

Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«КАЛИНИНГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ»

А. В. Чернова, М. П. Белова

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОЕКТИРОВАНИЯ  
ПРОДУКТОВ ПИТАНИЯ

Учебно-методическое пособие по лабораторным работам для студентов,  
обучающихся в магистратуре по направлениям подготовки:  
19.04.02 «Продукты питания из растительного сырья»,  
19.04.03 «Продукты питания животного происхождения»,  
19.04.04 «Технология продукции и организация общественного питания»  
(очная, заочная формы обучения)

Калининград  
Издательство ФГБОУ ВО «КГТУ»  
2021

УДК 664:51-7

Рецензент:

декан механико-технологического факультета, зав. кафедрой технологии продуктов питания ФГБОУ ВО «Калининградский государственный технический университет», канд. техн. наук, доцент И. М. Титова

**Чернова А. В., Белова М. П.** Математическое моделирование проектирования продуктов питания: учебно-методическое пособие по лабораторным работам / А. В. Чернова, М. П. Белова – Калининград: Издательство ФГБОУ ВО «КГТУ», 2021. – 74 с.

Учебно-методическое пособие содержит основы математической статистики и теории вероятностей, необходимые для формирования умений и навыков в их использовании при разработке нового ассортимента продуктов питания и технологий с заданными составом и свойствами.

Пособие предназначено для студентов вузов, обучающихся в магистратуре по направлениям подготовки «Продукты питания животного происхождения», «Продукты питания из растительного сырья», «Технология продукции и организация общественного питания» очной и заочной форм обучения.

Ил. 35, табл. 25, список лит. – 8 наименований, прилож. 5

Учебно-методическое пособие рассмотрено и одобрено кафедрой технологии продуктов питания ФГБОУ ВО «Калининградский государственный технический университет» 30 августа 2021 г., протокол № 01

Учебно-методическое пособие рекомендовано к изданию методической комиссией механико-технологического факультета ФГБОУ ВО «Калининградский государственный технический университет» 31 августа 2021 г., протокол № 01

УДК 664:51-7

© Федеральное государственное  
бюджетное образовательное  
учреждение высшего образования  
«Калининградский государственный  
технический университет», 2021 г.  
© Чернова А. В., Белова М. П., 2021 г.

## Оглавление

Лабораторная работа № 1. Исследование вероятностей исхода эксперимента, распределений случайных величин.....	4
Лабораторная работа № 2. Построение матрицы планирования полного факторного эксперимента .....	15
Лабораторная работа № 3. Построение матрицы планирования дробного факторного эксперимента .....	29
Лабораторная работа № 4. Моделирование рецептур методом линейного программирования.....	32
Лабораторная работа № 5. Обработка результатов прямых и косвенных измерений, проверка однородности независимых выборок и воспроизводимости опытов в MS Excel ....	40
Лабораторная работа № 6. Построение и анализ регрессии в MS Excel .....	49
Лабораторная работа № 7. Нахождение оптимальных параметров технологических процессов методом крутого восхождения по поверхности отклика.....	58
Лабораторная работа № 8. Прогнозирование сроков хранения при помощи обобщенной функции желательности Харрингтона.....	62
Список литературы .....	67
ПРИЛОЖЕНИЯ.....	68
Приложение А. Квантили распределения $\chi^2$ Пирсона .....	68
Приложение Б. Квантили распределения Стьюдента .....	69
Приложение В. Отбор промахов по критерию Шовене .....	69
Приложение Г. Квантили распределения Фишера при уровне значимости $\alpha = 0,05$ .....	70
Приложение Д. Квантили распределения Кохрена при уровне значимости $\alpha = 0,05$ .....	71

## **Лабораторная работа № 1**

### **Исследование вероятностей исхода эксперимента, распределений случайных величин**

**Цель работы** - ознакомиться с основными понятиями теории вероятностей и математической статистики, получить навыки их использования в решении задач математического моделирования.

**Задание:**

1. Изучить теоретический материал.
2. Решить задачи, используя Microsoft Office Excel (MS Excel).

