

Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«Оренбургский государственный университет»

Кафедра технологии пищевых производств

П.В. Медведев,  
В.А. Федотов

# **УЧЕБНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА СТУДЕНТОВ**

## **Методические указания**

Рекомендовано к изданию редакционно-издательским советом федерального государственного бюджетного образовательного учреждения высшего образования «Оренбургский государственный университет» для обучающихся по образовательной программе высшего образования по направлению подготовки 19.03.02 Продукты питания из растительного сырья

Оренбург  
2019

УДК 664.65.05 (075.8)  
ББК 36.83-5я73  
М 42

Рецензент – доктор технических наук, профессор В. Ю. Полищук

**Медведев, П.В.**  
М 42 Учебно-исследовательская работа студентов: методические указания /  
П.В. Медведев, В.А. Федотов; Оренбургский гос. ун-т. – Оренбург:  
ОГУ, 2019. – 51 с.

Методические указания для выполнения лабораторных работ по дисциплине «Учебно-исследовательская работа студентов» содержат основы использования корреляционного анализа при проведении исследований, описание основных статистических характеристик исследуемых показателей.

Методические указания предназначены для обучающихся по образовательной программе высшего образования по направлению подготовки 19.03.02 Продукты питания из растительного сырья.

УДК 664.65.05 (075.8)  
ББК 36.82-5я73

© Медведев П.В.,  
Федотов В.А., 2019  
© ОГУ, 2019

## Содержание

Введение .....	4
1 Основы научных исследований.....	5
2 Лабораторные работы .....	9
2.1 Лабораторная работа № 1. Классификация типов данных .....	9
2.2 Лабораторная работа № 2. Статистические характеристики случайных величин .....	13
2.3 Лабораторная работа № 3. Нахождение ошибки многократных измерений.....	15
2.4 Лабораторная работа № 4. Проверка нормальности распределения случайной величины графическим способом.....	20
2.5 Лабораторная работа № 5. Проверка нормальности распределения случайной величины алгебраическим способом.....	24
2.6 Лабораторная работа № 6. Корреляционный анализ .....	26
Список использованных источников .....	49

## Введение

Научные исследования являются источником технологий, материалов и механизмов, при помощи которых становится возможным создавать продукты лучшего качества, меньшей стоимости, создавать методы лечения болезней, бороться со стихийными бедствиями и т.д.

Однако заниматься наукой является большой роскошью, так как вероятность получения практического результата из результатов исследований весьма невелика, а стоимость исследований может достигать колоссальных величин в связи с потребностью в экспериментальном оборудовании и сырье. Таким образом, позволить себе содержать собственное исследовательское подразделение могут лишь немногие коммерческие компании.

Подавляющая часть научных исследований финансируется Государством через различные фонды (российский фонд фундаментальных исследований РФФИ, фонд министерства образования и др.) и целевые отраслевые программы (Космическая программа, программа развития ОПК и др.). [2, 3]

За все время существования споров по поводу того - является ли математика наукой, является ли наукой литература, история или искусствоведение, было сформулировано множество различных определений термина Наука. С точки зрения авторов данной статьи наиболее логичным является определение К.Поппера, согласно которому мысль является научной, если она проходит три стадии:

- 1) постановку вопроса;
- 2) формулирование теории;
- 3) проведение эксперимента, подтверждающего или опровергающего теорию.

Такое определение является функциональным с точки зрения государства, являющегося основным источником финансирования научных работ и требующего максимальной эффективности затраченных средств.

# 1 Основы научных исследований

Если работа прошла три указанных стадии, то отчет по работе позволяет:

- наглядно увидеть то - на решение какой проблемы направлена исследовательская работа (по пункту «Формулирование вопроса»);
- использовать теорию или аналитическую модель, которая получила подтверждение в ходе проверочного эксперимента (пункты «Формулирование теории» и «Проведение эксперимента»), в других работах и изысканиях, экономя при этом средства на локальные опыты;
- исключать теорию и модель, опровергнутую в ходе подтверждающих экспериментов, при анализе рисков;
- использовать сведения о результатах эксперимента (пункт «Проведение эксперимента») при апробации других теорий и гипотез, экономя средства на проведение дублирующих опытов [7, 8, 9]

Модуль STATISTICA «Планирование эксперимента» предлагает исчерпывающий набор процедур для построения и анализа экспериментальных планов, используемых в промышленных исследованиях.

Вниманию пользователя предлагаются « $2^{k-p}$ » факторные планы с блоками (для планов, содержащих более 100 факторов, имеются высокоэффективные алгоритмы для нахождения планов с минимальной абберацией и максимально несмешанных планов, в которых пользователь может задавать эффекты взаимодействия, которые должны быть несмешанными), отсеивающие планы (для более 100 факторов предусмотрены планы Плакетта-Бермана), « $3^{k-p}$ » факторные планы с блоками (в том числе планы Бокса-Бенкена), смешанные планы, (малые) центральные композиционные планы (или поверхности отклика), планы на Латинских квадратах, планы для смесей и тернарных поверхностей, вершины и центры для поверхностей и смесей с ограничениями, оптимальные планы для факторных планов, поверхностей и смесей. Для анализа всех факторных планов, планов для поверхностей отклика и для смесей имеются сходные опции. Они могут

работать с несбалансированными и неполными планами и дают пользователю полный контроль при выборе модели, подгоняемой к данным. Программа вычислит обобщенную обратную матрицу для оцениваемых эффектов, а также выявления эффектов, являющихся псевдонимами других эффектов.

В качестве вспомогательного средства для определения точности подгонки соответствующей модели и для выявления выбросов вы можете просмотреть гистограммы остатков (и удаленных остатков) и предсказанных значений, диаграммы рассеяния (удаленных) остатков по предсказанным значениям или нормальные и полунормальные вероятностные графики (удаленных) остатков. Для проверки сериальной корреляции остатков вы можете также отобразить на графике (удаленные) остатки по номерам наблюдений. На всех графиках, на которых отображаются отдельные наблюдения (например, значения остатков наблюдений), точки обозначаются соответствующими номерами наблюдений или метками, что позволяет легко определить выбросы в наборе данных. Окончательно, могут быть вычислены значения максимального правдоподобия лямбда для преобразования Бокса-Кокса переменных отклика; результатам преобразования Бокса-Кокса сопутствует также график зависимости сумм квадратов остатков от лямбда, вместе с доверительными пределами для лямбда.

Более того, функция желательности может быть отображена на трехмерном графике поверхности или на контурном графике (контурные желательности), и пользователь может запросить матрицы таких графиков для всех факторов в анализе. Все установки, такие как сетка факторов или функции желательности, могут быть легко изменены для интерактивного анализа (например, можно быстро исключить отдельные отклики из анализа, и наблюдать эффекты на общей функции желательности). Спецификации для сложных функций желательности для многих переменных отклика могут быть сохранены в файле, а затем быстро восстановлены, если вы захотите анализировать другие эксперименты с использованием тех же переменных отклика. Кроме того, имеются опции для определения оптимального значения функции желательности, как с использованием сеточного поиска по экспериментальной области, так и с использованием эффективного алгоритма

оптимизации функции (который обычно применяется для оптимизации функций желательности в экспериментах с большим числом факторов).

Стандартные дробные факторные планы с блоками. Модуль STATISTICA Планирование эксперимента предоставляет полный набор всех стандартных планов (или так называемых планов с минимальной абберацией). Как выглядит подобный план в модуле STATISTICA показано на рисунке 1.

Пользователь может просмотреть планы в таблице результатов; опыты могут быть рандомизированы (в целом или по блокам), и в таблицу результатов могут быть добавлены пустые столбцы. Предоставляются опции для задания верхних и нижних установок факторов, для просмотра и сохранения планов в терминах перекодированных уровней факторов или в исходных метриках факторов. Пользователь может также запросить реплики, добавить центральные точки в план, а также запросить инверсию исходного плана. Дробные генераторы плана и блоковые генераторы плана, также как и матрица псевдонимов главных эффектов и взаимодействий, могут быть легко просмотрены.

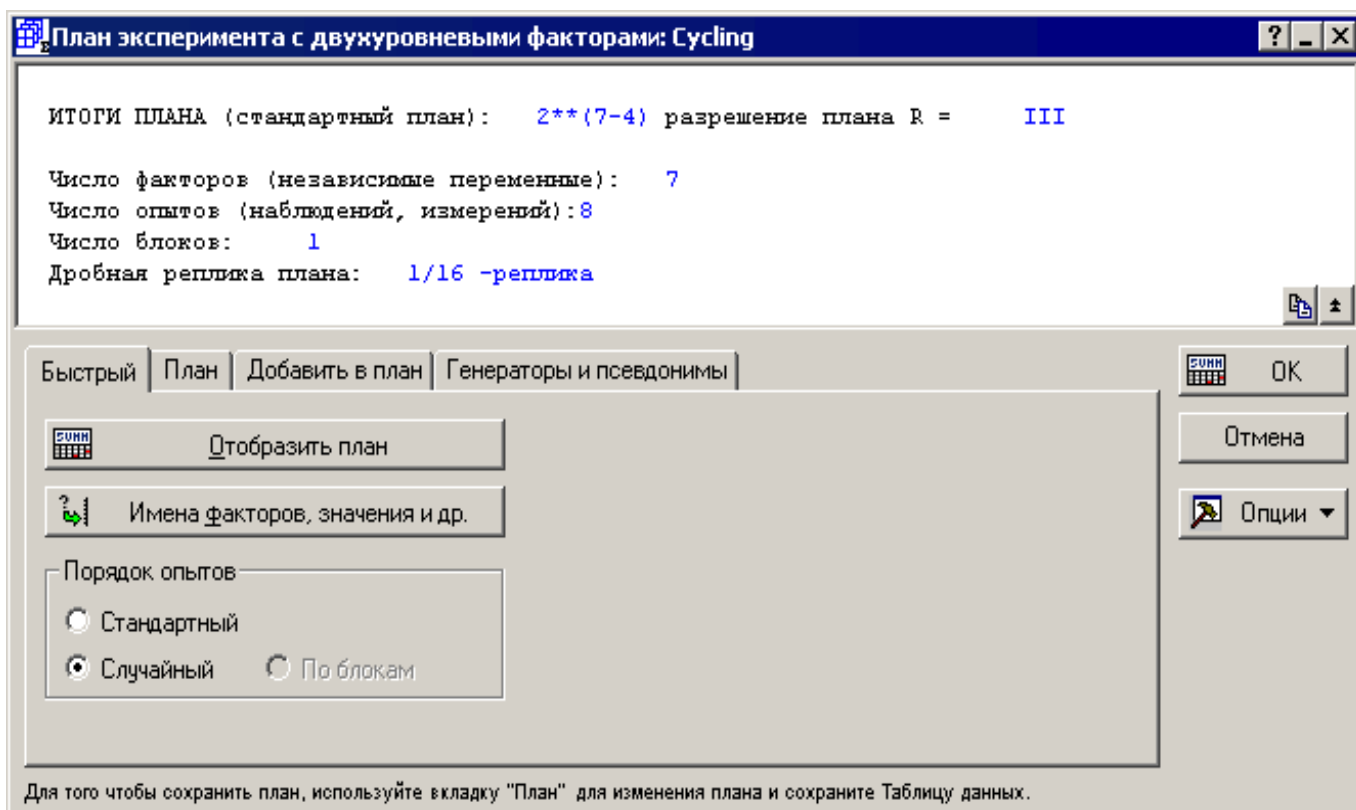


Рисунок 1 - План эксперимента с двухуровневыми факторами

Модуль STATISTICA Планирование эксперимента автоматически проведет исчерпывающий дисперсионный анализ плана. Пользователь полностью управляет включением эффектов и взаимодействий в модель, он может также просмотреть корреляции между столбцами матрицы плана (X), матрицу (т.е. ковариационную и корреляционную матрицу оценок параметров). Программа вычислит оценки параметров дисперсионного анализа, их стандартные ошибки и доверительные интервалы, коэффициенты перекодированных (-1, +1) значений факторов, их стандартные ошибки и доверительные интервалы, коэффициенты (стандартные ошибки, доверительные интервалы) для непреобразованных значений факторов.

