sum_{k=1}^{n} x_k\, y_k.$$

4.7.3. Возьмите $\bar{x}$, s_x^2 и $\bar{y}$, s_y^2 для n_x=n_y=10 из второго задания (таблица №5).

4.7.4. Вычислите

$$\hat{r}_{10} = \frac{(\hat{\alpha}_{xy} - \bar{x}\bar{y})\frac{n}{n-1}}{\sqrt{s_x^2 s_y^2}}$$

4.7.5. Постройте интервальные оценки для z $\rightarrow$ ($\hat{z}_1\hat{z}_2]$, используя $z-$преобразование Фишера, ошибку $\hat{z} \rightarrow \sigma_{\hat{z}} = \frac{1}{\sqrt{n-3}}$ и статистику $t_{1-\frac{\alpha}{2}}$ для $n = 10$ и трех уровней значимости α: 0,1; 0,05; 0,01 . По таблицам обратного преобразования Фишера (таб.1Х приложений) найдите интервальные оценки $(\hat{r}_1, \hat{r}_2]$

Полученные результаты впишите в таблицу № 27.

Таблица №27. Интервальные оценки коэффициента корреляции.

α	0,1	0,05	0,01
$t_{1-\frac{\alpha}{2}}$	$t_{0,95} =$	$t_{0,975} =$	$t_{0,995} =$
n=10	$(\hat{z}_1, \hat{z}_2)$	$(\hat{z}_1, \hat{z}_2)$	$(\hat{z}_1, \hat{z}_2)$
	$(\hat{r}_1, \hat{r}_2)$	$(\hat{r}_1, \hat{r}_2)$	$(\hat{r}_1, \hat{r}_2)$

4.7.6. Проверьте для $\hat{r}_{10}$ гипотезы о значении КК и о равенстве его нулю.

4.7.7. Запишите выводы (табл. № 28; 29), сравните с результатами для $n = 50$.

Таблица № 28 (n=10)

t α	0,1	0,05	0,01
$t_{1-\alpha/2}$			
t			
H:			

Таблица № 29 (n=10)

t ⟍ α	0,1	0,05	0,01
$t_{\nu,1-\alpha/2}$			
t_{ν}			
H:			

4.8. Коэффициент ранговой корреляции Спирмена

Коэффициент ранговой корреляции Спирмена ρ вычисляется так же, как КК Пирсона r, но вместо самих выборочных значений берутся их ранги – порядковые номера вариант, расположенных в виде возрастающего вариационного ряда. Это позволяет сделать КК при отсутствии статистической связи между случайными величинами (гипотеза H_0) независимым от их распределения. Кроме того, КК Спирмена близок по модулю к 1 не только в случае линейной, но и в случае любой монотонной нелинейной зависимости.

Таблица № 30. Пример вычисления ранговых коэффициентов корреляции.

u_x	x	y	u_y	d_i^2	$P_>$	$Q_<$
1	*523*	*519*	*6*	*25*	*4*	*5*
2	*466*	*877*	*1*	*1*	*8*	*0*
3	*438*	*806*	*2*	*1*	*7*	*0*
4	*436*	*802*	*3*	*1*	*6*	*0*
5	*409*	*235*	*9*	*16*	*1*	*4*
6	*408*	*231*	*10*	*16*	*0*	*4*
7	*379*	*661*	*4*	*9*	*3*	*0*
8	*351*	*590*	*5*	*9*	*2*	*0*
9	*294*	*448*	*7*	*4*	*1*	*0*
10	*264*	*373*	*8*	*4*	*0*	*0*
			Σ	*86*	*32*	*13*

4.8.1. Составьте и заполните таблицу, аналогичную табл. №30.

4.8.2. Выберите первые 10 пар $(n = 10)$ выборочных значений (x_i, y_i)с номерами i = 00÷ 09 из исходной выборки. Расположите их в порядке убывания x_i. Обратите внимание, что сортируется только x, а y следует за своей парой, так как у них одинаковые индексы. Ни в коем случае нельзя разрывать пары (x_i, y_i)!

4.8.3. Каждому x_i присвойте ранг U_x, а каждому y_i – ранг U_y по правилу: большему значению варианты $x_i(y_i)$ присваивается меньший ранг (у минимального значения x_i (y_i) будет ранг - n, у максимального -1). Если среди чисел есть одинаковые, им присваиваются одинаковые ранги, равные среднему арифметическому суммы их рангов.

4.8.4. Если все ранги различны, можно вычислить ρ по упрощенной формуле. Для этого запишите квадраты разностей рангов $d_i^2 = (U_x - U_y)^2$ и воспользуйтесь формулой

$$\rho = 1 - \frac{6\sum d_i^2}{n(n-1)(n+1)}$$

4.8.5. Если в обоих сопоставляемых ранговых рядах U_x и U_y присутствуют группы одинаковых рангов, то перед подсчетом коэффициента ранговой корреляции необходимо внести поправки на одинаковые ранги T_x и T_y

$T_x = \sum(u^3 - u)/12$; $T_y = \sum(v^3 - v)/12$, где u – объем каждой группы одинаковых рангов в ранговом ряду U_x , v - объем каждой группы одинаковых рангов в ранговом ряду U_y . Тогда для подсчета рангового коэффициента корреляции Спирмена используем формулу:

$$\rho = 1 - 6\frac{\sum d^2{}_i + T_x + T_y}{n(n^2-1)}.$$

Для вычисления рангового коэффициента корреляции ρ можно использовать также общую формулу, аналогичную формуле линейного коэффициента корреляции $r_{x.y}$

$$\rho = \frac{\sum U_x U_y - \frac{1}{n}\sum U_x \sum U_y}{\sqrt{(\sum U_x^2 - \frac{1}{n}(\sum U_x)^2)(\sum U_y^2 - \frac{1}{n}(\sum U_y)^2)}}$$

4.8.1.6. Проверьте гипотезу о равенстве ρ нулю:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Найдите в таблице критические значения для n =10 и трех уровней значимости α: 0,1;0,05;0,01. *Таблицы XX приложений.* Гипотеза $H_1: \rho \neq 0$ принимается, если ρ превышает критическое значение. Запишите вывод. Сравните с аналогичным выводом для КК Пирсона.

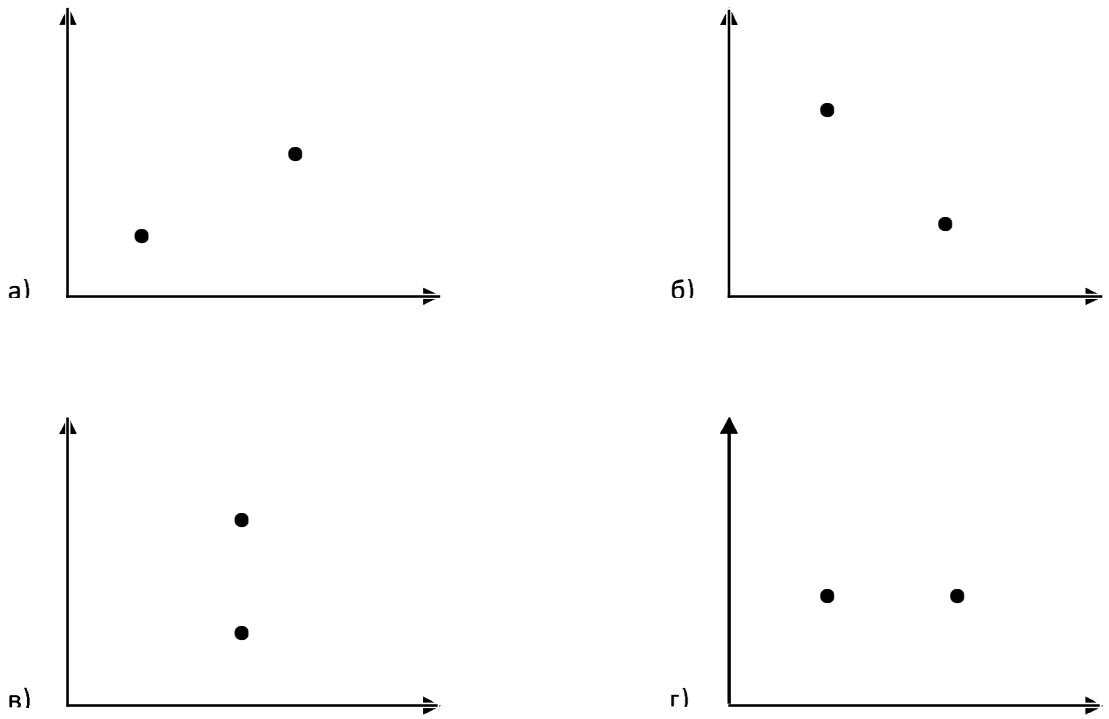


Рисунок. 7. Интерпретация коэффициента корреляции Кенделла. а) конкордантная пара точек; б) дискордантная пара точек; в), г) такие пары не считаются ни конкордантными, ни дискордантными.