#### **Теоретический материал, методические указания и примеры выполнения заданий**

Теория вероятностей и математическая статистика – основа вероятностно-статистических методов обработки данных. Для принятия решений необходимо выразить рассматриваемые задачи в терминах вероятностно-статистических моделей. Применение конкретного вероятностно-статистического метода состоит из трех этапов:

- переход от технологической реальности к абстрактной математико-статистической схеме, т. е. построение вероятностной модели системы технологического процесса / продукта;
- проведение расчетов и получение выводов математическими средствами в рамках вероятностной модели;
- интерпретация математико-статистических выводов применительно к реальной ситуации и принятие соответствующего решения (например, о соответствии или несоответствии качества продукции установленным требованиям, необходимости изменения технологического процесса, рецептуры, заключение о доле дефектных единиц продукции в партии).

Понятия и методы теории вероятностей широко используются в процессе проектирования новой пищевой продукции, а также управления качеством и безопасностью технологического процесса ее производства. Например, чтобы решить, соответствует или не соответствует контролируемая партия продукции установленным требованиям, из нее отбирается выборка. По результатам контроля выборки делается заключение обо всей партии. В этом случае очень важно избежать субъективизма при формировании выборки, т. е. необходимо, чтобы каждая единица продукции в контролируемой партии имела одинаковую вероятность быть отобранной в выборку. В производственных условиях отбор единиц продукции в выборку осуществляют по рекомендациям, выведенным с использованием математической статистики, или по специальным таблицам / генераторам случайных чисел. На основе методов математической статистики разрабатываются правила и планы статистического контроля процессов, направленные на своевременное обнаружение ошибок технологических процессов и принятие мер к предотвращению выпуска продукции, не

соответствующей установленным требованиям. Лишь те методы обработки данных являются доказательными, которые опираются на вероятностные модели соответствующих реальных явлений и процессов. Соответствие вероятностной модели реальности, т. е. ее адекватность, обосновывают, в частности, с помощью статистических методов проверки гипотез.

В данном курсе изучается прикладная математическая статистика. С помощью ее методов проводится статистический анализ точности и стабильности технологических процессов и статистическая оценка качества. К специфическим методам относятся методы статистического приемочного контроля качества продукции, статистического регулирования технологических процессов, оценки и контроля надежности.

Особенно важно применять статистические методы на начальном этапе жизненного цикла продукции, т. е. на этапе разработки нового ассортимента продукции, технического задания. Это объясняется ограниченностью информации, доступной на начальном этапе, и необходимостью прогнозирования технических возможностей и экономической ситуации на будущее. Статистические методы должны также применяться на всех этапах проведения экспериментов.

Исходное понятие при построении вероятностных моделей – опыт (испытание). Его называют элементарным событием. Совокупность всех возможных исходов опыта, т. е. всех элементарных событий, называется пространством элементарных событий. С математической точки зрения пространство всех элементарных событий, возможных в опыте, – это множество, а элементарные события – его элементы.

Основное понятие теории вероятностей – понятие вероятности события. Это мера возможности осуществления события. Вероятность события  $A$  будет обозначаться как  $P(A)$ . Вероятность выражается в долях единицы или в процентах. Вероятность события  $A$  – число от 0 до 1 (или от 0 до 100 %), которое представляет собой предел частоты реализации события  $A$  при неограниченном числе повторений одного и того же комплекса условий. Вероятность события  $A$ , состоящего из элементарных событий  $\omega_i$ , равна сумме вероятностей элементарных событий, входящих в  $A$ , т. е. определяется равенством (1.1):

$$P(A) = \sum_1^n P(\omega_i). \quad (1.1)$$

При практическом применении вероятностно-статистических методов принятия решений постоянно используется понятие *независимости*. Например, при применении статистических методов управления качеством продукции говорят о независимых измерениях значений контролируемых параметров  $y$  включенных в выборку единиц продукции, о независимости появления дефектов одного вида от появления дефектов другого вида. События  $A$  и  $B$  называются независимыми, если вероятность их совместного осуществления равна произведению вероятностей осуществления каждого из них в отдельности:  $P(AB) = P(A) \cdot P(B)$ . Осуществление или неосуществление одного события не должно влиять на осуществление или неосуществление другого.

Испытания называются независимыми, если любые два события  $A$  и  $B$ , из которых  $A$  определяется по исходу первого испытания, а  $B$  – по исходу второго, являются независимыми.