Основываясь на этих оценках, программа может вычислить предсказанные значения (стандартные ошибки, доверительные интервалы) для заданных пользователем значений факторов. Программа выведет полную таблицу дисперсионного анализа, основанную на среднеквадратичном (MS) разностном члене, или, когда план хотя бы частично реплицирован, основываясь на оценке чистой ошибки (рисунок 2).

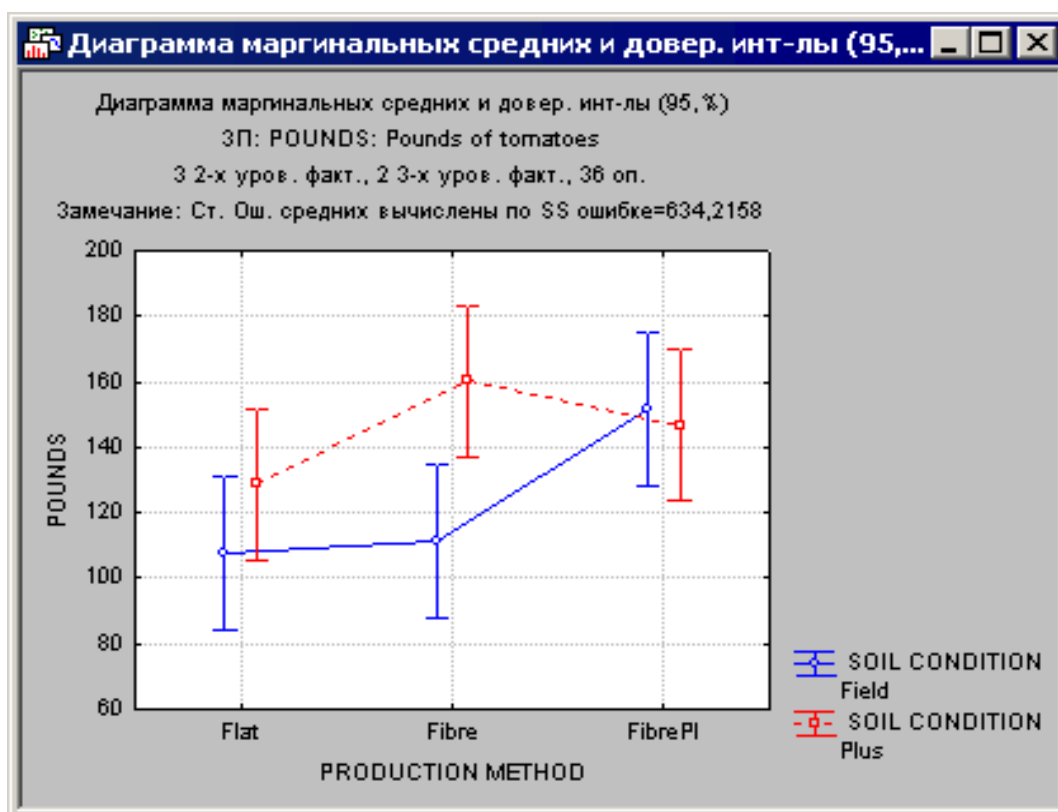


Рисунок 2 - Диаграмма маргинальных средних



Пользователь может просмотреть таблицу средних и маргинальных средних, а также их доверительные интервалы. Доступны многочисленные опции для просмотра результатов в графическом виде: диаграммы Парето эффектов нормальные и полунормальные вероятностные графики эффектов, квадратичные и кубические диаграммы, графики средних, графики взаимодействий (с доверительными интервалами для маргинальных средних), графики поверхности отклика, контуры поверхности отклика. В дополнение к этому, при проведении детального анализа остатков доступны все описанные выше основные процедуры для оценивания качества подгонки модели, а также для нахождения оптимальных установок факторов для одной или нескольких переменных отклика.

## **2 Лабораторные работы**

### **2.1 Лабораторная работа № 1. Классификация типов данных**

Значения переменных, которые регистрируются с помощью чисел, имеющих содержательный смысл, называют количественными данными. Данные, приведенные в таблице 1, в столбцах с 3 по 8, являются количественными. В то же время числа, приведенные в первом столбце этой таблицы, не являются количественными данными: они только указывают на номер государства при их алфавитном расположении. Эти числа не имеют содержательной интерпретации. С количественными данными можно выполнять все обычные операции над числами, такие, как вычисление среднего и оценку изменчивости.

В зависимости от того, какие значения может потенциально принимать переменная, выделяют два типа количественных данных: дискретные и непрерывные.

Дискретная - это такая переменная, которая может принимать значения только из некоторого списка определенных чисел. Примерами дискретной переменной

являются число детей в семье; число вызовов «скорой помощи», поступающих в больницу; число отказов изделия; число клиентов, обратившихся в фирму за определенный промежуток времени, и т. д.

Непрерывной будем считать любую переменную, не являющуюся дискретной. Она принимает значения из некоторого промежутка. Примерами непрерывной переменной является рост взрослого человека (например, от 140 до 230 см), фактическая масса буханки хлеба (например, от 750 до 830 г), дальность полета снаряда, урожайность культуры, выращенной в хозяйстве и т. п.

Есть данные, которые регистрируют определенное качество, которым обладает объект. Такие данные называют качественными. Даже если значениям этого качества можно приписать числа (например, полу человека приписать соответственно числа 0 и 1), то обрабатывать эти числа как количественные данные нельзя. Примерами качественных данных являются тип школы, где обучается ребенок (лицей, гимназия, специализированная физико-математическая школа); должность, которую занимает сотрудник на предприятии; названия газет, которые читают в определенном городе, и т. п.

Качественные данные бывают двух типов: порядковые, для которых существует имеющий содержательный смысл порядок, и номинальные, для которых нет содержательно интерпретируемого порядка.

Порядковые данные можно ранжировать и использовать это ранжирование при проведении статистического анализа. Примером порядковых данных являются ответы на вопросы анкеты, содержащей следующие варианты ответов: да; больше да, чем нет; больше нет, чем да; нет. Хотя и можно выразить эти ответы числами (например, 4, 3, 2, 1), но предложенная шкала оценок носит субъективный характер. Нельзя считать, что разница между ответами 4 и 3 такая же, как и между ответами 2 и 1. Также нельзя считать, что ответ 3 в три раза лучше ответа 1.

Таблица 1 - Производство пшеницы по годам, в миллионах тонн

Номер государства	Страна	2010	2011	2012	2013	2014	2015	2016
1	Россия	30,1	34,5	50,6	34,1	45,4	47,7	45,0
2	Австрия	1,3	1,3	1,4	1,2	1,7	1,5	1,4
3	Албания	0,4	0,3	0,3	0,3	0,3	0,3	0,2
4	Беларусь	0,4	1,0	1,0	0,8	1,1	1,2	1,1
5	Бельгия	-	1,7	1,7	1,6	1,9	1,8	1,6
6	Болгария	3,4	2,8	4,1	2,0	4,0	3,5	3,3
7	Венгрия	4,6	3,7	3,9	2,9	6,0	5,1	4,4
8	Германия	17,8	21,6	20,8	19,3	25,4	23,7	22,4
9	Греция	2,3	2,3	2,0	1,7	2,1	2,0	1,4
10	Дания	4,6	4,7	4,1	4,7	4,8	4,9	4,8
11	Ирландия	0,6	0,7	0,9	0,8	1,0	0,8	0,8
12	Испания	3,1	7,3	6,8	6,3	7,1	4,0	5,6
13	Италия	7,9	7,5	7,5	6,2	8,6	7,7	7,1
14	Латвия	0,2	0,4	0,5	0,5	0,5	0,7	0,6
15	Литва	0,6	1,2	1,2	1,2	1,4	1,4	0,8
16	Нидерланды	1,2	1,1	1,1	1,1	1,2	1,2	1,2
17	Норвегия	0,3	0,3	0,3	0,3	0,4	0,4	0,4
18	Республика Македония	0,4	0,3	0,3	0,2	0,4	0,3	0,3
19	Республика Молдова	1,3	0,7	1,1	0,1	0,9	1,1	0,7

Продолжение таблицы 1

Номер государства	Страна	2010	2011	2012	2013	2014	2015	2016
20	Румыния	7,7	4,5	4,4	2,5	7,8	7,3	5,5
21	Словакия	1,9	1,3	1,6	0,9	1,8	1,6	1,3
22	Словения	0,2	0,2	0,2	0,1	0,1	0,1	0,1
23	Соединенное Королевство (Великобритания)	14,3	16,7	16,0	14,3	15,5	14,9	14,7
24	Украина	16,3	10,2	20,6	3,6	17,5	18,7	13,9
25	Финляндия	0,4	0,5	0,6	0,7	0,8	0,8	0,7
26	Франция	30,9	37,4	38,9	30,5	39,7	36,9	35,4
27	Чешская Республика	3,8	4,1	3,9	2,6	5,0	4,1	3,5
28	Швейцария	0,6	0,6	0,5	0,4	0,5	0,5	0,5
29	Швеция	1,6	2,4	2,1	2,3	2,4	2,2	2,0
30	Эстония	0,1	0,1	0,1	0,1	0,2	0,3	0,2

Задание для выполнения: составить таблицы количественных и качественных показателей, характеризующих технологические свойства зерна, круп, муки, пшеницы, ржи и других зерновых культур, готовых макаронных, кондитерских и хлебобулочных изделий. Для качественных характеристик представить возможные альтернативные количественные диапазоны значений.

Например, в зависимости от влажности теста различают 3 типа замеса: твердый (влажность менее 30 %); средний (диапазон влажности от 30 до 32 %); мягкий (влажность более 32 %). Также, существует три типа замеса макаронного теста, в зависимости от температуры используемой воды:

- горячий (температура воды более 65 °С),
- теплый (температура воды от 35 до 65 °С),
- холодный (температура воды ниже 35 °С).

## 2.2 Лабораторная работа № 2. Статистические характеристики случайных величин

2.2.1 Среднее арифметическое ряда чисел – это сумма данных чисел, поделенная на количество слагаемых. Среднее арифметическое называют средним значением числового ряда.

Пример: Найдем среднее арифметическое чисел 2, 6, 9, 15.

Решение. У нас четыре числа. Значит, надо их сумму разделить на 4. Это и будет среднее арифметическое данных чисел

$$\frac{2 + 6 + 9 + 15}{4} = 8.$$

2.2.2 Среднее геометрическое ряда чисел – это корень n-й степени из произведения этих чисел.

Пример: Найдем среднее геометрическое чисел 2, 4, 8.

Решение. У нас три числа. Значит, надо найти корень третьей степени из их произведения. Это и будет среднее геометрическое данных чисел

$$\sqrt[3]{(2 \cdot 4 \cdot 8)} = \sqrt[3]{64} = 4.$$

2.2.3 Размах ряда чисел – это разность между наибольшим и наименьшим из этих чисел.

Пример: Найти размах чисел 2, 5, 8, 12, 33.

Решение: Наибольшее число здесь 33, наименьшее 2. Значит, размах составляет

$$33 - 2 = 31.$$

2.2.4 Мода ряда чисел – это число, которое встречается в данном ряду чаще других.

Пример: Найти моду ряда чисел 1, 7, 3, 8, 7, 12, 22, 7, 11, 22, 8.

Решение: Чаще всего в этом ряду чисел встречается число 7 (3 раза). Оно и является модой данного ряда чисел.

2.2.5 Медиана. В упорядоченном ряде чисел: медиана нечетного количества чисел – это число, записанное посередине.

Пример: В ряде чисел 2, 5, 9, 15, 21 медианой является число 9, находящееся посередине.