4.9. Коэффициент ранговой корреляции Кенделла

Коэффициент ранговой корреляции Кенделла τ обладает теми же преимуществами, что и КК Спирмена, но также имеет простую интерпретацию. Каждую пару выборочных значений можно изобразить точкой на плоскости. В свою очередь, точки тоже можно брать парами. Всего можно составить $C_n^2 = n(n-1)/2$ пар точек.

Если все числа разные, то пары точек могут быть двух типов. Если при переходе от первой точки ко второй обе переменные меняются в одну сторону, такую пару называют конкордантной, если в разные – дискордантной (рис.7). Так вот τ равен разности между количеством конкордантных пар P и дискордантных Q, деленной на их общее количество: $\tau = \frac{P-Q}{n(n-1)/2}$. Соответственно, если все пары конкордантные, $\tau = 1$, если дискордантные, $\tau = -1$, если тех и других поровну, $\tau = 0$.

Если же среди чисел есть одинаковые, то появляются пары такие как на рис. 7 в) и г). Они не учитываются при подсчете τ. А это значит, что общее количество пар уменьшается и знаменатель формулы для τ должен быть уже другим

4.9.1. Заполните оставшиеся два столбца табл. №30.

4.9.2. Столбец $P_>$ сформируйте по правилу: последовательно просматривая строки табл. №30 сверху вниз, подсчитайте число строк, лежащих ниже первой, в которых U_x строго больше U_{x1}, а U_y строго больше U_{y1}. Затем число строк, лежащих ниже второй, в которых U_x строго больше U_{x2}, а U_y строго больше U_{y2} и т. д. вплоть до последней строки.

4.9.3. Столбец $Q_<$ сформируйте по аналогичному правилу, только теперь считайте строки, в которых U_x строго больше U_{xi}, а U_y строго меньше U_{yi}.

4.9.4. Вычислите суммы значений P и Q по обоим столбцам.

4.9.5 Если все ранги разные, вычислите τ по формуле $\tau = \frac{2(P-Q)}{n(n-1)}$. Если все сосчитано правильно, то P+Q должно быть равно n(n-1)/2.

4.8.6. Если же среди рангов есть повторяющиеся,

$$\tau = \frac{2(P-Q)}{\sqrt{[n(n-1) - \sum_i t_{xi}(t_{xi}-1)][n(n-1) - \sum_i t_{yi}(t_{yi}-1)]}},$$

где t_{xi} – это количество одинаковых чисел для i-го повторяющегося значения x, а t_{yi} – то же самое для y.

Допустим, например, что в выборке X ранг 3,5 повторяется два раза и два раза повторяется ранг 7,5, а в выборке Y - три раза повторяется ранг 6. Это значит, что $t_{x1} = 2;\ t_{x2} = 2\ и\ t_{y1} = 3$. Тогда знаменатель формулы для τ будет равен $\sqrt{10 \cdot 9 - 2 \cdot 1 - 2 \cdot 1)(10 \cdot 9 - 3 \cdot 2)} = \sqrt{7224} \approx 84{,}99$.

4.9.7. Проверьте гипотезу о равенстве τ нулю. Найдите в таблице критические значения для $n = 10$ и трех α. Гипотеза $H_1: \tau \neq 0$ принимается, если τ превышает критическое значение. Запишите вывод, сравните с аналогичным выводом для КК Пирсона и непараметрического КК ρ Спирмена.

4.10. Построение линий регрессий

Итак, мы вычислили для двумерной случайной выборки (x_k, y_k), где $k: 00 \div 49$ коэффициенты корреляции $\hat{r}_{x,y}$ для большого объема выборки $(n_x = n_y = 50)$ и для малого объема $(n_x = n_y = 10)$. Полученные коэффициенты корреляции являются точечными оценками неизвестного параметра генеральной совокупности r_{xy}. Для него построили интервальные оценки $P(r_{x,y} \in (\hat{r}_1, \hat{r}_2]) = 1 - \alpha$ и проверили гипотезы о значении коэффициента корреляции, которые указывают на наличие и силу линейной зависимости между переменными X и Y. Остается построить две линии регрессии, чтобы наглядно увидеть характер линейной зависимости между двумя изучаемыми признаками X и Y. Сделаем это для объема двумерной выборки n=50 с помощью точечной оценки $\hat{r}_{x,y}$.

Известно, что линейный характер зависимости между X и Y определяется уравнением прямой $y = kx + b$, где k – угловой коэффициент прямой к оси абсцисс, а b – отрезок, отсекаемый на оси ординат. Так как переменные X и Y можно менять местами, то есть $y = f(x)$, где x – аргумент, а y – функция, или $x = f(y)$, где y – аргумент, а x – функция, то существуют две прямые линии, которые определяют линейный характер зависимости между переменными X и Y и называются линиями регрессии. Таким образом, для каждого коэффициента корреляции $\hat{r}_{x,y}$, можно построить две линии регрессии.

Уравнение линии регрессии находится по методу наименьших квадратов и имеет вид: $y = k_1(x - \bar{x}) + \bar{y}$ для $\hat{r}_{y,x}$ и $x = k_2(y - \bar{y}) + \bar{x}$ для $\hat{r}_{x,y}$, где $k_1 = \hat{r}_{y,x}s_y/s_x$; $k_2 = \hat{r}_{x,y}s_x/s_y$ –коэффициенты регрессии Y по X и X по Y соответственно.

Из этих уравнений видно, что линии регрессии проходят через одну и ту же точку $(\bar{x}, \bar{y})$, если уравнение представить в виде:

$$y - \bar{y} = k_1(x - \bar{x})$$

$$x - \bar{x} = k_2(y - \bar{y})$$

(уравнение прямой, проходящей через заданную точку $(\bar{x}, \bar{y})$ с заданным угловым коэффициентом).

Приведем уравнения к виду:

$y = k_1x - k_1\bar{x} + \bar{y}$; $\quad x = k_2y - k_2\bar{y} + \bar{x}$, где

$b_1 = -k_1\bar{x} + \bar{y}$ и $b_2 = -k_2\bar{y} + \bar{x}$, *то есть получим две линии регрессии в виде* $y = k_1x + b_1$ *и* $x = k_2y + b_2$.

Поведение линий регрессии легко увидеть, если на корреляционной решетке найти усредненную точку с координатами $(\bar{x}, \bar{y})$; на оси ординат отложить отрезок b_1, а на оси абсцисс b_2.

Через точки полученные таким образом, проведем две линии регрессии: первую через точку b_1 на оси ординат и усредненную точку $(\bar{x}, \bar{y})$, вторую - через точку b_1 на оси абсцисс и точку $(\bar{y}, \bar{x})$. На корреляционной решетке $(\bar{x}, \bar{y})$ и $(\bar{y}, \bar{x})$ одна и та же точка.

Усредненная точка $(\bar{x}, \bar{y})$ является центром эллипса рассеивания, около которой группируются остальные точки с координатами (x_k, y_k), k: 00÷49.

При стремлении $|\hat{r}_{xy}|$ к единице эллипс рассеивания вытягивается вдоль диагонали корреляционной решетки, в зависимости от знака коэффициента корреляции, а угол между линиями регрессии стремится к нулю. При стремлении $|\hat{r}_{xy}|$ к нулю, эллипс рассеивания стремится к окружности, внутри которой находится основная масса случайных точек с координатами (x_k, y_k), угол между линиями регрессии стремится к $\pi/2$, а сами линии регрессии стремятся занять положение параллельное осям координат.

Расположение линий регрессий полностью отражает распределение частот $m_{uv} = m_{xy}$ внутри корреляционной решетки, в которой проходят линии регрессии.

Примечание: Каждый студент получает индивидуальное задание в виде двумерной выборки объема n=50, представленное выборочными значениями, как показано в Приложениях.

Приложения.