Вероятность наступления события  $A$ , при условии наступления события  $B$ , называется *условной вероятностью* и обозначается как  $P(A|B)$ .

Набор событий  $A_i$ , хотя бы одно из которых обязательно (с единичной вероятностью) наступит в результате испытания, называется полным. *Полная вероятность* вычисляется по формуле (1.2):

$$P(B) = \sum_{i=1}^n P(A_i) \times P(B|A_i). \quad (1.2)$$

Для проверки статистических гипотез также используется формула Байеса (1.3):

$$P(A_i|B) = \frac{P(A_i) \times P(B|A_i)}{P(B)}, \quad (1.3)$$

где  $P(A_i|B)$  – апостериорная (лат. - на основе опыта) оценка гипотезы;  $P(A_i)$  – априорная (лат. – до опыта) оценка гипотезы, вероятность независимого события  $A_i$ ;  $P(B|A_i)$  – вероятность зависящего от  $A_i$  события  $B$ .

**Пример 1.1.** На трех шкуроемных машинах обрабатывается филе сельди с кожей, поступающее после обработки на общий конвейер. Первая машина дает 2 % брака, вторая – 7 %, третья – 10 %. Производительность первой машины в 3 раза больше производительности второй, а третьей – в 2 раза меньше, чем второй.

а) Каков процент брака на конвейере?

б) Каковы доли каждой машины среди бракованных филе сельди на конвейере?

Возьмем с конвейера наудачу одно филе и рассмотрим событие  $A$  – филе повреждено в процессе обесшкуривания и подлежит отбраковке. Оно связано с гипотезами относительно того, где было обработано это филе:  $H_k$  – взятое наудачу филе обработано на  $k$ -й машине,  $k = 1, 2, 3$ .

Условные вероятности (в условии задачи они даны в форме процентов):

$$P(A|H_1) = 0,02; P(A|H_2) = 0,07; P(A|H_3) = 0,10.$$

Зависимости между производительностями машин следующие:

$$P(H_1) = 3 \cdot P(H_2); P(H_3) = 0,5 \cdot P(H_2),$$

а так как гипотезы образуют полную группу, то  $P(H_1) + P(H_2) + P(H_3) = 1$ .

Решив полученную систему уравнений, найдем:  $P(H_1) = 6/9 = 2/3$ ;  $P(H_2) = 2/9$ ;  $P(H_3) = 1/9$ .

а) Полная вероятность того, что взятое наудачу с конвейера филе бракованное:

$$P(A) = \sum P(A|H_i) \cdot P(H_i) = 2/3 \cdot 0,02 + 2/9 \cdot 0,07 + 1/9 \cdot 0,1 = 0,04.$$

Другими словами, в массе филе без кожи, сходящего с конвейера, брак составляет 4 %.

б) Пусть известно, что взятое наудачу филе бракованное. Пользуясь формулой Байеса, найдем условные вероятности гипотез:

$$P(H_1|A) = \frac{P(A|H_1) \cdot P(H_1)}{P(A)} = \frac{2/3 \cdot 0,02}{0,04} = 0,33.$$

$$P(H_2|A) = \frac{P(A|H_2) \cdot P(H_2)}{P(A)} = \frac{2/9 \cdot 0,07}{0,04} = 0,39$$

$$P(H_3|A) = \frac{P(A|H_3) \cdot P(H_3)}{P(A)} = \frac{1/9 \cdot 0,1}{0,04} = 0,28.$$

Таким образом, в общей массе поврежденного филе на конвейере доля первой машины составляет 33 %, второй – 39 %, третьей – 28 %.

Случайная величина – это величина, значение которой зависит от случая, т. е. от элементарного события. Она в результате испытаний, проводимых в одних и тех же условиях, принимает различные значения, зависящие от не учитываемых случайных факторов.

*Законом распределения дискретной случайной величины* называется задаваемое в любой форме соответствие между возможными значениями случайной величины и их вероятностями. Закон распределения случайной величины  $X$  может быть записан в виде таблицы, которую называют *рядом распределения* дискретной случайной величины:

$X$	$x_1$	$x_2$	...	$x_n$
$p$	$p_1$	$p_2$	...	$p_n$

*Математическим ожиданием* случайной величины  $X$  называется величина  $M(X)$ , вычисляемая по формуле (1.4):

$$M(X) = \sum_{i=1}^n X_i \cdot P(X_i), \quad (1.4)$$

т. е. математическое ожидание случайной величины – это взвешенная сумма значений случайной величины с весами, равными вероятностям соответствующих элементарных событий. Математическое ожидание показывает, вокруг какой точки группируются значения случайной величины.