Медиана четного количества чисел – это среднее арифметическое двух чисел, находящихся посередине.

Пример: Найти медиану чисел 4, 5, 7, 11, 13, 19.

Решение: Здесь четное количество чисел (6). Поэтому ищем не одно, а два числа, записанных посередине. Это числа 7 и 11. Находим среднее арифметическое этих чисел

$$\frac{7+11}{2} = 9.$$

Число 9 и является медианой данного ряда чисел.

Медианой произвольного ряда чисел называется медиана соответствующего упорядоченного ряда.

Пример: Найдем медиану произвольного ряда чисел 5, 1, 3, 25, 19, 17, 21.

Решение: Располагаем числа в порядке возрастания: 1, 3, 5, 17, 19, 21, 25.

Посередине оказывается число 17. Оно и является медианой данного ряда чисел.

Пример: Добавим к нашему произвольному ряду чисел еще одно число, чтобы ряд стал четным, и найдем медиану: 5, 1, 3, 25, 19, 17, 21, 19.

Решение: Снова выстраиваем упорядоченный ряд: 1, 3, 5, 17, 19, 19, 21, 25.

Посередине оказались числа 17 и 19. Находим их среднее значение

$$\frac{17+19}{2}=18.$$

Число 18 и является медианой данного ряда чисел [4].

#### 2.2.6 Задания для выполнения.

Найти среднее арифметическое для значений влажности зерна 8 %, 9 %, 10 %, 15 %, 12 %, 9 %, 14 %, 25 %, 21 %, 27 %.

Найти медиану значений 5, 9, 10, 14, 21, 89, 54, 25, 9, 10, 14, 47, 74, 21.

Найти размах значений 89, 54, 25, 9, 1, 47, 97, 74, 9, 10, 14, 32, 47, 20, 15, 14.

Найти размах значений 9, 1, 89, 54, 25, 47, 97, 74, 9, 10, 14, 21, 12, 78, 23.

Найти среднее арифметическое значений 89, 54, 25, 9, 1, 47, 97, 74, 9, 10, 14, 32, 47, 20, 15, 14.

Найти среднее арифметическое значений 5, 9, 10, 14, 21, 89, 54, 25, 9, 10, 14, 47, 74, 21.

При помоле зерна были получены следующие 5 значений твердости зерновки: 35,1, 37,0 и 36,5 кг/мм<sup>2</sup>. Определить медиану, моду, среднее, выборочную дисперсию, доверительный интервал для математического ожидания, среднеквадратичное отклонение.

### **2.3 Лабораторная работа № 3. Нахождение ошибки многократных измерений**

Предварительная обработка результатов измерений или наблюдений необходима для того, чтобы отсеять так называемые грубые промахи, убрать из рассмотрения статистически незначимые факторы и в дальнейшем использовать оставшиеся данные для построения адекватных моделей исследуемого процесса.

Грубые промахи могут быть вызваны ошибками измерительных приборов, субъективными ошибками исследователя, методом обработки данных, влиянием неучтенных случайных факторов, округлением при вычислениях, другими причинами.

Другим важным моментом предварительной обработки данных является проверка соответствия распределения результатов измерения закону нормального распределения. Если эта гипотеза неприемлема, то следует определить, какому закону распределения подчиняются опытные данные, и если это возможно, преобразовать данное распределение к нормальному.

Только после предварительного анализа и исключения грубых промахов можно использовать оставшиеся данные для получения правильных решений. Существуют различные рекомендации для проведения отсева грубых погрешностей наблюдения. Чашу всего грубые погрешности можно заметить на визуальных отображениях - графиках величин (рисунок 3).

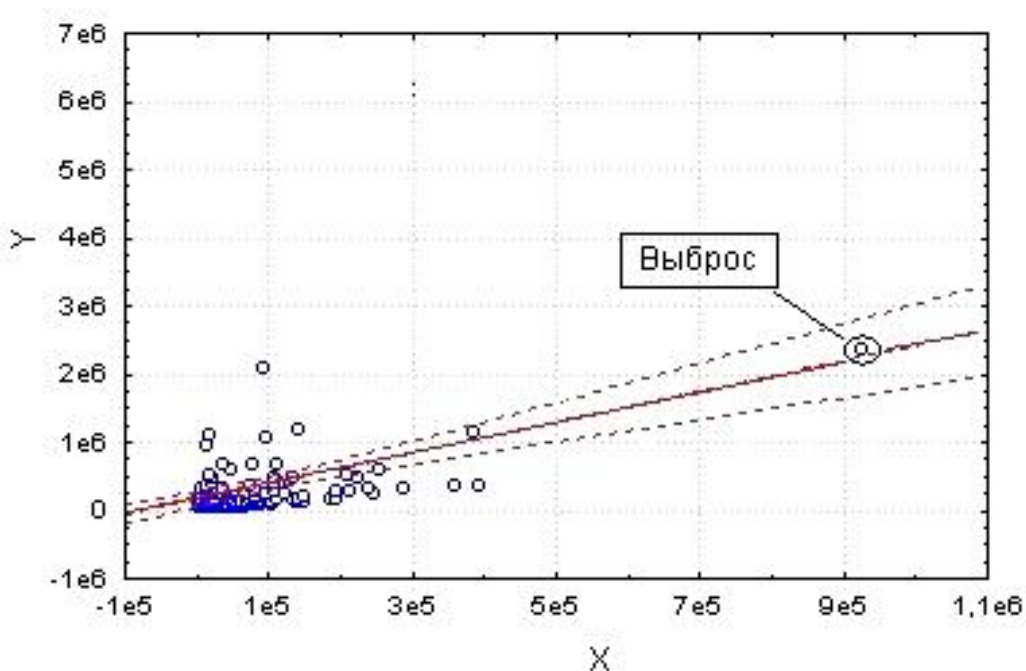


Рисунок 3 – Распределение случайной величины

Эксперимент является основным и наиболее совершенным методом познания. Он может быть активным и пассивным. Осуществление пассивного эксперимента не



зависит от экспериментатора, и ему приходится довольствоваться лишь ролью наблюдателя. Основной вид эксперимента – активный, проводится в контролируемых и управляемых условиях.

Все факторы, влияющие на исследуемые параметры объекта, предусмотреть, как правило, не удастся. Так, в сложных системах, зависящих от множества факторов, некоторые воздействия не могут контролироваться или управляться. Воздействие этих факторов рассматриваются как белый шум, наложенный на истинные результаты эксперимента. Чтобы отделить факторы, интересующие экспериментатора, от шумового фона, применяются специальные методы, называемые рандомизацией эксперимента.

Проведение активного эксперимента зачастую требует больших материальных затрат. Поэтому важной задачей является получение необходимых сведений при минимальном числе опытов. Решением этой проблемы занимается теория планирования эксперимента, представляющая собой раздел математической статистики.

Каждый эксперимент содержит элемент неопределенности вследствие ограниченности экспериментального материала. Постановка повторных (или параллельных) опытов не дает полностью совпадающих результатов, потому что всегда существует ошибка опыта (ошибка воспроизводимости). Эту ошибку и нужно оценить по параллельным опытам. Для этого опыт воспроизводится по возможности в одинаковых условиях несколько раз и затем берется среднее арифметическое всех результатов. Среднее арифметическое  $\bar{y}$  равно сумме всех  $n$  отдельных результатов, деленной на количество параллельных опытов  $n$

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n} = \frac{1}{n} \cdot \sum_{q=1}^n y_q \quad (2)$$

Отклонение результата любого опыта от среднего арифметического можно представить как разность  $y_q - \bar{y}$ , где  $y_q$  – результат отдельного опыта. Наличие

отклонения свидетельствует об изменчивости, вариации значений повторных опытов. Для измерения этой изменчивости чаще всего используют дисперсию. Дисперсией называется среднее значение квадрата отклонений величины от ее среднего значения. Дисперсия обозначается  $s^2$  и выражается формулой

$$s^2 = \frac{1}{n-1} \cdot \sum_{q=1}^n (y_q - \bar{y})^2, \quad (3)$$

где  $(n-1)$  – число степеней свободы, равное количеству опытов минус единица. Одна степень свободы использована для вычисления среднего.

Корень квадратный из дисперсии, взятый с положительным знаком, называется средним квадратическим отклонением, стандартом или квадратичной ошибкой

Стандарт имеет размерность той величины, для которой он вычислен. Дисперсия и стандарт – это меры рассеяния, изменчивости. Чем больше дисперсия и стандарт, тем больше рассеяны значения параллельных опытов около среднего значения.

Ошибка опыта являемся суммарной величиной, результатом многих ошибок: ошибок измерений факторов, ошибок измерений параметра оптимизации и др. Каждую из этих ошибок можно, в свою очередь, разделить на составляющие.

Вопрос о классификации ошибок довольно сложный и вызывает много дискуссий. В качестве примера одной из возможных схем классификации приведем схему (рисунок 4).

Все ошибки принято разделять на два класса: систематические и случайные.

Систематические ошибки порождаются причинами, действующими регулярно, в определенном направлении. Чаще всего эти ошибки можно изучить и определить количественно.

Систематические ошибки обуславливаются причинами, действующими вполне определённым образом. Примером систематической ошибки при взвешивании может являться смещение стрелки ненагруженных весов относительно

нулевой отметки на некоторую постоянную величину. Зная это смещение (например, взвесив гирю, масса которой точно известна), можно, всякий раз измеряя массу на этих весах, вычитать из показаний прибора. Таким образом, систематические ошибки могут быть устранены или достаточно точно учтены.

Систематические ошибки находят, калибруя измерительные приборы и сопоставляя опытные данные с изменяющимися внешними условиями (например, при градуировке термометры по реперным точкам, при сравнении с эталонным прибором). Если систематические ошибки вызываются внешними условиями (переменной температуры, сырья и т. д.), следует компенсировать их влияние. Как это делать, будет показано ниже. Случайными ошибками называются те, которые появляются нерегулярно, причины возникновения которых неизвестны и которые невозможно учесть заранее.