Пример двумерной выборки для разных типов объектов

LYMPH, 7

m_0 = 83.2; 18.1

σ_0^2= 908.1; 0.9007

r_o = 0.353

FFACT	AREA
18.720	155.18
17.884	135.94
17.718	149.80
17.634	100.98
18.045	129.39
17.318	130.47
17.383	117.29
18.297	137.11
18.665	139.45
17.366	159.77
18.568	112.89
18.092	114.55
17.989	111.13
17.453	126.37
17.453	154.30
17.157	102.93
18.399	107.52
17.120	144.14
18.808	110.94
17.664	153.42
17.327	112.50
18.119	93.36
19.030	131.45
18.855	119.73
17.800	120.70
17.750	149.61
17.625	142.29
18.682	111.23
17.043	104.88
18.521	129.59
19.411	113.57
18.154	139.55
16.971	166.02
19.191	109.08
17.911	116.21
17.738	114.55
19.156	110.64
17.163	153.91
17.638	144.73
17.957	130.86
16.978	129.10
17.716	59.18
17.184	157.42
18.129	115.92
17.593	103.12
17.641	137.99
18.430	120.80
18.525	124.80
18.228	118.85
18.259	103.42

SEGM, 7

m_0 = 130; 18.5

σ_0^2= 882.5; 1.221

r_o = 0.295

FFACT	FFACT
18.851	18.851
17.687	17.687
17.970	17.970
17.750	17.750
17.683	17.683
18.346	18.346
17.934	17.934
18.747	18.747
19.312	19.312
18.988	18.988
18.104	18.104
18.312	18.312
17.910	17.910
18.025	18.025
18.705	18.705
18.129	18.129
17.094	17.094
18.751	18.751
18.384	18.384
17.731	17.731
18.159	18.159
18.364	18.364
18.642	18.642
18.908	18.908
17.062	17.062
18.529	18.529
18.563	18.563
17.455	17.455
18.382	18.382
18.457	18.457
18.008	18.008
18.074	18.074
18.344	18.344
18.206	18.206
18.080	18.080
17.872	17.872
17.980	17.980
18.451	18.451
18.683	18.683
18.840	18.840
18.087	18.087
18.025	18.025
18.851	18.851
18.322	18.322
18.121	18.121
17.852	17.852
17.895	17.895
18.225	18.225
18.156	18.156
17.237	17.237

Таблица I. Таблица случайных чисел

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1534	7106	2836	7873	5574	7545	7590	5574	1202	7712
6128	8993	4102	2551	0330	2358	6427	7067	9325	2454
6047	8566	8644	9343	9297	6751	3500	8754	2913	1258
0806	5201	5705	7355	1448	9562	7514	9205	0402	2427
9915	8274	4525	5695	5752	9630	7172	6988	0227	4264
2882	7158	4341	3463	1178	5786	1173	0670	0820	5067
9213	1223	4388	9760	6691	6861	8214	8813	0611	3131
8410	9836	3899	3683	1253	1683	6988	9978	8026	6751
9974	2362	2103	4326	3825	9079	6187	2721	1489	4216
3402	8162	8226	0782	3364	7871	4500	5598	9421	3816
8188	6596	1492	2139	8823	6878	0613	7161	0241	3834
3825	7020	1124	7483	9155	4919	3209	5959	2364	2555
9801	8788	6338	5899	3309	0807	0968	0539	4205	8257
5603	1251	6352	6467	0231	3556	2569	9446	4174	9219
0714	3757	0378	8266	8864	1374	6687	1221	0678	3714
4617	5652	7627	0372	8151	3668	1994	4402	2124	0016
6789	6279	7306	1856	7028	9043	7161	7526	6913	6396
6705	4978	8621	1790	4433	6298	0854	9127	3445	1111
3840	1086	0774	9241	9297	4239	1739	7734	0119	2436
7662	3939	2965	3273	0551	1645	8477	1877	5327	8629
7639	2868	4391	2950	7122	7325	9727	0080	7464	7947
3237	7203	4246	7329	7936	0065	4146	0866	4916	8648
3917	6271	1721	5469	1914	8653	0387	2756	6073	8984
9138	9395	6005	6423	7977	1873	7103	4267	9316	7206
8358	5896	6286	9242	5040	8509	2941	3913	3028	1563
1030	5094	1745	2975	2018	7340	6547	0207	5587	0300
6606	6305	1564	6668	7822	7142	6564	1659	5369	1659
4533	8841	4922	9365	1361	6692	1633	6764	0747	3881
4258	2012	0992	0106	1542	4760	0392	4057	0092	5203
5224	5128	8949	7928	7267	0116	1476	2009	1772	3860
6872	7492	7962	1867	7437	1526	3516	9129	4153	8084
8638	8407	7198	0956	0950	7753	5144	3914	5596	6104
9958	7172	5822	4224	6701	7559	4985	4856	4461	6147
0265	3086	2996	0699	3584	9702	1665	0446	9107	6437
8987	5441	7878	9404	0487	2939	3805	9172	7887	5197
5552	3529	9627	9362	6298	6021	0024	9520	9154	0643
9383	6640	7394	9592	9903	7699	8939	9972	1257	0994
9903	4059	0332	9109	0182	6721	9163	9008	2542	4461
6530	5070	7589	6928	6014	1832	9307	5107	1354	9257
8679	8953	8310	2060	6277	1773	7979	6741	6033	3588
5765	4987	1639	3512	9843	5286	3786	2384	4919	5611
7198	2447	6716	0291	5585	1106	5330	0504	6346	3679
2385	0605	2678	1399	2371	7968	1212	9569	8650	5841
0732	8732	8660	5836	9065	4603	0029	8042	0159	0345
1642	6094	3795	2600	4532	9740	0376	4384	9203	5387
4514	1956	7212	0687	7632	2106	0846	7055	4106	9157
8744	5580	8038	9087	7222	0424	0028	4511	3191	9846
3729	6225	5397	6790	2157	3414	6509	5204	4779	5641
8858	3147	8410	2783	1290	9796	8873	7585	7185	4726
3522	5601	6197	6051	3470	8283	5702	0103	8726	5282

Продолжение *таблицы 1*

(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
5489	5583	3156	0835	1988	3912	0938	7460	0869	4420
3522	0935	7877	5665	7020	9255	7379	7124	7878	5544
7555	7579	2550	2487	9477	0864	2349	1012	8250	2633
5759	3554	5080	9074	7001	6249	3224	6368	9102	2672
6303	6895	3371	3196	7231	2918	7380	0438	7547	2644
7351	5634	5323	2623	7803	8374	2191	0464	0696	9529
7068	7803	8832	5119	6350	0120	5026	3684	5657	0304
3613	1428	1796	8447	0503	5654	3254	7336	9536	1944
5143	4534	2105	0368	7890	2473	4240	8652	9435	1422
9815	5141	7649	8638	6137	8070	5345	4865	2456	5708
5780	1277	6316	1013	2867	9938	3930	3203	5696	1769
1187	0951	5991	5245	5700	5564	7352	0891	6249	6568
4184	2179	4554	9083	2254	2435	2965	5154	1209	7069
2916	2972	9885	0275	0144	8034	8122	3213	7666	0230
5524	1341	9860	6565	6981	9842	0171	2284	2707	3008
0146	5291	2354	5694	0377	5336	6460	9585	3415	2358
4920	2826	5238	5402	7937	1993	4332	2327	6875	5230
7978	1947	6380	3425	7267	7285	1130	7722	0164	8573
7453	0653	3645	7497	5969	8682	4191	2976	0361	9334
1473	6938	4899	5348	1641	3652	0852	5296	4538	4456
8162	8797	8000	4707	1880	9660	8446	1883	9768	0881
5642	4219	0807	3301	4279	4168	4305	9937	3120	5547
2042	1192	1175	8851	6432	4635	5757	6656	1660	5389
5470	7702	6958	9080	5925	8519	0127	9233	2452	7341
4045	1730	6005	1704	0345	3275	4738	4862	2556	8333
5880	1257	6163	4439	7276	6353	6912	0731	9033	5294
9083	4260	5277	4998	4298	5204	3965	4028	8936	5148
1762	8713	1189	1090	8989	7273	3213	1935	9321	4820
2023	2589	1740	0424	8924	0005	1969	1636	7237	1227
7965	3855	4765	0703	1678	0841	7543	0308	9732	1289
7690	0480	8098	9629	4819	7219	7241	5128	3853	1921
9292	0426	9573	4903	5916	6576	8368	3270	6641	0033
0867	1656	7016	4220	2533	6345	8227	1904	5138	2537
0505	2127	8255	5276	2233	3956	4118	8199	6380	6340
6295	9795	1112	5761	2575	6837	3336	9322	7403	8345
6323	2615	3410	3365	1117	2417	3176	2434	5240	5455
8672	8536	2966	5773	5412	8114	0930	4697	6919	4569
1422	5507	7596	0670	3013	1351	3886	3268	9469	2584
2653	1472	5113	5735	1469	9545	9331	5303	9914	6394
0438	4376	3328	8649	8327	0110	4549	7955	5275	2890
2851	2157	0047	7085	1129	0460	6821	8373	2572	8962
7962	2753	3077	8718	7418	8004	3425	3706	8822	1494
3837	4098	0220	1217	4732	0150	1637	1097	1040	7372
8542	4126	9274	2251	0607	4301	8730	7690	6235	3477
0139	0765	8039	9484	2577	7859	1976	0623	1418	6685
6687	1943	4307	0579	8171	8224	8641	7034	3595	3875
6242	5582	5872	3197	4919	2792	5991	4058	9769	1918
6859	9606	0522	4993	0345	8958	1289	8825	6941	7685
6590	1932	6043	3623	1973	4112	1795	8465	2110	8045
3482	0478	0221	6738	7323	5643	4767	0106	2372	9862

Таблица II. Значения функции $F(x)$.

x	0	1	2	3	4	5	6	7	8	9
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84850	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92786	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99998	0,99999	0,99999	0,99999	—	—	—	—	—

Таблица III. Значение t при разных уровнях значимости и данном числе степеней свободы ν.