*Дисперсией* случайной величины  $X$  называется величина  $D(X)$ , которая также обозначается как  $\sigma^2$  и вычисляется по формуле (1.5):

$$D(X) = M(X^2) - [M(X)]^2. \quad (1.5)$$

На практике при анализе выборок математическое ожидание, как правило, неизвестно, поэтому вместо него используют оценку – среднее арифметическое и применяют следующую формулу дисперсии (1.6):

$$\delta^2(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}, \quad (1.6)$$

где  $\delta^2(X)$  – выборочная дисперсия, рассчитанная по данным наблюдений (примерно до 30-ти наблюдений);  $\bar{X}$  – среднее арифметическое по выборке;  $n$  – количество значений в анализируемой совокупности данных.

Соответствующая функция в программе MS Excel - ДИСП.Г (для  $n \geq 30$ ) и ДИСП.В (для  $n < 30$ ) – рисунок 1.1.

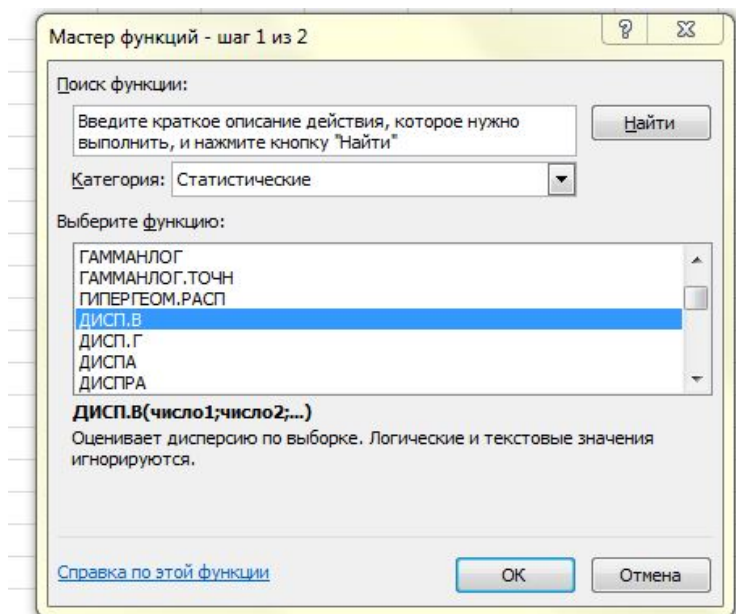


Рисунок 1.1 – Дисперсия в MS Excel

Дисперсия в статистике - очень важный показатель, но ее обычно используют не в чистом виде, а для дальнейших расчетов.

*Среднеквадратичное отклонение* – это корень из дисперсии. Этот показатель, его также называют стандартным отклонением, обозначают как  $S$  и рассчитывают по формуле (1.7) по генеральной совокупности (термин «генеральная совокупность» используется, когда речь идет о большой, но конечной совокупности изучаемых единиц, например, партия продукции), и как  $\sigma$  и рассчитывают по формуле (1.8) по выборке:

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}} \quad (1.7)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}. \quad (1.8)$$

В MS Excel для среднеквадратичного отклонения есть готовые функции: СТАНДОТКЛОН.Г и СТАНДОТКЛОН.В (по генеральной и выборочной совокупности соответственно) – рисунок 1.2.

Стандартное и среднеквадратичное отклонение - синонимы. Среднеквадратичное отклонение также характеризует меру рассеяния данных, но, в отличие от дисперсии, его можно сравнивать с исходными данными, так как единицы измерения у них одинаковые. Свойства данного показателя хорошо изучены, он участвует во многих статистических расчетах. С его помощью устанавливают степень точности различных оценок и прогнозов.