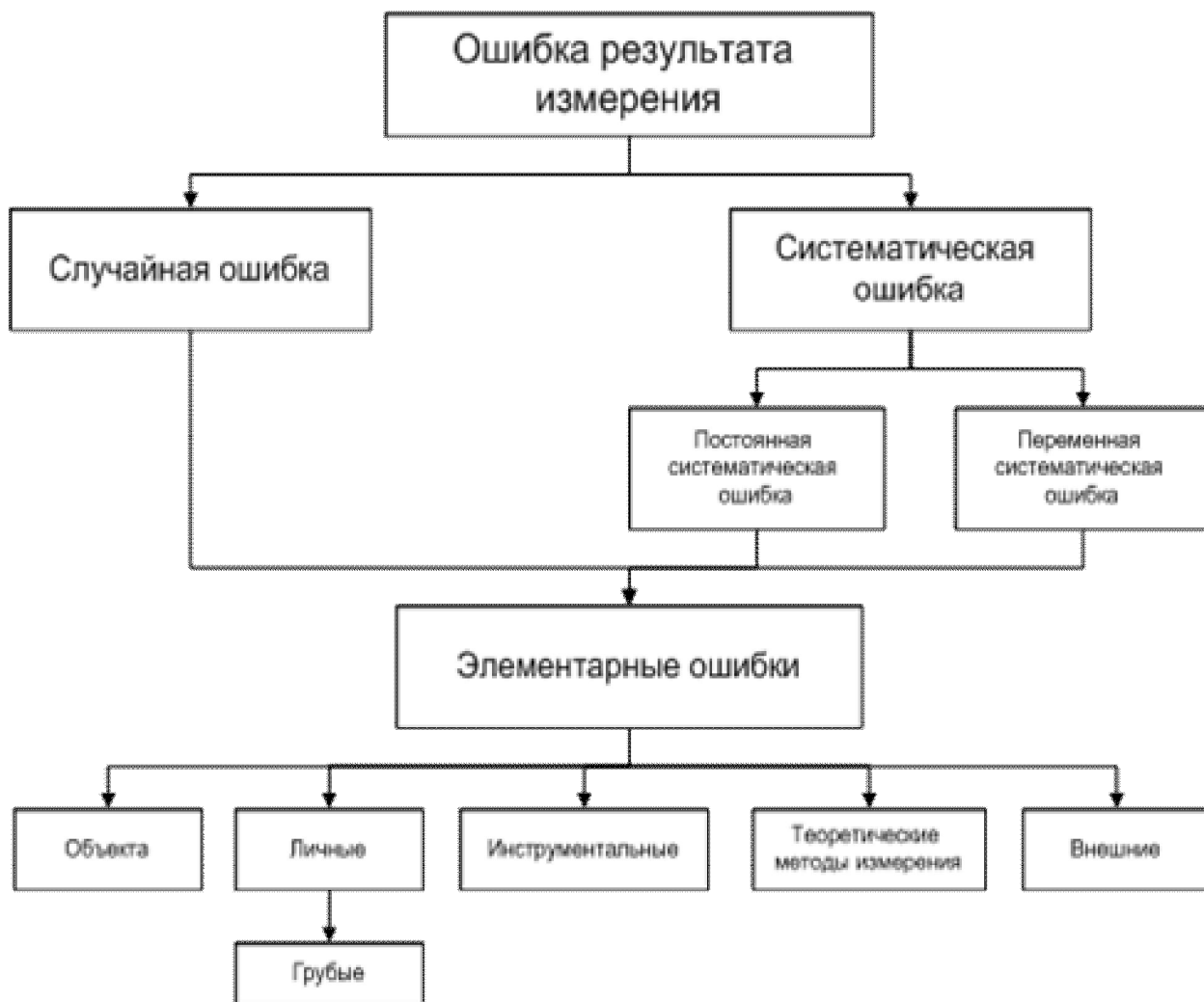


Рисунок 4 – Классификация ошибок измерений

При повторении опытов такие ошибки могут вызвать большой разброс экспериментальных результатов [10].

Очень важно исключить из экспериментальных данных грубые ошибки, так называемый брак при повторных опытах. Для отброса ошибочных опытов существуют правила. Для определения брака используют, например, критерий Стьюдента (таблица А.1, рисунок А.1)

$$\frac{y - \bar{y}}{s} \geq t. \quad (4)$$

Значение  $t$  берут из таблицы  $t$ -распределения Стьюдента. Опыт считается бракованным, если экспериментальное значение критерия  $t$  по модулю больше табличного значения.

Задание для выполнения: изучить основные положения изучаемой темы, провести исследования для получения сведений о необходимых показателях качества по стандартным или общепринятым методикам (ГОСТ, ТУ), сделать вывод по проделанной работе. Измерить влажность одного образца зерна, муки, хлеба в 5 повторностях, построить кривую распределения значений и записать средние значения с вычисленными погрешностями измерений.

#### **2.4 Лабораторная работа № 4. Проверка нормальности распределения случайной величины графическим способом**

Встречаемые в обыденной жизни большое число наблюдений, измеряемые значения параметров, погрешности измерений и ошибки подчиняются нормальному закону. Этот закон можно считать основой математической статистики. Нормальным называется закон распределения случайной величины  $x$ , если плотность распределения вероятности определяется по формуле

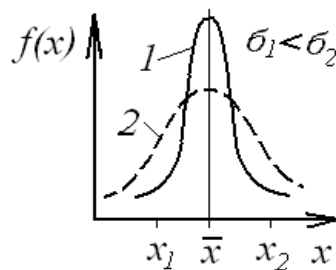
$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\bar{x})^2}, \quad (-\infty < x < +\infty, \sigma > 0), \quad (5)$$

где  $\sigma$  – среднее квадратическое отклонение,

$\bar{x}$  - среднее значение (математическое ожидание),

$\pi = 3,141593$ ,  $e = 2,718282$  – математические константы.

Из формулы следует, что нормальное распределение полностью определяется параметрами  $\bar{x}$  и  $\sigma$ . Среднее квадратическое отклонение определяет форму кривой: чем больше  $\sigma$  (разброс данных), тем кривая становится пологой (рисунок 5). Математическое ожидание  $\bar{x}$  определяет положение кривой на оси абсцисс, кривая симметрична относительно этого значения.



1 – с меньшим разбросом данных, 2 – с большим разбросом данных

Рисунок 5 – Виды графиков нормального распределения случайной величины

Нормальное распределение часто встречается в природе. Например, следующие случайные величины хорошо моделируются нормальным распределением:

- отклонение при стрельбе;
- погрешности измерений (однако погрешности некоторых измерительных приборов имеют не нормальные распределения);
- рост человека;
- давление крови в течение дня;
- экзаменационные оценки;

- некоторые характеристики живых организмов в популяции.

Многие непрерывные случайные величины не являются ни точно, ни приближенно нормальными. Свойства таких величин довольно сильно отличаются от свойств нормального распределения, перечисленных выше (рисунок 6).

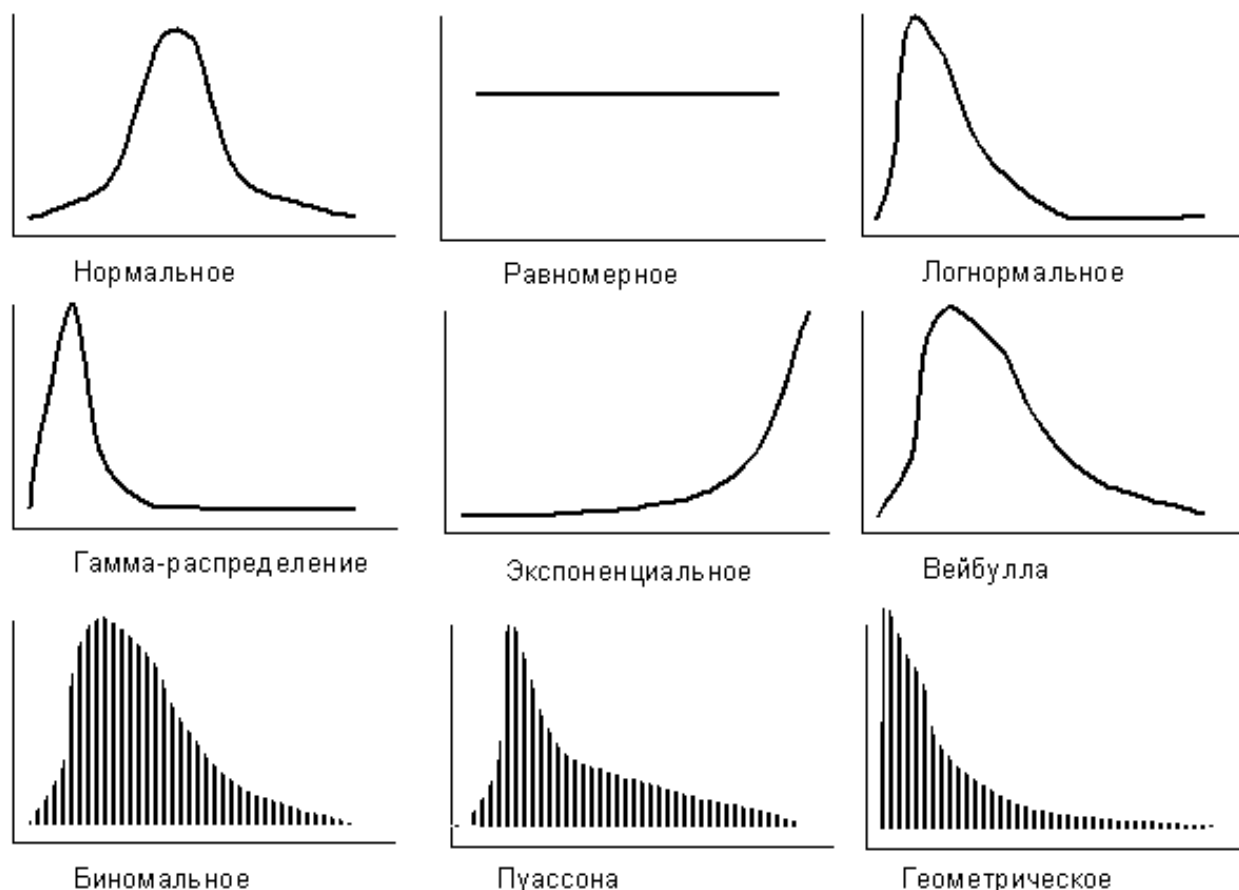


Рисунок 6 – Виды распределений величин

Вероятность, что подброшенная монета упадет орлом вверх 50 %, что при броске шестигранного кубика выпадет сторона «4» - 16,7 %, что завтра на кого-нибудь упадет метеорит – 0,00000000294 %. Это простые примеры, достаточно разделить количество желаемых событий на общее количество случаев и мы получаем вероятность события, но когда результаты эксперимента могут быть не только орлом или решкой (что эквивалентно да/нет), а большим набором данных. Например, вес батона хлеба, если мы возьмём в магазине 1000 буханок хлеба и взвесим каждую, то мы узнаем, что на самом деле батон не весит 400 грамм, результаты будут варьироваться в диапазоне от 384 до 416 грамм (допуск разброса

веса предусмотрен ГОСТ). Если Вы построите график «Количество буханок – Вес», то график будет иметь форму напоминающую колокол, что-то похожее на следующий график (рисунок 7).

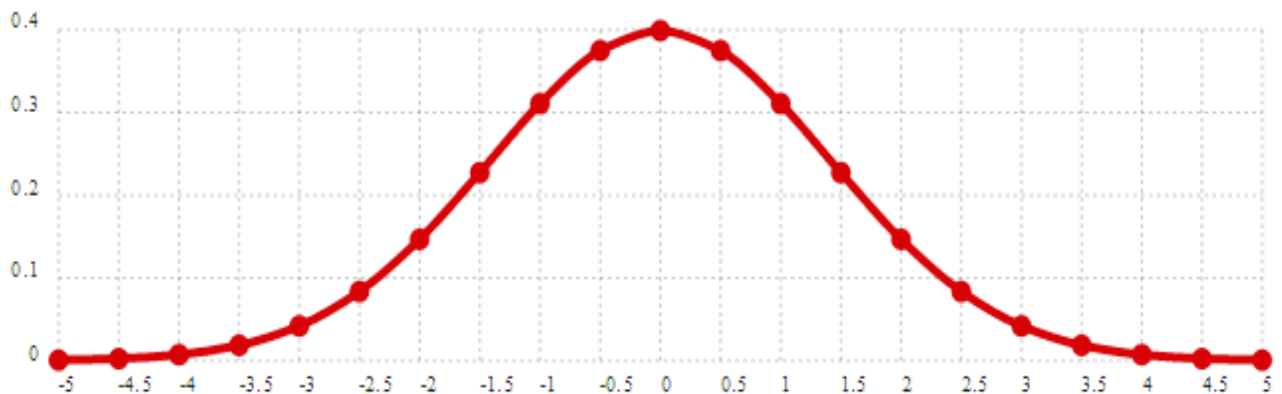


Рисунок 7 – Плотность вероятности нормального распределения

Такую форму график получит потому, что большинство значений близко к 400. Это - пример нормального распределения, множество событий имеют закон нормального распределения, например, вес или рост для определённого возраста, или среднее время Вашего похода до магазина и многие другие события также подчиняются закону нормального распределения. Так работают маркетологи: проводят опрос 1000 человек и получают представление о всём населении.

В случае таблицы Вы имеете дело с дискретными данными, т.е. для каждого веса есть определённая вероятность, но в случае графика дело немного меняется, теперь мы говорим не о 1000 буханок, которые мы взвесили, а обо всех буханках в мире сразу, чтобы не взвешивать все буханки. Имея закон распределения, который мы получили взвесив 1000 буханок (мы могли взвесить 100, 200, 500, сколько угодно), мы можем предположить, что сколько бы мы буханок не взяли, замерив их, мы получим ту же форму колокола. Используя термины статистики, все буханки хлеба - это генеральная совокупность, 1000 замеренных буханок - выборка.