ν	Уровни значимости для одностороннего критерия (α) (Q)							
	0,1	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
	Уровни значимости для двустороннего критерия (2α) ($2Q$)							
	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
1	3,08	6,31	12,71	31,82	63,66	127,32	318,30	636,61
2	1,89	2,92	4,30	6,96	9,92	14,09	22,33	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	10,21	12,92
4	1,53	2,13	2,78	3,75	4,60	5,60	7,17	8,61
5	1,48	2,02	2,57	3,36	4,03	4,77	5,89	6,87
6	1,44	1,94	2,45	3,14	3,71	4,32	5,21	5,96
7	1,41	1,89	2,36	3,00	3,50	4,03	4,79	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	4,50	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,30	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,14	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,02	4,44
12	1,36	1,78	2,18	2,68	3,05	3,43	3,93	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	3,85	4,22
14	1,34	1,76	2,14	2,62	2,98	3,33	3,79	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	3,73	4,07
16	1,34	1,75	2,12	2,58	2,92	3,25	3,69	4,02
17	1,33	1,74	2,11	2,57	2,90	3,22	3,65	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,61	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,58	3,88
20	1,33	1,72	2,09	2,53	2,85	3,15	3,55	3,85
21	1,32	1,72	2,08	2,52	2,83	3,14	3,53	3,82
22	1,32	1,72	2,07	2,51	2,82	3,12	3,51	3,79
23	1,32	1,71	2,07	2,50	2,81	3,10	3,48	3,77
24	1,32	1,71	2,06	2,49	2,80	3,09	3,47	3,75
25	1,32	1,71	2,06	2,49	2,79	3,08	3,45	3,73
26	1,32	1,71	2,06	2,48	2,78	3,07	3,44	3,71
27	1,31	1,70	2,05	2,47	2,77	3,06	3,42	3,69
28	1,31	1,70	2,05	2,47	2,76	3,05	3,41	3,67
29	1,31	1,70	2,05	2,46	2,76	3,04	3,40	3,66
30	1,31	1,70	2,04	2,46	2,75	3,03	3,39	3,65
40	1,30	1,68	2,02	2,42	2,70	2,97	3,31	3,55
60	1,30	1,67	2,00	2,39	2,66	2,91	3,23	3,46
120	1,29	1,66	1,98	2,36	2,62	2,86	3,16	3,37
∞	1,28	1,64	1,96	2,33	2,58	2,81	3,09	3,29

Таблица IV. Значение F при $\alpha = 0{,}05$

ν_2	ν_1									
	1	2	3	4	5	6	8	12	24	□
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2 36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,52
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,02	1,83	1,61	1,25
∞	3,84	2,90	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

Таблица V. Значения F при $\alpha = 0{,}025$

ν_2	ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	647,79	799,50	864,16	899,58	921,85	937,11	956,66	976,71	997,25	1018,3
2	38,51	39,00	39,16	39,25	39,30	39,33	39,37	39,42	39,46	39,50
3	17,43	16,04	15,44	15,10	14,88	14,74	14,54	14,34	14,12	13,90
4	12,22	10,65	9,98	9,60	9,36	9,20	8,98	8,75	8,51	8,26
5	10,01	8,43	7,76	7,39	7,15	6,98	6,76	6,52	6,28	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,60	5,37	5,12	4,85
7	8,07	6,54	5,89	5,52	5,28	5,12	4,90	4,67	4,42	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,43	4,20	3,95	3,67
9	7,21	5,72	5,08	4,72	4,48	4,32	4,10	3,87	3,61	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,86	3,62	3,36	3,08
11	6,72	5,26	4,64	4,28	4,04	3,88	3,66	3,43	3,17	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,51	3,28	3,02	2,72
13	6,41	4,96	4,35	4,00	3,77	3,60	3,39	3,15	2,89	2,60
14	6,30	4,86	4,24	3,89	3,68	3,50	3,28	3,05	2,79	2,49
15	6,20	4,76	4,15	3,80	3,58	3,42	3,20	2,96	2,70	2,40
16	6,12	4,69	4,08	3,73	3,50	3,34	3,12	2,89	2,62	2,32
17	6,04	4,62	4,01	3,66	3,44	3,28	3,06	2,82	2,56	2,25
18	5,98	4,56	3,95	3,61	3,38	3,22	3,00	2,77	2,50	2,19
19	5,92	4,51	3,90	3,56	3,33	3,17	2,96	2,72	2,45	2,13
20	5,87	4,46	3,86	3,52	3,29	3,13	2,91	2,68	2,41	2,08
21	5,83	4,42	3,82	3,48	3,25	3,09	2,87	2,64	2,37	2,04
22	5,79	4,38	3,78	3,44	3,22	3,06	2,84	2,60	2,33	2,00
23	5,75	4,35	3,75	3,41	3,18	3,02	2,81	2,57	2,30	1,97
24	5,72	4,32	3,72	3,38	3,16	3,00	2,78	2,54	2,27	1,94
25	5,69	4,29	3,69	3,35	3,13	2,97	2,75	2,52	2,24	1,91
26	5,66	4,27	3,67	3,33	3,10	2,94	2,73	2,49	2,22	1,88
27	5,63	4,24	3,65	3,31	3,08	2,92	2,71	2,47	2,20	1,85
28	5,61	4,22	3,63	3,29	3,06	2,90	2,69	2,45	2,17	1,83
29	5,59	4,20	3,61	3,27	3,04	2,88	2,67	2,43	2,15	1,81
30	5,57	4,18	3,59	3,25	3,03	2,87	2,65	2,41	2,14	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,53	2,29	2,01	1,64
60	5,29	3,92	3,34	3,01	2,79	2,63	2,41	2,17	1,88	1,48
120	5,15	3,80	3,23	2,89	2,67	2,52	2,30	2,06	1,76	1,31
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,19	1,94	1,64	1,00

Таблица VI. Значения F при $\alpha = 0{,}01$

ν_2	ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50
3	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46
5	16,26	13,27	12,05	11,39	10,97	10,67	10,29	9,89	9,47	9,02
6	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,67	5,28	4,86
9	10,56	8,02	6,99	6,42	6,05	5,80	5,47	5,11	4,73	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91
11	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36
13	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16
14	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65
18	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,00	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,85	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36
22	7,94	5,72	4,82	4,31	3,99	3,76	3,45	3,12	2,75	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21
25	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	2,99	2,66	2,29	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,66	2,34	1,95	1,38
∞	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,00

Таблица VII. Значения F при $\alpha = 0{,}005$

ν_2	ν_1									
	1	2	3	4	5	6	8	12	24	∞
1	16211	20 000	21 615	22 500	23 056	23 437	23 925	24 426	24 940	25 465
2	198,50	199,00	199,17	199,25	199,30	199,33	199,37	199,42	199,46	199,51
3	55,55	49,80	47,47	46,20	45,39	44,84	44,13	43,39	42,62	41,83
4	31,33	26,28	24,26	23,16	22,46	21,98	21,35	20,70	20,03	19,32
5	22,78	18,31	16,53	15,56	14,94	14,51	13,96	13,38	12,78	12,14
6	18,64	14,54	12,92	12,03	11,46	11,07	10,57	10,03	9,47	8,88
7	16,24	12,40	10,88	10,05	9,52	9,16	8,68	8,18	7,64	7,08
8	14,69	11,04	9,60	8,80	8,30	7,95	7,50	7,02	6,50	5,95
9	13,61	10,11	8,72	7,96	7,47	7,13	6,69	6,23	5,73	5,19
10	12,83	9,43	8,08	7,34	6,87	6,54	6,12	5,66	5,17	4,64
11	12,23	8,91	7,60	6,88	6,42	6,10	5,68	5,24	4,76	4,23
12	11,75	8,51	7,23	6,52	6,07	5,76	5,34	4,91	4,43	3,90
13	11,37	8,19	6,93	6,23	5,79	5,48	5,08	4,64	4,17	3,65
14	11,06	7,92	6,68	6,00	5,56	5,26	4,86	4,43	3,96	3,44
15	10,80	7,70	6,48	5,80	5,37	5,07	4,67	4,25	3,79	3,26
16	10,58	7,51	6,30	5,64	5,21	4,91	4,52	4,10	3,64	3,11
17	10,38	7,35	6,16	5,50	5,08	4,78	4,39	3,97	3,51	2,98
18	10,22	7,22	6,03	5,38	4,96	4,66	4,28	3,86	3,40	2,87
19	10,07	7,09	5,92	5,27	4,85	4,56	4,18	3,76	3,31	2,78
20	9,94	6,99	5,82	5,17	4,76	4,47	4,09	3,68	3,22	2,69
21	9,83	6,89	5,73	5,09	4,68	4,39	4,01	3,60	3,15	2,61
22	9,73	6,81	5,65	5,02	4,61	4,32	3,94	3,54	3,08	2,55
23	9,64	6,73	5,58	4,95	4,54	4,26	3,88	3,48	3,02	2,48
24	9,55	6,66	5,52	4,89	4,49	4,20	3,83	3,42	2,97	2,43
25	9,48	6,60	5,46	4,84	4,43	4,15	3,78	3,37	2,92	2,38
26	9,41	6,54	5,41	4,78	4,38	4,10	3,73	3,32	2,87	2,33
27	9,34	6,49	5,36	4,74	4,34	4,06	3,69	3,28	2,83	2,29
28	9,28	6,44	5,32	4,70	4,30	4,02	3,65	3,25	2,79	2,24
29	9,23	6,40	5,28	4,66	4,26	3,98	3,61	3,21	2,76	2,21
30	9,18	6,36	5,24	4,62	4,23	3,95	3,58	3,18	2,73	2,18
40	8,83	6,07	4,98	4,37	3,99	3,71	3,35	2,95	2,50	1,93
60	8,50	5,80	4,73	4,14	3,76	3,49	3,13	2,74	2,29	1,69
120	8,18	5,54	4,50	3,92	3,55	3,28	2,93	2,54	2,09	1,43
∞	7,88	5,30	4,28	3,72	3,35	3,09	2,74	2,36	1,90	1,00