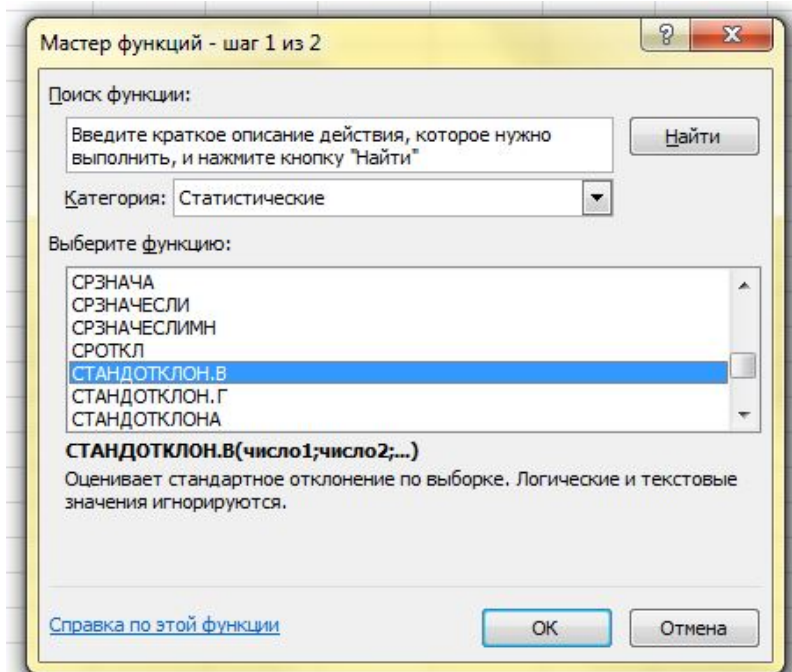


Рисунок 1.2 – Среднеквадратичное отклонение в MS Excel

Для проверки статистических гипотез часто используют *теорему Бернулли*: пусть  $m$  – число наступлений события  $A$  в  $k$  независимых (попарно) испытаниях и  $p$  – вероятность наступления события  $A$  в каждом испытании. Тогда при любом  $\varepsilon > 0$  справедливо неравенство (1.9):

$$P \left\{ \left| \frac{m}{k} - p \right| \geq \varepsilon \right\} \leq \frac{p \times (1-p)}{k \times \varepsilon^2}. \quad (1.9)$$

**Пример 1.2.** Пусть из 100000 единиц продукции 30000 оказались дефектными. Согласуется ли это с гипотезой о том, что вероятность дефектности равна 0,23?

Считаем, что испытания (попарно) независимы и что в каждом испытании вероятность того, что единица продукции является дефектной, равна  $p$ . Для проверки гипотезы воспользуемся неравенством (2.9). В рассматриваемом случае  $k = 100000$ ,  $m = 30000$ ,  $m/k = 0,3$ ,  $p = 0,23$ ,  $m/k - p = = 0,07$ . Оценим вероятность того, что  $m/k$  отличается от  $p$  так же, как в рассматриваемом случае, или больше, т. е. оценим вероятность выполнения неравенства  $|m/k - 0,23| > 0,07$ . Положим в неравенстве (2.9)  $p = 0,23$ ,  $\varepsilon = 0,07$ . Тогда:

$$P \left\{ \left| \frac{30000}{100000} - 0,23 \right| \geq \varepsilon \right\} \leq \frac{0,23 \times 0,77}{0,0049 \times 100000} \sim 0,00036.$$

Из расчета видно, что неравенство не выполняется, а, значит, гипотезу необходимо отвергнуть. Таким образом, вероятность дефектности не равна 0,23.

Одна из основных характеристик метода проверки гипотезы – *уровень значимости*, т. е. вероятность отвергнуть проверяемую гипотезу, когда она верна. Уровень значимости также называют надежностью и обозначают  $\alpha$ . Если вероятность меньше  $\alpha$ , гипотезу отвергают. Если эта вероятность больше или равна  $\alpha$ , гипотезу принимают.

Обычно в вероятностно-статистических методах принятия решений (в том числе и при лабораторных испытаниях продуктов питания) выбирают  $\alpha = 0,05$  (или 95 %), значительно реже  $\alpha = 0,01$  или  $\alpha = 0,1$ , в зависимости от конкретной практической ситуации.