Нормальное распределение получило своё название абсолютно справедливо: по статистике, большинство событий происходят именно с вероятностью нормального распределения, но что это значит? Это означает, например, что когда

Вы видите на упаковке хлеба обозначение «Вес:  $400 \pm 16$  г» - вес батона имеет нормальное распределение со средним значением 400 г и стандартным отклонением 16 г.

Предположим, что математическое ожидание зольности в зерне равно 0,6 %, а среднеквадратичное отклонение 0,15 %. В этом случае мы можем быть уверены в том, что величина фактически измеренного значения процентного содержания минеральных веществ в зерне будет находиться в интервалах:

$$0,6 \pm 1,00 \cdot 0,15 = 0,6 \pm 0,15 \text{ с вероятностью } 0,68;$$

$$0,6 \pm 1,64 \cdot 0,15 = 0,6 \pm 0,25 \text{ с вероятностью } 0,90;$$

$$0,6 \pm 1,96 \cdot 0,15 = 0,6 \pm 0,29 \text{ с вероятностью } 0,95;$$

$$0,6 \pm 3,00 \cdot 0,15 = 0,6 \pm 0,45 \text{ с вероятностью } 0,997,$$

т.е. из 1000 проб только 3 пробы по содержанию минеральных веществ в зерне будут выходить из диапазона от 0,15 % до 1,05 %.

Задание: использовать для исследований не менее 10 образцов одного вида хлеба для получения сведений о распределении случайной величины (веса буханки). Отобразить в отчете графическое представление данных о плотности вероятности нормального распределения.

## **2.5 Лабораторная работа № 5. Проверка нормальности распределения случайной величины алгебраическим способом**

Проверку нормальности распределения случайной величины можно производить разными способами, рассмотрим простую рекомендацию. Надо вычислить среднее абсолютное по обычной формуле



$$e = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}. \quad (6)$$

Если выполняется условие

$$\left| \frac{e}{S} - 0,7979 \right| < \frac{0,4}{\sqrt{n}}, \quad (7)$$

то гипотеза нормальности распределения выборки данных принимается. В противном случае надо постараться преобразовать распределение к нормальному или применить другое распределение.

Задание для выполнения: занести в таблицу 2 данные анализа показателей качества зерна имеющихся образцов пшеницы (желательно более 5). Построить диаграммы распределения этих показателей качества и оценить «нормальность» распределения данных.

Таблица 2 – Данные анализа показателей качества зерна

№ образца	Число падения, с	Количество клейковины, %	Качество клейковины, ед. ИДК
1			
2			
3			
4			
5			
6			
7			

Проверить на нормальное распределение ряд значений 21, 89, 54, 25, 9, 10, 14, 47, 74, 21, 89, 54, 25, 9, 1, 47, 97.

Установить, соответствие критериям нормальности численного ряда 89, 54, 25, 9, 1, 47, 97, 74, 9, 10, 14, 32, 47, 20, 15, 14.

Построить диаграмму нормальности для ряда значений 97, 74, 9, 10, 14, 32, 47, 54, 78, 12, 45, 56, 54.

## **2.6 Лабораторная работа № 6. Корреляционный анализ**

### 2.6.1 Общие сведения

Корреляционный анализ – метод математической статистики, используемый для изучения, исследования взаимосвязи между (генеральными) экономическими показателями на основе их наблюдаемых статистических (выборочных) аналогов.

Парный корреляционный анализ – изучение взаимосвязи между двумя показателями, описывающими свойства однотипных объектов из некоторой совокупности.

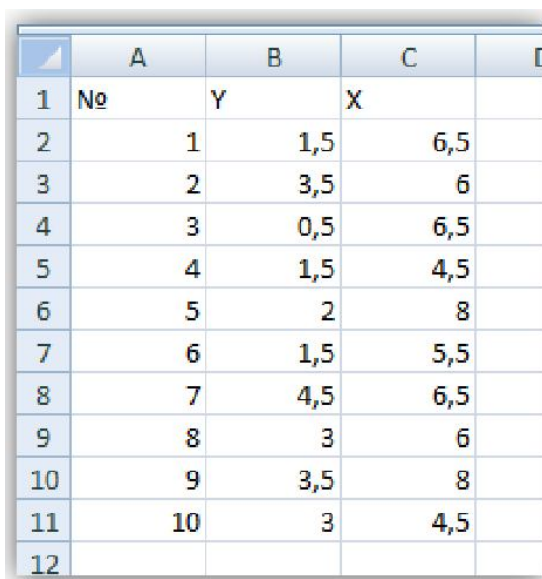
Шаги корреляционного анализа:

- 1) постановка задачи.
- 2) сбор и анализ данных; определение формы корреляционной связи (линейная, криволинейная).
- 3) вычисление показателя тесноты корреляционной связи.
- 4) оценка статистической значимости показателя тесноты корреляционной связи.

Сбор данных. Сбор данных осуществляется методом случайной выборки некоторого количества наблюдаемых объектов из некоторой однородной

совокупности, фиксации для каждого выбранного объекта пары признаков (свойств), взаимосвязь которых будет предметом исследования (рисунок 8).

Визуальный анализ данных. Визуальная оценка осуществляется на основе графического анализа. Данные на графике можно представить в виде точечной диаграммы в MS Excel. В результате такой оценки может быть сделана гипотеза о наличии линейной корреляционной связи, о нелинейной корреляционной связи или об отсутствии корреляционной связи.



	A	B	C	D
1	№	Y	X	
2		1	1,5	6,5
3		2	3,5	6
4		3	0,5	6,5
5		4	1,5	4,5
6		5	2	8
7		6	1,5	5,5
8		7	4,5	6,5
9		8	3	6
10		9	3,5	8
11		10	3	4,5
12				

Рисунок 8 – Таблица данных

Вычисление показателей тесноты корреляционной связи. Если визуальный анализ позволяет принять гипотезу о линейной форме связи между показателями – для оценки степени тесноты связи применяется линейный коэффициент корреляции  $r$ .

Границы измерения:  $-1 \leq r \leq 1$ .

Если взаимосвязь между показателями обратная (отрицательная) то корреляционная связь отрицательная:  $-1 < r < 0$ .

Если взаимосвязь между показателями прямая, то корреляционная зависимость положительная:  $0 < r < 1$ .

Если  $r = 0$ , линейная корреляционная зависимость отсутствует (рисунок 9).

В крайних случаях  $|r| = 1$  имеется функциональная линейная зависимость между показателями  $x$  и  $y$ . [12] Если визуальный анализ не позволяет принять гипотезу о линейной форме связи, коэффициент корреляции в этом случае применять неправомерно (рисунок 10).

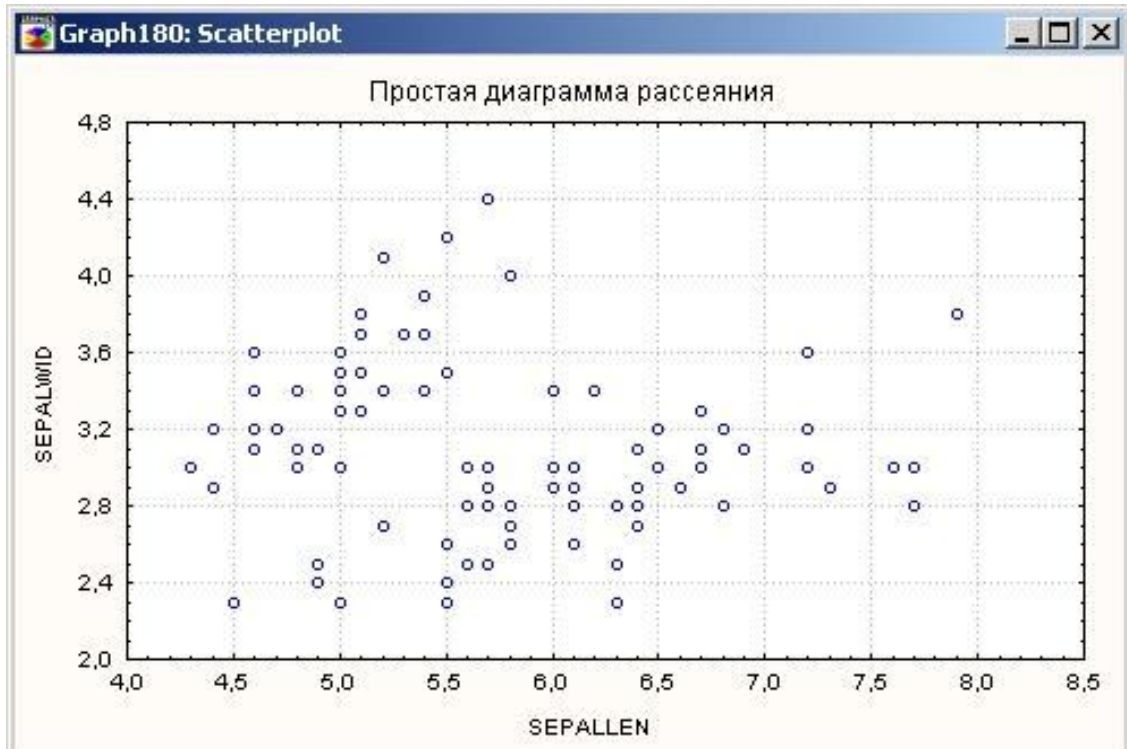


Рисунок 9 – Пример отсутствия зависимости

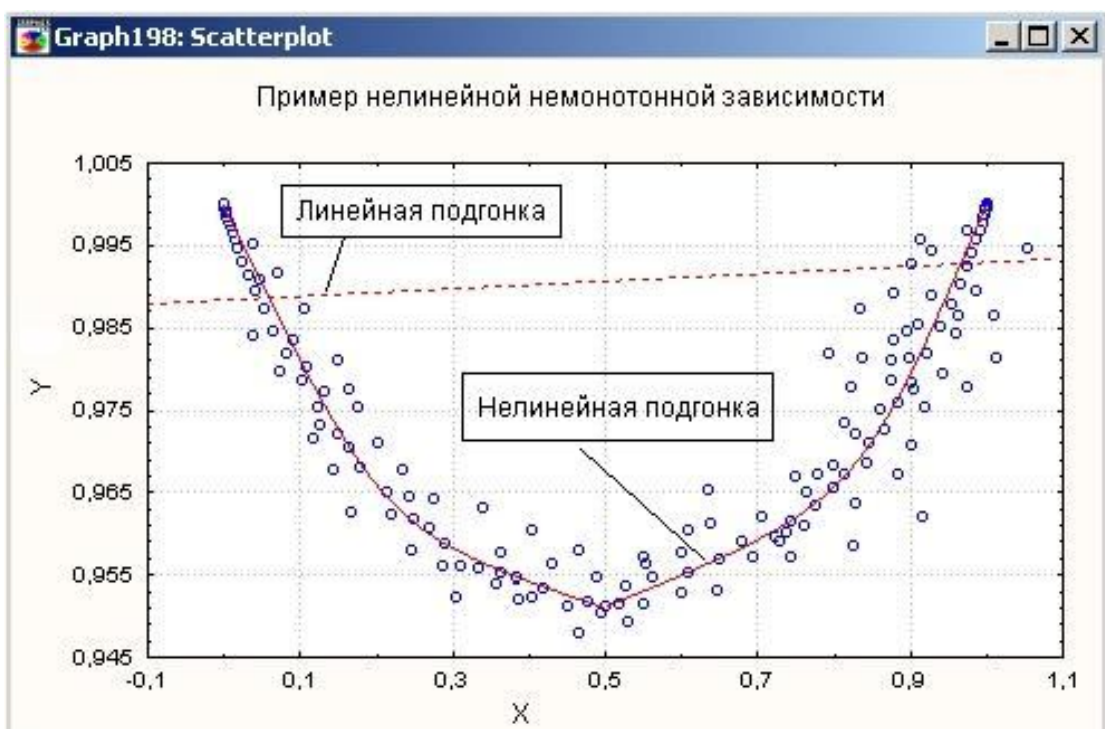


Рисунок 10 – Пример нелинейной корреляции

В некоторых случаях, несмотря на высокие значения коэффициента корреляции обнаруженная связь оказывается ложной (рисунок 11).