Таблица VIII. Критерий согласия Колмогорова. Значения $\alpha = 1-K(\lambda)$

λ	0	1	2	3	4	5	6	7	8	9
0,3	0,99999	0,99998	0,99995	0,99991	0,99983	0,99970	0,99949	0,99917	0,99872	0,99807
0,4	0,99719	0,99603	0,99452	0,99262	0,99027	0,98741	0,98400	0,97998	0,97532	0,96998
0,5	0,96394	0,95719	0,94969	0,94147	0,93250	0,92282	0,91242	0,90134	0,88960	0,87724
0,6	0,86428	0,85077	0,83678	0,82225	0,80732	0,79201	0,77636	0,76042	0,74422	0,72781
0,7	0,71124	0,69453	0,67774	0,66089	0,64402	0,62717	0,61036	0,59363	0,57700	0,56050
0,8	0,54414	0,52796	0,51197	0,49619	0,48063	0,46532	0,45026	0,43545	0,42093	0,40668
0,9	0,39273	0,37907	0,36571	0,35266	0,33992	0,32748	0,31536	0,30356	0,29206	0,28087
1,0	0,27000	0,25943	0,24917	0,23922	0,22957	0,22021	0,21114	0,20236	0,19387	0,18566
1,1	0,17772	0,17005	0,16264	0,15550	0,14861	0,14196	0,13556	0,12939	0,12345	0,11774
1,2	0,11225	0,10697	0,10190	0,09703	0,09235	0,08787	0,08357	0,07944	0,07550	0,07171
1,3	0,06809	0,06463	0,06132	0,05815	0,05513	0,05224	0,04949	0,04686	0,04435	0,04196
1,4	0,03968	0,03751	0,03545	0,03348	0,03162	0,02984	0,02815	0,02655	0,02503	0,02359
1,5	0,02222	0,02092	0,01969	0,01852	0,01742	0,01638	0,01539	0,01446	0,01357	0,01274
1,6	0,01195	0,01121	0,01051	0,00985	0,00922	0,00864	0,00808	0,00756	0,00707	0,00661
1,7	0,00618	0,00577	0,00539	0,00503	0,00469	0,00438	0,00408	0,00380	0,00354	0,00330
1,8	0,00307	0,00285	0,00265	0,00247	0,00229	0,00213	0,00198	0,00186	0,00170	0,00158
1,9	0,00146	0,00136	0,00126	0,00116	0,00108	0,00100	0,00092	0,00085	0,00079	0,00073
2,0	0,00067	0,00062	0,00057	0,00053	0,00048	0,00045	0,00041	0,00038	0,00035	0,00032
2,1	0,00030	0,00027	0,00025	0,00023	0,00021	0,00019	0,00018	0,00016	0,00015	0,00014
2,2	0,00013	0,00011	0,00010	0,00010	0,00009	0,00008	0,00007	0,00007	0,00006	0,00006
2,3	0,00005	0,00005	0,00004	0,00004	0,00004	0,00003	0,00003	0,00003	0,00002	0,00002
2,4	0,00002	0,00002	0,00002	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001	0,00001

Таблица IX. Значения величины ***r*** для значений ***z*** от 0,00 до 2,99

z	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0500	0,0599	0,0699	0,0798	0,0898
0,1	0,0997	0,1096	0,1194	0,1293	0,1391	0,1489	0,1586	0,1684	0,1781	0,1877
0,2	0,1974	0,2070	0,2165	0,2260	0,2355	0,2449	0,2543	0,2636	0,2729	0,2821
0,3	0,2913	0,3004	0,3095	0,3185	0,3275	0,3364	0,3452	0,3540	0,3627	0,3714
0,4	0,3800	0,3885	0,3969	0,4053	0,4136	0,4219	0,4301	0,4382	0,4462	0,4542
0,5	0,4621	0,4699	0,4777	0,4854	0,4930	0,5005	0,5080	0,5154	0,5227	0,5299
0,6	0,5370	0,5441	0,5511	0,5580	0,5649	0,5717	0,5784	0,5850	0,5915	0,5980
0,7	0,6044	0,6107	0,6169	0,6231	0,6291	0,6351	0,6411	0,6469	0,6527	0,6584
0,8	0,6640	0,6696	0,6751	0,6805	0,6858	0,6911	0,6963	0,7014	0,7064	0,7114
0,9	0,7163	0,7211	0,7259	0,7306	0,7352	0,7398	0,7443	0,7487	0,7531	0,7574
1,0	0,7616	0,7658	0,7699	0,7739	0,7779	0,7818	0,7857	0,7895	0,7932	0,7969
1,1	0,8005	0,8041	0,8076	0,8110	0,8144	0,8178	0,8210	0,8243	0,8275	0,8306
1,2	0,8337	0,8367	0,8397	0,8426	0,8455	0,8483	0,8511	0,8538	0,8565	0,8591
1,3	0,8617	0,8643	0,8668	0,8692	0,8717	0,8741	0,8764	0,8787	0,8810	0,8832
1,4	0,8854	0,8875	0,8896	0,8917	0,8937	0,8957	0,8977	0,8996	0,9015	0,9033
1,5	0,9051	0,9069	0,9087	0,9104	0,9121	0,9138	0,9154	0,9170	0,9186	0,9201
1,6	0,9217	0,9232	0,9246	0,9261	0,9275	0,9289	0,9302	0,9316	0,9329	0,9341
1,7	0,9354	0,9366	0,9379	0,9391	0,9402	0,9414	0,9425	0,9436	0,9447	0,9458
1,8	0,9468	0,9478	0,9488	0,9498	0,9508	0,9518	0,9527	0,9536	0,9545	0,9554
1,9	0,9562	0,9571	0,9579	0,9587	0,9595	0,9603	0,9611	0,9618	0,9626	0,9633
2,0	0,9640	0,9647	0,9654	0,9661	0,9668	0,9674	0,9680	0,9686	0,9693	0,9699
2,1	0,9704	0,9710	0,9716	0,9722	0,9727	0,9732	0,9738	0,9743	0,9748	0,9753
2,2	0,9757	0,9762	0,9767	0,9771	0,9776	0,9780	0,9785	0,9789	0,9793	0,9797
2,3	0,9801	0,9805	0,9809	0,9812	0,9816	0,9820	0,9823	0,9827	0,9830	0,9834
2,4	0,9837	0,9840	0,9843	0,9846	0,9849	0,9852	0,9855	0,9858	0,9861	0,9864
2,5	0,9866	0,9869	0,9871	0,9874	0,9876	0,9879	0,9881	0,9884	0,9886	0,9888
2,6	0,9890	0,9892	0,9894	0,9897	0,9899	0,9901	0,9903	0,9904	0,9906	0,9908
2,7	0,9910	0,9912	0,9914	0,9915	0,9917	0,9919	0,9920	0,9922	0,9923	0,9925
2,8	0,9926	0,9928	0,9929	0,9931	0,9932	0,9933	0,9935	0,9936	0,9937	0,9938
2,9	0,9940	0,9941	0,9942	0,9943	0,9944	0,9945	0,9946	0,9947	0,9948	0,9949

Таблица X. Значения величины ***z*** для значений ***r*** от 0,00 до 0,99

r	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0100	0,0200	0,0300	0,0400	0,0501	0,0601	0,0701	0,0802	0,0902
0,1	0,1003	0,1105	0,1206	0,1308	0,1409	0,1511	0,1614	0,1717	0,1820	0,1923
0,2	0,2027	0,2132	0,2237	0,2342	0,2448	0,2554	0,2661	0,2769	0,2877	0,2986
0,3	0,3095	0,3206	0,3317	0,3428	0,3541	0,3654	0,3769	0,3884	0,4001	0,4118
0,4	0,4236	0,4356	0,4477	0,4599	0,4722	0,4847	0,4973	0,5101	0,5230	0,5361
0,5	0,5493	0,5627	0,5763	0,5901	0,6042	0,6184	0,6328	0,6475	0,6625	0,6777
0,6	0,6931	0,7089	0,7250	0,7414	0,7582	0,7753	0,7928	0,8107	0,8291	0,8480
0,7	0,8673	0,8872	0,9076	0,9287	0,9505	1,9730	0,9962	1,0203	1,0454	1,0714
0,8	1,0986	1,1270	1,1568	1,1881	1,2212	1,2562	1,2933	1,3331	1,3758	1,4219
0,9	1,4722	1,5275	1,5890	1,6584	1,7380	1,8318	1,9459	2,0923	2,2976	2,6467