Статистическое исследование, как правило, начинается с анализа закона распределения исследуемой случайной величины. *Распределение числовой случайной величины* – это функция, которая однозначно определяет вероятность того, что случайная величина принимает заданное значение или принадлежит к некоторому заданному интервалу. *Функцией распределения* случайной величины  $X$  называется функция  $F(x)$ , определенная на всей числовой оси следующим образом:  $F(x) = P(X < x)$ , т. е.  $F(x)$  – вероятность того, что случайная величина  $X$  примет значение меньше, чем  $x$ . Используемые в прикладных исследованиях функции распределения бывают либо дискретными, либо непрерывными, либо их комбинациями.

Функцию распределения можно представить графически. Для дискретной случайной величины график имеет ступенчатый вид (рисунок 1.3).

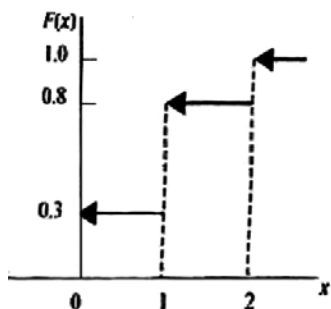


Рисунок 1.3 - График функции распределения дискретной случайной величины

График функции распределения непрерывной случайной величины представляет собой непрерывную кривую (рисунок 1.4).

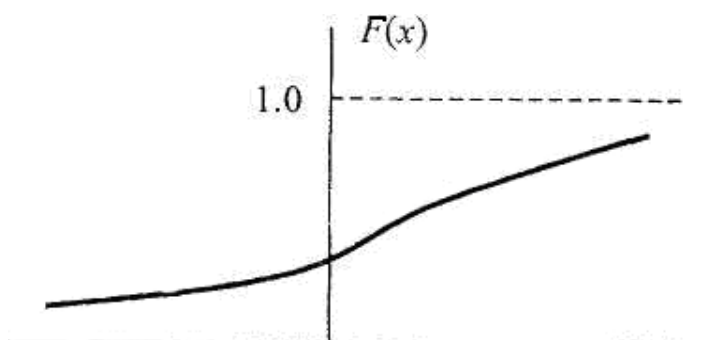


Рисунок 1.4 - График функции распределения непрерывной случайной величины

Смешанные функции распределения встречаются, в частности, тогда, когда наблюдения в какой-то момент прекращаются.

Большое значение имеет *нормальное распределение* (рисунок 1.5). Если результат наблюдения является суммой многих случайных слабо взаимозависимых величин, каждая из которых вносит малый вклад относительно общей суммы, то при увеличении числа слагаемых распределение центрированного и нормированного результата стремится к нормальному. Данное распределение лежит в основе измерений, разработки тестовых шкал, методов проверки гипотез. Чем больше объем выборки, тем достовернее полученное эмпирическое распределение приближается к нормальному. Эмпирическим путем было установлено, что многие биологические параметры распределены подобным образом (рост, масса).