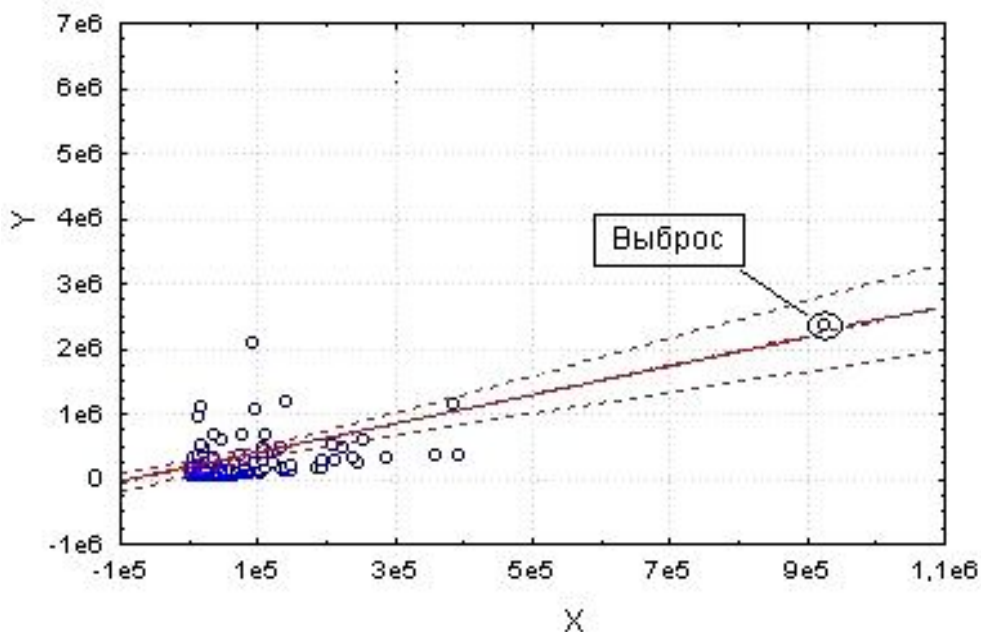


Рисунок 11 – Диаграмма сильной корреляции, обусловленной единственным выбросом ( $r = 0,8284$ )

Внутри каждой группы корреляция окажется существенной (рисунок 12).

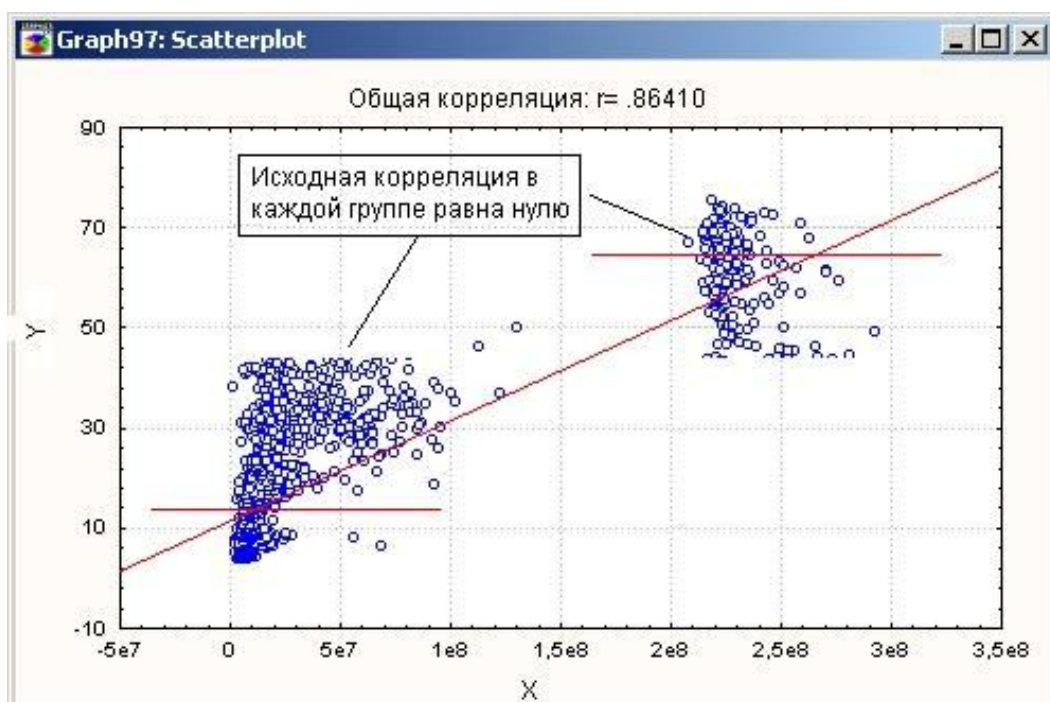


Рисунок 12 – Диаграмма данных, нуждающихся в группировании для оценки их корреляции

В некоторых случаях, низкий коэффициент корреляции изначально ошибочен, и исследуемую группу образцов или явления следует разделить на две-три разные группы.

## 2.6.2 Интерпретация коэффициента корреляции

Как интерпретировать значение коэффициента корреляции Пирсона? Для оценки тесноты, или силы, корреляционной связи обычно используют общепринятые критерии, согласно которым абсолютные значения  $r < 0,3$  свидетельствуют о слабой связи, значения  $r$  от 0,3 до 0,7 – о связи средней тесноты, значения  $r > 0,7$  – о сильной связи. Более точную оценку силы корреляционной связи можно получить, если воспользоваться таблицей Чеддока (таблица 2).

Таблица 2 – Таблица Чеддока для интерпретации силы корреляционной связи

Абсолютное значение $r$	Теснота (сила) корреляционной связи
менее 0,3	слабая
от 0,3 до 0,5	умеренная
от 0,5 до 0,7	заметная
от 0,7 до 0,9	высокая
более 0,9	весьма высокая

Другим способом оценки силы связи по коэффициенту корреляции является его сравнение с некоторым критическим, который в свою очередь зависит от количества измерений, опытов, наблюдений (таблица А.2).

Допустим, проводится независимое измерение различных параметров у одного типа объектов. Из этих данных можно получить качественно новую информацию - о взаимосвязи этих параметров. [13]

Например, измеряем рост и вес человека, каждое измерение представлено точкой в двумерном пространстве (рисунок 13). Несмотря на то, что величины носят случайный характер, в общем наблюдается некоторая зависимость - величины коррелируют.

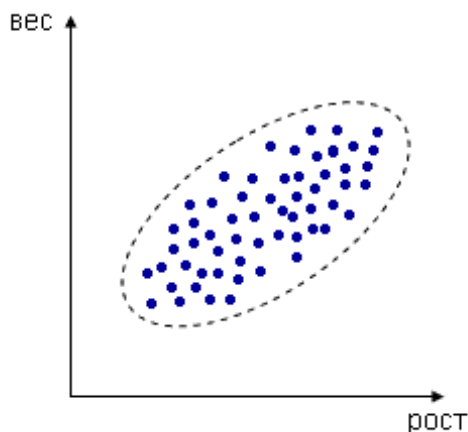


Рисунок 13 – Положительная корреляция

В данном случае это положительная корреляция (при увеличении одного параметра второй тоже увеличивается). Возможны также такие случаи (рисунки 14, 15).

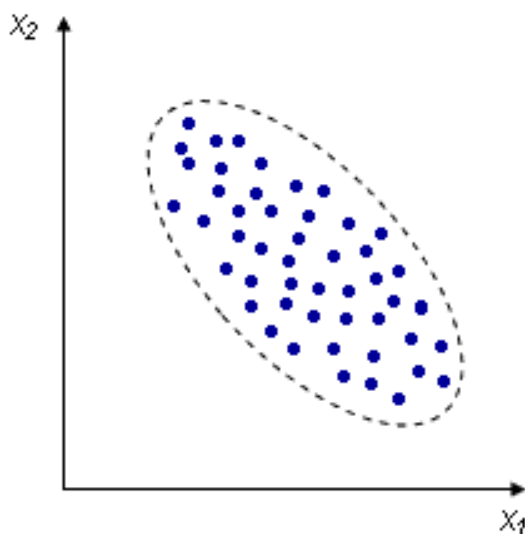


Рисунок 14 – Отрицательная корреляция

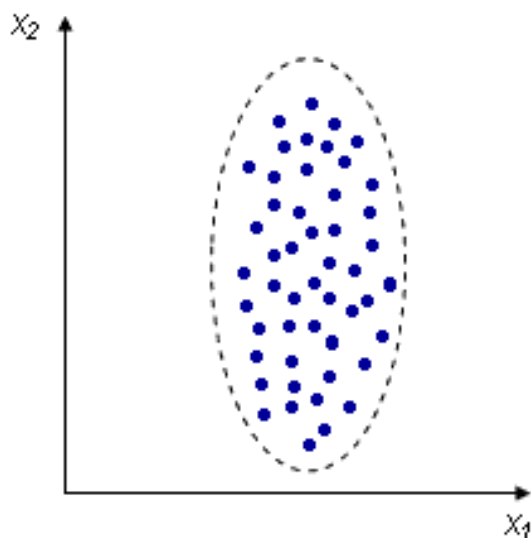


Рисунок 15 – Отсутствие корреляции

Взаимосвязь между переменными необходимо охарактеризовать численно, чтобы, например, различать такие случаи (рисунки 16, 17).

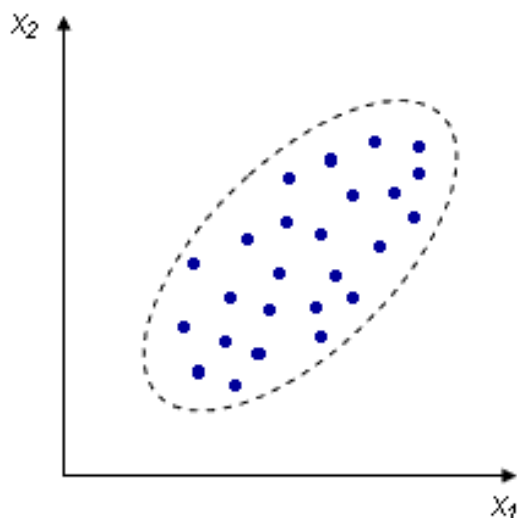


Рисунок 16 – Слабая связь

Оценка статистической значимости линейного коэффициента корреляции проводится с помощью теста Стьюдента по t-статистике

$$t_r = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}}. \quad (8)$$



Если значение коэффициента  $t$  будет больше или равно табличного значения (таблица А.1), то коэффициент корреляции статистически значим.

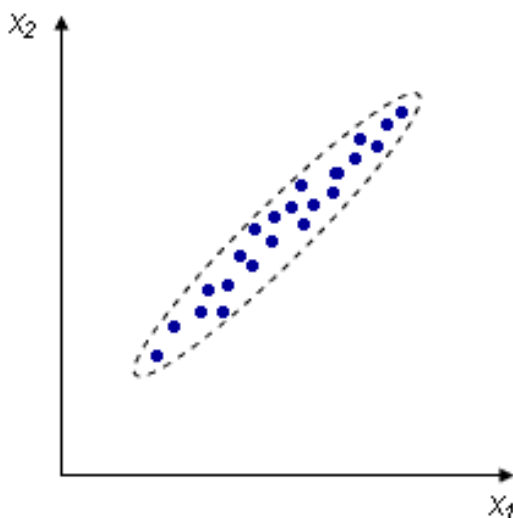


Рисунок 17 – Сильная связь

Таким образом, при определенном уровне значимости проверяется гипотеза о том, что в генеральной совокупности нет корреляционной зависимости между анализируемыми показателями.