Таблица XI. χ^2-распределение. Критические точки для разных значений вероятности α и чисел степеней свободы ν

N	α						α					
	0,1	0,05	0,025	0,01	0,005	0,001	0,999	0,995	0,990	0,975	0,950	0,9
1	2,71	3,84	5,02	6,64	7,88	10,83	—	—	0,0001	0,001	0,004	0,0158
2	4,61	5,99	7,38	9,21	10,60	13,82	0,002	0,01	0,02	0,05	0,10	0,21
3	6,25	7,82	9,35	11,34	12,84	16,27	0,02	0,07	0,12	0,22	0,35	0,58
4	7,78	9,49	11,14	13,28	14,86	18,47	0,09	0,21	0,30	0,48	0,71	1,06
5	9,24	11,07	12,83	15,09	16,75	20,52	0,21	0,41	0,55	0,83	1,14	1,61
6	10,65	12,59	14,15	16,81	18,55	22,46	0,38	0,68	0,87	1,24	1,64	2,20
7	12,02	14,07	16,01	18,48	20,28	24,32	0,60	0,99	1,24	1,69	2,17	2,83
8	13,36	15,51	17,54	20,09	21,96	26,12	0,86	1,44	1,65	2,18	2,73	3,49
9	14,68	16,92	19,02	21,67	23,59	27,88	1,15	1,54	2,09	2,70	3,32	4,17
10	15,99	18,31	20,48	23,21	25,19	29,59	1,48	2,65	2,56	3,25	3,94	4,87
11	17,28	19,68	21,92	24,72	26,76	31,26	1,83	2,31	3,05	3,82	4,58	5,58
12	18,55	21,03	23,34	26,22	28,30	32,91	2,21	3,47	3,57	4,40	5,23	6,30
13	19,81	22,36	24,74	27,69	29,82	34,53	2,62	3,56	4,11	5,01	5,89	7,04
14	21,06	23,68	26,12	29,14	31,32	36,12	3,04	4,57	4,66	5,63	6,57	7,79
15	22,31	25,00	27,49	30,58	32,80	37,70	3,48	4,11	5,23	6,26	7,26	8,55
16	23,54	26,30	28,84	32,00	34,27	39,25	3,94	5,24	5,81	6,91	7,96	9,31
17	24,77	27,59	30,19	33,41	35,72	40,79	4,42	5,80	6,41	7,56	8,67	10,09
18	25,99	28,87	31,53	34,80	37,16	42,31	4,91	6,56	7,02	8,23	9,39	10,87
19	27,20	30,14	32,85	36,19	38,18	43,82	5,41	6,45	7,63	8,91	10,12	11,65
20	28,41	31,41	34,17	37,57	40,00	45,32	5,92	7,43	8,27	9,59	10,85	12,44
21	29,62	32,67	35,48	38,93	41,40	46,80	6,45	8,43	8,90	10,28	11,59	13,24
22	30,81	33,92	36,78	40,29	42,80	48,27	6,98	8,35	9,54	10,98	12,34	14,04
23	32,01	35,17	38,08	41,64	44,18	49,73	7,53	9,06	10,20	11,69	13,09	14,85
24	33,20	36,42	39,36	42,98	45,56	51,18	8,09	9,69	10,86	12,40	13,85	15,66
25	34,38	37,65	40,65	44,31	46,93	52,62	8,65	10,02	11,52	13,12	14,61	16,47
26	35,56	38,88	41,92	45,64	48,29	54,05	9,22	11,06	12,20	13,84	15,38	17,29
27	36,74	40,11	43,19	46,95	49,64	55,48	9,80	11,81	12,88	14,57	16,15	18,11
28	37,92	41,34	44,46	48,28	50,99	56,89	10,39	12,46	13,56	15,31	16,93	18,94
29	39,09	42,56	45,72	49,59	52,34	58,30	10,99	13,12	14,25	16,05	17,71	19,77
30	40,26	43,77	46,98	50,89	53,67	59,70	11,59	13,79	14,95	16,79	18,49	20,60
31	41,42	44,93	48,23	52,19	55,00	61,10	12,20	14,46	15,66	17,54	19,28	21,43
32	42,59	46,19	49,48	53,49	56,33	62,49	12,81	15,13	16,36	18,29	20,07	22,27
33	43,75	47,40	50,72	54,78	57,65	63,87	13,43	15,82	17,07	19,05	20,88	23,11
34	44,90	48,60	51,97	56,06	58,96	65,25	14,06	16,50	17,79	19,81	21,66	23,95
35	46,06	49,80	53,20	57,34	60,28	66,62	14,69	17,19	18,51	20,57	22,46	24,80
36	47,21	51,00	54,44	58,62	61,58	67,98	15,32	17,89	19,23	21,34	23,27	25,64
37	48,36	52,19	55,67	59,89	62,88	69,35	15,94	18,59	19,96	22,11	24,08	26,49
38	49,51	53,38	56,90	61,18	64,18	70,70	16,61	19,29	20,69	22,88	24,88	27,34
39	50,66	54,57	58,12	62,43	65,48	72,06	17,26	20,00	21,43	23,65	25,70	28,20
40	51,81	55,76	59,34	63,69	66,77	73,40	17,92	20,71	22,16	24,43	26,51	29,05
41	52,95	56,94	60,56	64,95	68,05	74,74	18,58	21,42	22,91	25,22	27,33	29,01
42	54,09	58,12	61,78	66,21	69,34	76,08	19,24	22,14	23,65	26,00	28,14	30,77
43	55,23	59,30	62,99	67,46	70,62	77,42	19,91	22,86	24,40	26,78	28,97	31,63
44	56,37	60,48	64,20	68,71	71,89	78,75	20,58	23,58	25,15	27,58	29,79	32,49
45	57,51	61,66	65,41	69,96	73,17	80,08	21,25	24,31	25,90	28,37	30,61	33,35
46	58,64	62,83	66,62	71,20	74,44	81,40	21,93	25,04	26,66	29,16	31,44	34,22
47	59,77	64,00	67,82	72,44	75,70	82,72	22,61	25,78	27,42	29,96	32,27	35,08
48	60,91	65,17	69,02	73,68	76,97	84,04	23,30	26,51	28,18	30,76	33,10	35,95
49	62,04	66,34	70,22	74,92	78,23	85,35	23,98	27,25	28,94	31,56	33,93	36,82
50	63,17	67,51	71,42	76,15	79,49	86,66	24,67	27,99	29,71	32,36	34,76	37,69
51	64,30	68,67	72,62	77,39	80,75	87,97	25,37	28,74	30,48	53,16	35,60	38,56