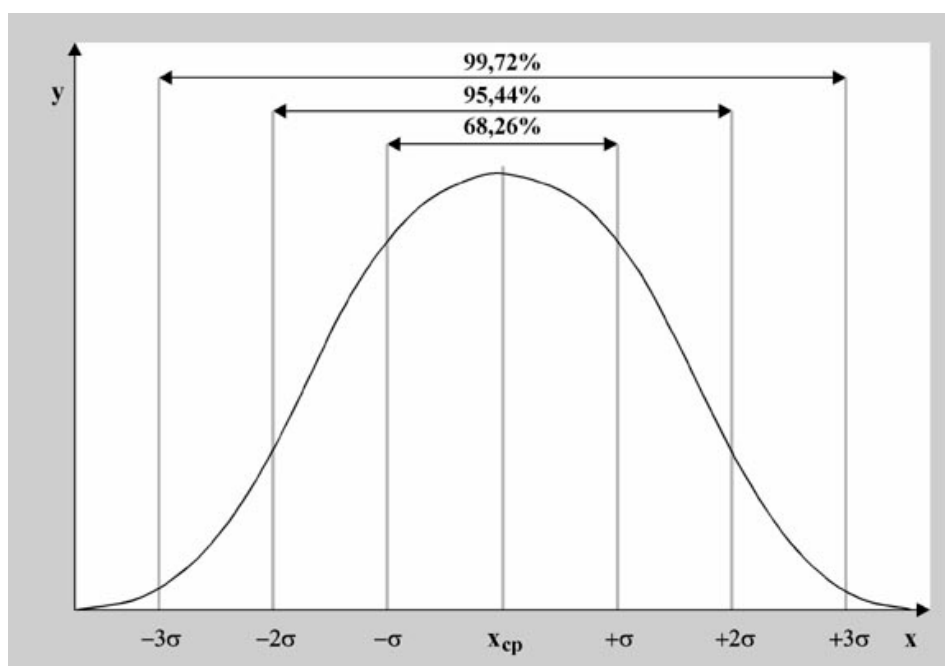


Рисунок 1.5 - График нормального распределения

При анализе достоверности проведенных измерений часто используют *правило трех сигм*, которое основано на законе нормального распределения. Вероятность того, что случайная величина отклонится от своего математического ожидания на величину, большую, чем утроенное среднее квадратичное отклонение, практически равна нулю. Значит, если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратичного отклонения. Это и есть правило трех сигм.

С помощью нормального распределения определяются три распределения, которые часто используют в статистической обработке данных экспериментов:

- Распределение Пирсона  $\chi^2$  (кси-квадрат) используют при оценивании дисперсии (с помощью доверительного интервала), при проверке гипотез согласия, однородности, независимости, прежде всего, для качественных переменных, принимающих конечное число значений.

– Распределение  $t$  Стьюдента применяют при оценивании математического ожидания, прогнозного значения и других характеристик с помощью доверительных интервалов, по проверке гипотез о значениях математических ожиданий, коэффициентов регрессионной зависимости, гипотез однородности выборок.

– Распределение Фишера используют при проверке гипотез об адекватности модели в регрессионном анализе, о равенстве дисперсий и в других задачах прикладной статистики.

*Экспоненциальное распределение* характеризует «поток событий», т. е. последовательность событий, происходящих одно за другим в какие-то моменты времени. Примерами могут служить: поток отказов оборудования в технологической цепочке; остановок производственной линии; поток отбракованных единиц при испытаниях продукции; поток покупателей. Данное распределение применяют в теории надежности, когда определяют вероятность безотказной работы за определенное время. Экспоненциальные распределения – частный случай *распределений Вейбулла - Гнеденко*. Весь график функции данного распределения можно разбить на три периода. На первом из них функция имеет высокие значения и явную тенденцию к убыванию (чаще всего она монотонно убывает). Это можно объяснить наличием в рассматриваемой партии продукции единиц продукции с явными и скрытыми дефектами, которые приводят к относительно быстрому выходу из строя этих единиц. Первый период называют «периодом обкатки» (например, период выпуска пробной партии новой продукции). Затем наступает период нормального течения технологического процесса, характеризующийся приблизительно постоянной и сравнительно низкой интенсивностью отказов / брака. Этот период можно также назвать сроком сохранения высоких потребительских характеристик продукта. Природа брака в этот период носит внезапный характер (аварии, ошибки работников, ненадлежащие условия хранения продуктов). Последний период - период старения и износа. Природа отказов / брака в этот период - в необратимых физико-механических и химических изменениях сырья и материалов, приводящих к прогрессирующему ухудшению качества изготавливаемой и хранимой продукции. Наступление данного периода указывает на окончание срока годности продукта или на необходимость замены технологического оборудования, используемого при производстве данного продукта.

Наиболее часто используют три семейства *дискретных распределений* - биномиальных, гипергеометрических и Пуассона, а также некоторые другие семейства - геометрических, отрицательных биномиальных, отрицательных гипергеометрических.