### 2.6.3 Условия и ограничения применения критерия Пирсона

У критерия Пирсона существует несколько допустимых к использованию условий.

1. Сопоставляемые показатели должны быть измерены в количественной шкале (например, стекловидность, содержание белка).
2. Посредством критерия корреляции Пирсона можно определить лишь наличие и силу линейной взаимосвязи между величинами.

Прочие характеристики связи, в том числе направление (прямая или обратная), характер изменений (прямолинейный или криволинейный), а также

наличие зависимости одной переменной от другой - определяются при помощи регрессионного анализа.

3. Количество сопоставляемых величин должно быть равно двум. В случае анализ взаимосвязи трех и более параметров следует воспользоваться методом факторного анализа.

4. Критерий корреляции Пирсона является параметрическим, в связи с чем условием его применения служит нормальное распределение сопоставляемых переменных. В случае необходимости корреляционного анализа показателей, распределение которых отличается от нормального, в том числе измеренных в порядковой шкале, следует использовать коэффициент ранговой корреляции Спирмена.

5. Следует четко различать понятия зависимости и корреляции. Зависимость величин обуславливает наличие корреляционной связи между ними, но не наоборот.

#### 2.6.4 Пример расчета коэффициента корреляции Пирсона

Одним из критериев, позволяющих определить, какова теснота (или сила) корреляционной связи между двумя показателями, измеренными в количественной шкале является критерий корреляции Пирсона.

Целью исследования явилось выявление, определение тесноты и статистической значимости корреляционной связи между двумя количественными показателями: газообразующей способностью муки ГОСМ (X) и показателя амилолитической активности муки («число падения» или ЧП) - (Y). Исходные данные для выборки, состоящей из 5 опытов ( $n = 5$ ), сведены в таблице 3.

Таблица 3 – Исходные данные по результатам экспериментов

Газообразующая способность муки, мл CO <sub>2</sub> на 100 г (X)	Число падения, % (Y)
951	83
874	76
957	84
1084	89
903	79

Вычислим суммы анализируемых значений X и Y

$$\Sigma(X) = 951 + 874 + 957 + 1084 + 903 = 4769,$$

$$\Sigma(Y) = 83 + 76 + 84 + 89 + 79 = 441.$$

Найдем средние арифметические для X и Y

$$M_x = \Sigma(X) / n = 4769 / 5 = 953,8,$$

$$M_y = \Sigma(Y) / n = 441 / 5 = 88,2.$$

Рассчитаем для каждого значения сопоставляемых показателей величину отклонения от среднего арифметического dx и dy (таблица 4).

Таблица 4 – Расчет отклонений значений экспериментов от среднего арифметического

Газообразующая способность муки, мл CO <sub>2</sub> на 100 г (X)	Число падения, % (Y)	Отклонение ГОСМ от среднего значения ( $d_x$ )	Отклонение % ЧП от среднего значения ( $d_y$ )
951	83	-2,8	0,8
874	76	-79,8	-6,2
957	84	3,2	1,8
1084	89	130,2	6,8
903	79	-50,8	-3,2

Возведем в квадрат каждое значение отклонения  $d_x$  и  $d_y$  – таблица 5.

Таблица 5 – Расчеты промежуточных статистических данных

Газообразующая способность муки, мл CO <sub>2</sub> на 100 г (X)	Число падения, % (Y)	Отклонение ГОСМ от среднего значения ( $d_x$ )	Отклонение % ЧП от среднего значения ( $d_y$ )	$d_x^2$	$d_y^2$
951	83	-2,8	0,8	7,84	0,64
874	76	-79,8	-6,2	6368,04	38,44
957	84	3,2	1,8	10,24	3,24
1084	89	130,2	6,8	16952,04	46,24
903	79	-50,8	-3,2	2580,64	10,24

Рассчитаем для каждой пары анализируемых значений произведение отклонений  $d_x$  и  $d_y$  (таблица 6).

Определим значения суммы квадратов отклонений  $\sum d_x^2$  и  $\sum d_y^2$

$$\sum d_x^2 = 25918,8,$$

$$\sum d_y^2 = 98,8.$$

Таблица 6 – Значения промежуточных статистических данных

Газообразующая способность муки, мл CO <sub>2</sub> на 100 г (X)	Число падений, % (Y)	Отклонение ГОСМ от среднего значения ( $d_x$ )	Отклонение % ЧП от среднего значения ( $d_y$ )	$d_x^2$	$d_y^2$	$d_x \cdot d_y$
951	83	-2,8	0,8	7,84	0,64	-2,24
874	76	-79,8	-6,2	6368,04	38,44	494,76
957	84	3,2	1,8	10,24	3,24	5,76
1084	89	130,2	6,8	16952,04	46,24	885,36
903	79	-50,8	-3,2	2580,64	10,24	162,56

Найдем значение суммы произведений отклонений  $\sum (d_x \cdot d_y)$

$$\sum (d_x \cdot d_y) = 1546,2$$

Рассчитаем значение коэффициента корреляции Пирсона  $r$  по приведенной выше формуле

$$r = \frac{\sum (d_x \cdot d_y)}{\sqrt{(\sum d_x^2 \cdot \sum d_y^2)}}, \quad (8)$$

$$r = \frac{1546,2}{\sqrt{25918,8 \cdot 98,8}} = 0,966.$$

Найдем значение t-критерия для оценки статистической значимости корреляционной связи

$$t_r = \frac{0,97 \cdot \sqrt{5-2}}{\sqrt{1-0,97^2}} = 7,0.$$

Критическое значение t-критерия найдем по таблице, где при числе степеней свободы  $f = n - 2 = 3$  и уровне значимости  $p = 0,01$  значение  $t_{table} = 5,84$ . Рассчитанное значение  $t_r$  (7,0) больше  $t_{table}$  (5,84), следовательно связь является статистически значимой.

Сделаем статистический вывод. Значение коэффициента корреляции Пирсона составило 0,97, что соответствует весьма высокой тесноте связи между газообразующей способностью муки и амилолитической активностью муки. Данная корреляционная связь является статистически значимой ( $p < 0,01$ ).

## 2.6.5 Задания для выполнения

2.6.5.1 Определите отсутствие или наличие линейной корреляционной взаимосвязи между основными характеристиками зернового анализа зерна Оренбургской области. Для анализа используйте данные Оренбургского государственного центра агрохимической службы (таблица 7).

Этапы выполнения задания:

- 1) постройте точечную диаграмму и выдвинете гипотезу о характере связи между рассматриваемыми переменными;
- 2) рассчитайте коэффициент корреляции;

3) проверьте значимость коэффициента корреляции;

4) сделайте выводы.

Таблица 7 – Основные характеристики зернового анализа пшеницы

Сорт	Зона произрастания	Объем зерновки, куб.мм	Натура, г/л	Масса 1000 зерен, г	Стекловидность , %
Твердые сорта					
Харьковская 3	восток	22,60	810	31,1	88
	центр	22,63	807	31,5	85
	запад	22,57	825	30,2	85
Оренбургская 10	запад	22,50	820	30	98
	восток	22,13	811	29,4	96
	центр	22,57	775	29,6	94
Оренбургская 21	восток	21,62	830	29,9	95
	центр	22,83	802	29,4	98
	запад	21,55	784	29	93
Безенчукская Янтарь	восток	22,45	789	30,1	90
	центр	23,82	803	30,2	94
	запад	23,32	798	30	94
Безенчукская 200	восток	23,45	830	31,7	95
	центр	21,4	830	30,7	96
	запад	24,53	825	28,2	95
Степь 3	восток	21,65	755	29,3	96
	центр	20,5	750	29,0	95
	запад	20,1	723	28,4	90

Продолжение таблицы 7

Сорт	Зона произрастания	Объем зерновки, куб.мм	Натура, г/л	Масса 1000 зерен, г	Стекловидность , %
Мягкие сорта					
Юго-Восточная 3	восток	22,65	781	30,2	98
	центр	23,54	780	30,7	95
	запад	24,32	772	31,6	96
Учитель	восток	23,85	745	30,5	93
	запад	22,94	768	29,3	90
	центр	21,56	798	29	95
Варяг	восток	21,26	803	31,3	98
	центр	20,54	800	30,7	95
	запад	21,7	811	31,4	97
Оренбургская 13	восток	21,87	820	31,6	94
	центр	21,64	812	30,5	95
	запад	21,49	815	31,4	95
Прохор	восток	23,8	820	30,4	94
	центр	23,58	804	29,5	93
	запад	23,7	800	30,1	93
Л 503	восток	25,7	790	32,5	95
	центр	25,4	786	32	92
	запад	24,53	775	32,2	91
Саратовская 42	восток	24,21	710	30,8	85
	центр	23,82	745	30,3	87
	запад	24,58	750	29,9	88



2.6.5.2 Определите отсутствие или наличие линейной корреляционной взаимосвязи между зараженностью зерна пшеницы спорами «картофельной» палочки (показатель выражается в единицах КОЕ/г – колониеобразующих единиц бактерий в 1 грамме зерна) и тепло- и влагообеспеченностью района произрастания (показатели суммы температур и коэффициента атмосферного увлажнения), стекловидности зерна (выражена в %). Для анализа используйте данные Оренбургского гос. центра агрохимической службы (таблица 8).

Таблица 8 – Входные данные для корреляционного анализа

№ образца зерна	Сумма температур, °С	Коэффициент атмосферного увлажнения	Обсемененность спорами «картофельной» палочки, КОЕ/г	Стекловидность, %
1	2270	0,22	400	98
2	2360	0,15	900	90
3	2350	0,25	500	67
4	2400	0,15	1100	69
5	2490	0,25	800	68
6	2610	0,15	1600	74
7	2670	0,22	1300	75
8	2665	0,12	2000	79
9	2730	0,16	1400	95
10	2770	0,12	2100	96
11	2790	0,25	1800	85
12	2770	0,15	2200	80
13	2900	0,25	1600	87
14	2970	0,15	2600	84
15	2240	0,25	380	89
16	2290	0,12	850	90
17	2450	0,25	510	95
18	2470	0,15	1090	96
19	2520	0,25	790	97
20	2540	0,15	1580	98
21	2580	0,22	1250	92
22	2640	0,12	1990	91
23	2630	0,15	1360	83
24	2675	0,12	2080	87
25	2850	0,22	1770	76
26	2900	0,12	2160	74

2.6.5.3 Провести статистическую обработку результатов антропометрических данных на примере группы людей.

2.6.5.4 Провести корреляционный анализ, используя в качестве исходных данных – значения церебрального индекса животных.

2.6.5.5 Провести корреляционный анализ, используя в качестве исходных данных – значения индекса массы группы людей.

2.6.5.6 Провести статистический анализ данных (вычислить описательные статистики, корреляции, экстраполяция - прогноз) на примере динамики курса валют за продолжительный период времени в связи с основными экономическими характеристиками.

2.6.5.7 В таблицах 9, 10 и 11 представлены данные анализа показателей качества зерна пшеницы. Необходимо найти основные статистические характеристики образцов зерна по годам урожая и по зонам произрастания, сравнить их, сделать выводы о преобладании влияния погодных или географических условий произрастания.