Продолжение *таблицы XI*

N	α						α					
	0,1	0,05	0,025	0,01	0,005	0,001	0,999	0,995	0,990	0,975	0,950	0,9
52	65,42	69,83	73,81	78,62	82,00	89,27	26,06	29,48	31,25	33,97	36,44	39,43
53	66,55	70,99	75,00	79,84	83,25	90,57	26,76	30,23	32,02	34,78	37,28	40,31
54	67,67	72,15	76,19	81,07	84,50	91,87	27,47	30,98	32,79	35,59	38,12	41,18
55	68,80	73,31	77,38	82,29	85,75	93,17	28,17	31,74	33,57	36,40	38,96	42,06
56	69,92	74,47	78,57	83,51	86,99	94,46	28,88	32,49	34,35	37,21	39,80	42,94
57	71,04	75,62	79,75	84,73	88,24	95,75	29,59	33,25	35,13	38,03	40,65	43,82
58	72,16	76,78	80,94	85,95	89,48	97,04	30,30	34,01	35,91	38,84	41,49	44,70
59	73,28	77,93	82,12	87,17	90,72	98,32	31,02	34,77	36,70	39,66	42,34	45,58
60	74,40	79,08	83,30	88,38	91,95	99,61	31,74	35,54	37,48	40,48	43,19	46,46
61	75,51	80,23	84,48	89,59	93,19	100,89	32,46	36,30	38,27	41,30	44,04	47,34
62	76,63	81,38	85,65	90,80	94,42	102,17	33,18	37,07	39,06	42,13	44,89	48,23
63	77,75	82,53	86,83	92,01	95,65	103,44	33,91	37,84	39,86	42,95	45,74	49,11
64	78,86	83,68	88,00	93,22	96,88	104,72	34,63	38,61	40,65	43,78	46,60	50,00
65	79,97	84,82	89,18	94,42	98,11	105,99	35,36	39,38	41,44	44,60	47,45	50,88
66	81,09	85,97	90,35	95,63	99,33	107,26	36,09	40,16	42,24	45,43	48,30	51,77
67	82,20	87,11	91,52	96,83	100,6	108,53	36,83	40,94	43,04	46,26	49,16	52,66
68	83,31	88,25	92,69	98,03	101,8	109,79	37,56	41,71	43,84	47,09	50,02	53,55
69	84,42	89,39	93,86	99,23	103,0	111,06	38,30	42,49	44,64	47,92	50,88	54,44
70	85,53	90,53	95,02	100,4	104,2	112,32	39,04	43,28	45,44	48,76	51,74	55,33
71	86,64	91,67	96,19	101,6	105,4	113,58	39,78	44,06	46,25	49,59	52,60	56,22
72	87,74	92,81	97,35	102,8	106,6	114,84	40,52	44,84	47,05	50,43	53,46	57,11
73	88,85	93,94	98,52	104,0	107,8	116,09	41,26	45,63	47,86	51,26	54,32	58,01
74	89,96	95,08	99,68	105,2	109,1	117,35	42,01	46,42	48,67	52,10	55,19	58,90
75	91,06	96,22	100,8	106,4	110,3	118,60	42,76	47,21	49,48	52,94	56,05	59,80
76	92,17	97,35	102,00	107,58	111,50	119,85	43,51	48,00	50,29	53,78	56,92	60,69
77	93,27	98,48	103,16	108,77	112,70	121,10	44,26	48,79	51,10	54,62	57,79	61,59
78	94,37	99,62	104,32	109,96	113,91	122,35	45,01	49,58	51,91	55,47	58,65	62,48
79	95,48	100,75	105,47	111,14	115,12	123,59	45,76	50,38	52,72	56,31	59,52	63,38
80	96,58	101,88	106,63	112,33	116,32	124,84	46,52	51,17	53,54	57,15	60,39	64,28
81	97,68	103,01	107,78	113,51	117,52	126,08	47,28	51,97	54,36	58,00	61,26	65,18
82	98,78	104,14	108,94	114,70	118,73	127,32	48,04	52,77	55,17	58,84	62,13	66,08
83	99,88	105,27	110,09	115,88	119,93	128,56	48,80	53,57	55,99	59,69	63,00	66,98
84	100,98	106,40	111,24	117,06	121,13	129,80	49,56	54,37	56,81	60,54	63,88	67,88
85	102,08	107,52	112,39	118,24	122,32	131,04	50,32	55,17	57,63	61,39	64,75	68,78
86	103,18	108,65	113,54	119,41	123,52	132,28	51,08	55,97	58,46	62,24	65,62	69,68
87	104,28	109,77	114,69	120,59	124,72	133,51	51,85	56,78	59,28	63,09	66,50	70,58
88	105,37	110,90	115,84	121,77	125,91	134,74	52,62	57,58	60,10	63,94	67,37	71,48
89	106,47	112,02	117,00	122,94	127,11	135,98	53,39	58,39	60,93	64,79	68,25	72,39
90	107,57	113,14	118,14	124,12	128,30	137,21	54,16	59,20	61,75	65,65	69,13	73,30
91	108,66	114,27	119,28	125,29	129,49	138,44	54,93	60,00	62,58	66,50	70,00	74,20
92	109,76	115,39	120,43	126,46	130,68	139,67	55,70	60,82	63,41	67,36	70,88	75,10
93	110,85	116,51	121,57	127,63	131,87	140,89	56,47	61,62	64,24	68,21	71,76	76,01
94	111,94	117,63	122,72	128,80	133,06	142,12	57,25	62,44	65,07	69,07	72,64	76,91
95	113,04	118,75	123,86	129,97	134,25	143,34	58,02	63,25	65,90	69,92	73,52	77,82
96	114,13	119,87	125,00	131,14	135,43	144,57	58,80	64,06	66,73	70,78	74,40	78,73
97	115,22	120,99	126,14	132,31	136,62	145,79	59,58	64,88	67,56	71,64	75,28	79,63
98	116,32	122,11	127,29	133,48	137,80	147,01	60,36	65,69	68,40	72,50	76,16	80,54
99	117,41	123,22	128,42	134,64	138,99	148,23	61,14	66,51	69,23	73,36	77,05	81,45
100	118,50	124,34	129,56	135,81	140,17	149,45	61,92	67,33	70,06	74,22	77,93	82,36

Таблица XII. Критические значения ***U***-критерия Уилкоксона (Манна-Уитни) (односторонний критерий, α=*0,01*)

n_2 \ n_1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6		3	4	6	7	8	9	11	12	14	15	16	18	19	20	22
7			6	7	9	11	12	14	16	18	19	21	23	24	26	28
8				9	11	13	15	17	20	22	24	26	28	30	32	34
9					14	16	19	21	23	26	28	31	33	36	38	40
10						19	22	24	27	30	33	36	38	41	44	47
11							25	28	31	34	37	41	44	47	50	53
12								31	35	38	42	46	49	53	56	60
13									39	43	47	51	55	59	63	67
14										47	51	56	60	65	69	73
15											56	61	66	70	75	80
16												66	71	76	82	87
17													77	82	88	94
18														88	94	100
19															101	107
20																114

Продолжение *таблицы XII* (односторонний критерий, α=*0,05*)

n_2 \ n_1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
4	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6		7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7			11	13	15	17	19	21	24	26	28	30	33	35	37	39
8				15	18	20	23	26	28	31	33	36	39	41	44	47
9					21	24	27	30	33	36	39	42	45	48	51	54
10						27	31	34	37	41	44	48	51	55	58	62
11							34	38	42	46	50	54	57	61	65	69
12								42	47	51	55	60	64	68	72	77
13									51	56	61	65	70	75	80	84
14										61	66	71	77	82	87	92
15											72	77	83	88	94	100
16												83	89	95	101	107
17													96	102	109	115
18														109	116	123
19															123	130
20																138

Таблица XIII. Критические значения z-критерия знаков при разных уровнях значимости α и объеме выборки n

n	α		n	α		n	α		n	α	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
6	0	–	30	9	7	54	19	17	78	29	27
7	0	–	31	9	7	55	19	17	79	30	27
8	0	0	32	9	8	56	20	17	80	30	28
9	1	0	33	10	8	57	20	18	81	31	28
10	1	0	34	10	9	58	21	18	82	31	28
11	1	0	35	11	9	59	21	19	83	32	29
12	2	1	36	11	9	60	21	19	84	32	29
13	2	1	37	12	10	61	22	20	85	32	30
14	2	1	38	12	10	62	22	20	86	33	30
15	3	2	39	12	11	63	23	20	87	33	31
16	3	2	40	13	11	64	23	21	88	34	31
17	4	2	41	13	11	65	24	21	89	34	31
18	4	3	42	14	12	66	24	22	90	35	32
19	4	3	43	14	12	67	25	22	91	35	32
20	5	3	44	15	13	68	25	22	92	36	33
21	5	4	45	15	13	69	25	23	93	36	33
22	5	4	46	15	13	70	26	23	94	37	34
23	6	4	47	16	14	71	26	24	95	37	34
24	6	5	48	16	14	72	27	24	96	37	34
25	7	5	49	17	15	73	27	25	97	38	35
26	7	6	50	17	15	74	28	25	98	38	35
27	7	6	51	18	15	75	28	25	99	39	36
28	8	6	52	18	16	76	28	26	100	39	36
29	8	7	53	18	16	77	29	26	–	–	–

Таблица XIV. Критические значения парного T-критерия Уилкоксона (односторонний критерий)

Число парных наблюдений n	Уровни значимости α		Число парных наблюдений n	Уровни значимости α	
	0,05	0,01		0,05	0,01
5	0	–	14	25	16
6	2	0	15	30	19
7	3	0	16	35	23
8	5	1	17	41	28
9	8	3	18	47	33
10	10	5	19	53	38
11	13	7	20	60	42
12	17	10	21	67	50
13	21	12	22	74	56

Продолжение *таблицы XIV*, (двусторонний критерий)

Число парных наблюдений ***n***	Уровни значимости α		Число парных наблюдений ***n***	Уровни значимости α	
	0,05	0,01		0,05	0,01
6	1	–	16	31	21
7	3	–	17	36	24
8	5	1	18	41	29
9	7	3	19	47	33
10	9	4	20	53	39
11	12	6	21	60	44
12	15	8	22	67	50
13	18	11	23	74	56
14	22	14	24	82	62
15	26	17	25	90	69

Примечание. Для n >25 критические значения T-критерия можно определить по формуле

$$T_{st} = \frac{n(n+1)}{4} \sqrt{\frac{n(n-1)(2n-1)}{24}},$$

где n- число парных наблюдений; t зависит от принятого уровня значимости, т. е. $t_{0,05}$=1,96 и $t_{0,01}$=2,58.

Таблица XV. Критические значения коэффициента асимметрии A_s

Объем выборки ***n***	Уровни значимости α		Объем выборки ***n***	Уровни значимости α	
	0,05	0,01		0,05	0,01
25	0,711	1,061	250	0,251	0,360
30	0,661	0,982	300	0,230	0,329
35	0,621	0,921	350	0,213	0,305
40	0,587	0,869	400	0,200	0,285
45	0,558	0,825	450	0,188	0,269
50	0,533	0,787	500	0,179	0,255
60	0,492	0,723	550	0,171	0,243
70	0,459	0,673	600	0,163	0,233
80	0,432	0,631	650	0,157	0,224
90	0,409	0,596	700	0,151	0,215
100	0,389	0,567	750	0,146	0,208
125	0,350	0,508	800	0,142	0,202
150	0,321	0,464	850	0,138	0,196
175	0,298	0,430	900	0,134	0,190
200	0,280	0,403	950	0,130	0,185
			1000	0,127	0,180

Таблица XVI. Критические значения коэффициента эксцесса E_x

Объем выборки n	Уровни значимости α		
	0,10	0,05	0,01
11	0,890	0,907	0,936
16	0,873	0,888	0,914
21	0,863	0,877	0,900
26	0,857	0,869	0,890
31	0,851	0,863	0,883
36	0,847	0,858	0,877
41	0,844	0,854	0,872
46	0,841	0,851	0,868
51	0,839	0,848	0,865
61	0,835	0,843	0,859
71	0,832	0,840	0,855
81	0,830	0,838	0,852
91	0,828	0,835	0,848
101	0,826	0,834	0,846
201	0,818	0,823	0,832
301	0,814	0,818	0,826
401	0,812	0,816	0,822
501	0,810	0,814	0,820

Таблица XVII. Критические значения величины нормированного отклонения при оценке сомнительных вариант с учетом объема выборки n и уровней значимости α

n	α		n	α		n	α	
	0,05	0,01		0,05	0,01		0,05	0,01
4	1,71	1,73	13	2,56	2,81	23	2,84	3,16
5	1,92	1,97	14	2,60	2,86	24	2,86	3,18
6	2,07	2,16	15	2,64	2,90	25	2,88	3,20
7	2,18	2,31	16	2,67	2,95	26	2,90	3,22
8	2,27	2,43	17	2,70	2,98	27	2,91	3,24
9	2,35	2,53	18	2,73	3,02	28	2,93	3,26
10	2,41	2,62	19	2,75	3,05	29	2,94	3,28
11	2,47	2,69	20	2,78	3,08	30	2,96	3,29
12	2,52	2,75	21	2,80	3,11			

Таблица XVIII. Критические значения критерия $t_1 = \frac{x_2 - x_1}{x_{n-1} - x_1}$.

n	Уровни значимости α		n	Уровни значимости α		n	Уровни значимости α	
	0,05	0,01		0,05	0,01		0,05	0,01
4	0,76	0,89	13	0,36	0,46	22	0,29	0,38
5	0,64	0,78	14	0,35	0,45	23	0,28	0,37
6	0,56	0,70	15	0,34	0,44	24	0,28	0,37
7	0,31	0,64	16	0,33	0,43	25	0,28	0,36
8	0,47	0,59	17	0,32	0,42	26	0,27	0,36
9	0,44	0,54	18	0,31	0,41	27	0,27	0,35
10	0,41	0,53	19	0,31	0,40	28	0,27	0,35
11	0,39	0,50	20	0,30	0,39	29	0,26	0,34
12	0,38	0,48	21	0,30	0,38	30	0,26	0,34

Таблица XIX. Критические значения коэффициента корреляции $r_{x,y}$

Степени свободы $\nu=n-2$	Уровни значимости α		Степени свободы $\nu=n-2$	Уровни значимости α	
	0,05	0,01		0,05	0,01
5	0,75	0,87	27	0,37	0,47
6	0,71	0,83	28	0,36	0,46
7	0,67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	35	0,33	0,42
10	0,58	0,71	40	0,30	0,39
11	0,55	0,68	45	0,29	0,37
12	0,53	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,56	125	0,17	0,23
19	0,43	0,55	150	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12
24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

Таблица XX. Критические значения коэффициента корреляции рангов при различных уровнях значимости α и объемах выборки n

n	α		n	α		n	α	
	0,05	0,01		0,05	0,01		0,05	0,01
5	0,94	–	17	0,48	0,62	29	0,37	0,48
6	0,85	–	18	0,47	0,60	30	0,36	0,47
7	0,78	0,94	19	0,46	0,58	31	0,36	0,46
8	0,72	0,88	20	0,45	0,57	32	0,36	0,45
9	0,68	0,83	21	0,44	0,56	33	0,34	0,45
10	0,64	0,79	22	0,43	0,54	34	0,34	0,44
11	0,61	0,76	23	0,42	0,53	35	0,33	0,43
12	0,58	0,73	24	0,41	0,52	36	0,33	0,43
13	0,56	0,70	25	0,40	0,51	37	0,33	0,42
14	0,54	0,68	26	0,39	0,50	38	0,32	0,41
15	0,52	0,66	27	0,38	0,49	39	0,32	0,41
16	0,50	0,64	28	0,38	0,48	40	0,31	0,40

Таблица XXI. Критические значения критерия Уилкоксона для несвязных (независимых) совокупностей (левосторонний односторонний критерий)

n_1	n_2	Уровни значимости α					
		0.01	0.02	0.05	0.10	0.20	0.40
4	4			10	11	13	14
	5		10	11	12	14	15
	6	10	11	12	13	15	17
	7	10	11	13	14	16	18
	8	11	12	14	15	17	20
	9	11	13	14	16	19	21
	10	12	13	15	17	20	23
	11	12	14	16	18	21	24
	12	13	15	17	19	22	26
5	5	15	16	17	19	20	22
	6	16	17	18	20	22	24
	7	16	18	20	21	23	26
	8	17	19	21	23	25	28
	9	18	20	22	24	27	30
	10	19	21	23	26	28	32
	11	20	22	24	27	30	34
	12	21	23	26	28	32	36
6	6	23	24	26	28	30	33
	7	24	25	27	29	32	35
	8	25	27	29	31	34	37
	9	26	28	31	33	36	40
	10	27	29	32	35	38	42
	11	28	30	34	37	40	44
	12	30	32	35	38	42	47
7	7	32	34	36	39	41	45
	8	34	35	38	41	44	48
	9	35	37	40	43	46	50
	10	37	39	42	45	49	53
	11	38	40	44	47	51	56
	12	40	42	46	49	54	59
8	8	43	45	49	51	55	59
	9	45	47	51	54	58	62
	10	47	49	53	56	60	65
	11	49	51	55	59	63	69
	12	51	53	58	62	66	72
9	9	56	59	62	66	70	75
	10	58	61	65	69	73	78
	11	61	63	68	72	76	82
	12	63	66	71	75	80	86
10	10	71	74	78	82	87	93
	11	73	77	81	86	91	97
	12	76	79	84	89	94	101
11	11	87	91	96	100	106	112
	12	90	94	99	104	110	117
12	12	105	109	115	120	127	134

ЛИТЕРАТУРА

1. Плохинский Н.А. Биометрия. - 2-е изд.- М: МГУ, 1970, 368 с.
2. Митропольский А.К. Техника статистических вычислений. - 2-е изд. - М: Наука, Главная редакция физико-математической литературы, 1971, 576 с.
3. Гублер Е.В., Генкин А.А. Применение непараметрических критериев в медико-биологических исследованиях. - 2-е изд.- Ленинград: Медицина, Л, 1973, 141 с.
4. СидоренкоЕ.В. Методы математической обработки в психологии.- Санкт-Петербург: Речь, 2002, 350 с.
5. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. - Под ред. В.В. Фигурнова. - М: Финансы и статистика, 1995, 384 с.
6. Урбах В.Ю. Статистический анализ в биологических и медицинских исследованиях. - М: Медицина, 1975, 295 с.
7. Кендел М. Ранговые корреляции. - М: Статистика, 1975, 214 с.
8. Холлендер М. Вулф Д.А. Непараметрические методы статистики. - Перевод с анг. под ред. Ю.П.Адлера и Ю.П.Тюрина. - М: Финансы и статистика, 1983, 518 с.
9. Гланц С. Медико-биологическая статистика. - Пер. с англ. доктора ф-м наук Ю.А.Данилова под ред. Н.Е.Бузикашвили и Д.В.Самойлова. - Москва: Практика, 1999, 459 с.
10. Реброва О.Ю. Статистический анализ медицинских данных. Применение прикладных программ СТАТИСТИКА. - М: Медиа Сфера, 2002, 305 с.
11. Ермолаев О.Ю. Математическая статистика для психологов. - 5-е изд. - М.: НОУ ВПО «МПСИ»: Флинта, 2011. - 336 с.
12. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. - 3-е изд. - М.: Наука, Главная редакция физико-математической литературы, 1983, 416 с.

СОДЕРЖАНИЕ

Задание 1. Построение интервального ряда и вычисление основных статистических характеристик……..... 3
Задание 2. Оценка параметров распределения..……... 17
Задание 3. Проверка статистических гипотез..…….…. 22
Задание 4. Элементы корреляционного анализа …….. 35
Приложения…….... 49
Литература…….. 68