Таблица 9 - Показатели качества макаронных изделий из зерна пшеницы

Образец муки №	Показатели качества макаронных изделий			
	Коэффициент увеличения массы	Сухие вещества, перешедшие в воду при варке, %	Время варки до готовности, мин	Прочность сухих изделий на срез, Н
1	1,4	4,3	10,8	4,9
2	1,6	4,7	11,2	5,4
3	1,7	4,1	9,7	5,9
4	1,6	4,8	12,3	5,8
5	1,5	4,8	12,2	5,9
6	1,8	4,4	11,8	3,2
7	1,2	5,6	11,3	5,8
8	1,5	5,4	13	6,6
9	1,8	5,7	13,2	6,6
10	1,8	5,8	10,1	6,7
11	1,9	5,7	10,7	5,4
12	1,9	5	13,6	5,3
13	1,6	5	12,5	7,9
14	2,1	5,8	10,3	6,7
15	1,6	4,9	12,6	5,6
16	1,8	6	10,9	4,6
17	2,1	5,6	11,8	7,9
18	2	6,3	13,3	7,6
19	2,1	5,9	13	7,6
20	1,9	5,9	14,3	6,8
21	2,1	6	13,6	6,6
22	1,8	5,4	12,8	8,7

## Продолжение таблицы 9

Образец муки №	Показатели качества макаронных изделий			
	Коэффициент увеличения массы	Сухие вещества, перешедшие в воду при варке, %	Время варки до готовности, мин	Прочность сухих изделий на срез, Н
23	2,3	6,1	13,9	8,2
24	2,3	5,9	12,7	7
25	1,9	5,3	13	8,5
26	2,1	5,3	10,4	6,1
27	1,8	6,3	13,7	6
28	2,1	5,7	14,2	6,9
29	2,4	6	13,5	7,6
30	1,9	5,2	11,9	8,1
31	2,1	3,1	13	6,3
32	2,3	5,8	14,4	8,9
33	1,7	6,4	12,3	7,6
34	2,3	5,7	15,3	7
35	2,4	6	15,2	7,8
36	2,3	5,6	14,1	8,4
37	2,1	7	11,7	8,5
38	2	6,2	16,1	8,9
39	1,7	6,8	15	5,4
40	2,3	5,7	12,7	7
41	2,4	5,1	14,0	6,9

Таблица 10 – Данные урожайности зерна за 5 лет (в ц/га)

Сорт	Зона произрастания / Год урожая	2012	2013	2014	2015	2016
		1	2	3	4	5
Харьковская 3	восточная	15,3	15,5	16,1	16,8	14,7
	центральная	15	16,7	15,2	17,4	17,8
	западная	16,5	17,5	16,5	16,7	16,8
Оренбургская 10	восточная	16,1	17,4	16,6	17,6	16,3
	центральная	15,9	18,7	17,9	17,5	18,7
	западная	15,7	18,4	18,1	16,7	15,3
Оренбургская 21	восточная	17,2	17,4	17,3	16,9	18,4
	центральная	16,9	18,1	18,8	19,3	19
	западная	19	16,4	18,3	17,2	17,5
Безенчукская Янтарь	восточная	17,1	17,9	15,5	15,8	18,6
	центральная	17,9	19,3	18,7	17,7	17,1
	западная	18,9	18,1	16,4	18,6	19,3
Безенчукская 200	восточная	18	15,7	20	19,3	18,9
	центральная	18,1	19,3	19	19,3	19,3
	западная	17	17	17,1	19,2	19,6

Продолжение таблицы 10

1	2	3	4	5	6	7
Степь 3	восточная	15,8	20,5	16,3	16,8	18
	центральная	18,4	19,8	17,8	17,5	19,2
	западная	18,8	20,4	19,3	18,3	17,2
Юго- Восточная 3	восточная	17,4	16,6	16,8	16,9	21,8
	центральная	20,8	21,7	20,8	20,9	16,7
	западная	19,7	21,3	19,6	20	18,5
Учитель	восточная	21,5	19,5	19,6	20,7	13,4
	центральная	15,9	19,1	17,8	16,8	16,2
	западная	17,5	21,8	17,6	21,7	17,7
Варяг	восточная	17,4	19,3	19,5	17,1	20,7
	центральная	17,9	21,7	21,2	20,1	21,6
	западная	19,1	19,8	17,8	21,7	16,8
Оренбургская 13	восточная	18,4	21,3	17,4	17,3	17,7
	центральная	20,8	19,6	22,2	19,3	20,1
	западная	18,3	17,6	21,4	22,8	16,5
Прохор	восточная	21,2	20,9	14,3	22,3	20,3
	центральная	21,5	20,7	21,1	19,8	18,5
	западная	14,9	22,6	18,5	19,1	19
Л 503	восточная	16,7	17,7	21,3	16,7	15,8
	центральная	22,2	20,7	20,7	21,8	19,3
	западная	18,9	21,7	21,4	20	20,9
Саратовская 42	восточная	17,9	20,2	24,7	22,1	18,8
	центральная	17,1	20,8	19,5	22,6	21,4
	западная	19,4	15,1	20,6	20,7	15,6

Таблица 11 – Показатель седиментации зерна пшеницы за 5 лет (в см<sup>3</sup>)

Сорт	Зона произрастания / Год урожая	2012	2013	2014	2015	2016
		1	2	3	4	5
Харьковская 3	восточная	12	18	10	10	21
	центральная	10	11	19	14	24
	западная	10	14	25	16	25
Оренбургская 10	восточная	16	15	22	18	18
	центральная	17	11	14	21	27
	западная	20	16	18	20	29
Оренбургская 21	восточная	17	18	13	25	26
	центральная	18	17	12	19	17
	западная	16	26	27	22	28
Безенчукская Янтарь	восточная	18	21	24	26	30
	центральная	17	27	20	26	33
	западная	30	14	17	17	10
Безенчукская 200	восточная	26	20	23	19	21
	центральная	19	24	17	30	25
	западная	17	27	14	14	21
Степь 3	восточная	22	14	20	15	24
	центральная	30	23	16	29	25
	западная	20	25	19	26	18
Юго- Восточная 3	восточная	37	16	36	18	14
	центральная	24	14	23	29	32
	западная	24	11	32	12	33
Учитель	восточная	34	20	27	26	30
	центральная	25	36	28	19	21
	западная	41	40	34	30	20

Продолжение таблицы 11

1	2	3	4	5	6	7
Варяг	восточная	18	11	10	38	19
	центральная	35	26	39	31	33
	западная	36	14	46	21	39
Оренбургская 13	восточная	19	27	30	25	33
	центральная	30	22	39	33	38
	западная	24	34	16	41	28
Прохор	восточная	26	48	30	20	16
	центральная	45	15	39	41	14
	западная	27	59	40	38	32
Л 503	восточная	28	20	15	21	33
	центральная	55	31	25	33	30
	западная	18	51	55	18	21
Саратовская 42	восточная	37	32	55	22	20
	центральная	24	21	36	43	19
	западная	24	15	23	28	33



## Список использованных источников

- 1 Сидняев, Н. И. Теория вероятностей и математическая статистика: учебник для бакалавров / Н. И. Сидняев. – Москва: Юрайт, ИД Юрайт, 2011. - 219 с.
- 2 Спирина, М. С. Теория вероятностей и математическая статистика: учебник для студ. учреждений сред. проф. образования / М. С. Спирина, П. А. Спирин. – Москва: ИЦ Академия, 2012. - 352 с.
- 3 Бычкова, С. Г. Социально-экономическая статистика: учебник для бакалавров / С. Г. Бычкова. – Москва: Юрайт, 2013. - 591 с.
- 4 Яковлева, А. В. Экономическая статистика: учебное пособие / А. В. Яковлева. – Москва: ИЦ РИОР, 2013. - 95 с.
- 5 Балдин, К. В. Общая теория статистики: учебное пособие / К. В. Балдин, А. В. Рукусуев. – Москва: Дашков и К, 2012. - 312 с.
- 6 Батракова, Л. Г. Теория статистики: учебное пособие / Л. Г. Батракова. – Москва: КноРус, 2013. - 528 с.
- 7 Громько, Г. Л. Теория статистики: Практикум / Г. Л. Громько. – Москва: НИЦ ИНФРА-М, 2013. - 238 с.
- 8 Ефимова, М. Р. Общая теория статистики: учебник / М. Р. Ефимова, Е. В. Петрова, В. Н. Румянцев. – Москва: ИНФРА-М, 2013. - 416 с.
- 9 Лысенко, С. Н. Общая теория статистики: учебное пособие / С. Н. Лысенко, И. А. Дмитриева. – Москва: ИД ФОРУМ, НИЦ ИНФРА-М, 2013. - 208 с.
- 10 Костин, В. Н. Теория эксперимента: учебное пособие / В. Н. Костин, В. В. Паничев; М-во образования и науки Рос. Федерации, Федер. гос. бюджет. образоват. учреждение высш. проф. образования «Оренбург. гос. ун-т», Каф. прогр. обеспечения вычисл. техники и автоматизир. систем. - Оренбург: Университет, 2014. - 212 с. : табл. - Библиогр.: с. 207-208. - Прил.: с. 209-212. - ISBN 978-5-4417-0415-1.
- 11 Килов, А. С. Планирование экспериментов и обработка экспериментальных данных : методические указания к практическому занятию для

студентов, обучающихся по программам высшего профессионального образования по направлению подготовки 150700.62 Машиностроение / А. С. Килов; М-во образования и науки Рос. Федерации, Федер. гос. бюджет. образоват. учреждение высш. проф. образования «Оренбург. гос. ун-т», Каф. материаловедения и технологии материалов. - Оренбург: ОГУ, 2014. - 35 с.

12 Годин, А. М. Статистика: учебник / А. М. Годин. – Москва: Дашков и К°, 2012. – 451 с.

13 Елисеева, И. И. Статистика: учебник для бакалавров / И. И. Елисеева. – Москва: Юрайт: ИД Юрайт, 2011. – 565 с.

14 Ниворожкина, Л. И. Статистика: учебник для бакалавров / Л. И. Ниворожкина. – Москва: Дашков и К°: Наука-Спектр, 2011. – 415 с.

15 Тумасян, А. А. Статистика промышленности: учебное пособие / А. А. Тумасян, Л. И. Василевская. – Минск: Новое знание, 2012. – 429 с.

16 Гмурман, В. Е. Теория вероятностей и математическая статистика: учебное пособие для бакалавров / В. Е. Гмурман. – Москва: Юрайт, 2013. - 479 с.

17 Горлач, Б. А. Теория вероятностей и математическая статистика: учебное пособие / Б. А. Горлач. – Санкт-Петербург: Лань, 2013. - 320 с.

18 Калинина, В. Н. Теория вероятностей и математическая статистика: учебник для бакалавров / В. Н. Калинина. – Москва: Юрайт, 2013. - 472 с.

19 Кобзарь, А. И. Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. – Москва: ФИЗМАТЛИТ, 2012. - 816 с.

20 Колемаев, В. А. Теория вероятностей и математическая статистика: учебник / В. А. Колемаев, В. Н. Калинина. – Москва: КноРус, 2013. - 376 с.

21 Кочетков, Е. С. Теория вероятностей и математическая статистика: учебник / Е. С. Кочетков, С. О. Смерчинская, В. В. Соколов. – Москва: Форум, НИЦ ИНФРА-М, 2013. - 240 с.

22 Краснов, М. Л. Вся высшая математика. Теория вероятностей. Математическая статистика. Теория игр: учебник / М. Л. Краснов, А. И. Киселев, Г. И. Макаренко. – Москва: ЛКИ, 2013. - 296 с.

23 Кремер, Н. Ш. Теория вероятностей и математическая статистика: учебник для студентов вузов / Н. Ш. Кремер. – Москва: ЮНИТИ-ДАНА, 2012. - 551 с.

24 Кричевец, А. Н. Математическая статистика для психологов: учебник для студ. учреждений высш. проф. образования / А. Н. Кричевец, А. А. Корнеев, Е. И. Рассказова. – Москва: ИЦ Академия, 2012. - 400 с.

25 Семенов, В. А. Теория вероятностей и математическая статистика: учебное пособие / В. А. Семенов. - Санкт-Петербург: Питер, 2013. - 192 с.