*Биномиальное распределение* имеет место при независимых испытаниях, в каждом из которых с вероятностью  $p$  появляется событие  $A$ . Они применяются при анализе данных выборочных исследований, в частности, при изучении предпочтений потребителей, выборочном контроле качества продукции по планам одноступенчатого контроля. Данное распределение

может быть представлено в виде гистограммы отрезков, ломаной линии или точечной диаграммы. Биномиальное распределение не обязательно должно быть симметричным. Формула Бернулли (1.10) позволяет находить вероятность появления события  $A$  при независимых испытаниях.

$$P_{k,n} = C_n^k \times p^k \times q^{n-k}, \quad (1.10)$$

где  $p$  – вероятность того, что случайная величина приобретет определенное значение  $X$  во всех испытаниях;  $q = 1 - p$ ;  $n$  – число независимых испытаний;  $k$  – сколько раз должно наступить событие (случайная величина приобретет определенное значение  $X$ );  $P$  – вероятность того, что случайная величина приобретет определенное значение  $X$   $k$ -раз в  $n$  испытаниях;  $C_n^k$  – биномиальный коэффициент, который вычисляется по формуле (1.11):

$$C_n^k = \frac{n!}{k! \times (n-k)!}. \quad (1.11)$$

**Пример 1.3.** В партии однотипных деталей стандартные составляют 97 %. Наугад из партии берут 400 деталей. Построить функцию распределения вероятности появления стандартной детали.

Случайная величина  $X$  может принимать значения  $X = k = 0, 1, 2, \dots, 400$ . Вероятности возможных значений для данной задачи определяются по формуле Бернулли (2.10) и составляют:

$$P_{k,400} = C_{400}^k \times p^k \times q^{400-k} = C_{400}^k \times 0,97^k \times 0,03^{400-k},$$

где  $p = 0,97$  – вероятность появления стандартной детали;  $q = 1 - p = 1 - 0,97 = 0,03$  – вероятность появления нестандартной детали.

*Распределение Пуассона* (2.12) часто используется, когда имеют дело с числом событий, появляющихся в промежутке времени или пространства, например: число остановок станков в неделю. Распределение Пуассона является предельным случаем биномиального распределения, когда вероятность  $p$  осуществления события мала, но число испытаний  $n$  велико (т. е. описывает редкие события).

Распределение Пуассона используется при анализе результатов выборочных маркетинговых обследований потребителей, расчете оперативных характеристик планов статистического приемочного контроля в случае малых значений приемочного уровня дефектности, для описания числа разладов статистически управляемого технологического процесса в единицу времени.

$$P(x) = \frac{\lambda^x}{x!} \times e^{-\lambda}, \quad (1.12)$$

где  $x$  – число появлений события в  $n$  независимых испытаниях;  $\lambda = np$  – среднее число появлений события в  $n$  испытаниях.

*Гипергеометрическое распределение* имеет место при выборочном контроле конечной совокупности объектов объема  $N$  по альтернативному признаку. Каждый контролируемый объект классифицируется либо как обладающий признаком  $A$ , либо как не обладающий этим признаком. Гипергеометрическое распределение имеет случайная величина  $Y$ , равная числу объектов, обладающих признаком  $A$  в случайной выборке объема  $n$ , где  $n < N$ . Например, число  $Y$  дефектных единиц продукции в случайной выборке объема  $n$  из партии объема  $N$  имеет гипергеометрическое распределение, если  $n < N$ .

Последовательно проводится несколько независимых испытаний до появления некоторого события  $A$ , вероятность которого в каждом испытании одна и та же и равна  $p$ . Тогда число  $X$  произведённых испытаний есть дискретная случайная величина с геометрическим распределением вероятности. Типичный пример представлен таблицей 1.1: осуществлена выработка партии продукции из  $N$  объектов, из которых  $D$  имеют дефект. Гипергеометрическое распределение описывает вероятность того, что в выборке из  $n$  различных объектов ровно  $k$  объектов являются бракованными.

Таблица 1.1 – Гипергеометрическое распределение

Объект	Отобранные	Неотобранные	Всего
С исследуемым признаком	$k$	$D - k$	$D$
Без исследуемого признака	$n - k$	$N + k - n - D$	$N - D$
Всего	$n$	$N - n$	$N$

Пусть  $Y$  - случайная величина, равная количеству выбранных элементов, обладающих нужным свойством. Тогда функция вероятности имеет вид (1.13):

$$P(Y = k) = \frac{C_D^k \times C_{N-D}^{n-k}}{C_N^n}, \quad (1.13)$$

где  $C_D^k, C_{N-D}^{n-k}, C_N^n$  – биномиальные коэффициенты (см. формулу (1.11)).

**Задание 1.1.** В магазине представлены консервы из рыбы, мяса, морепродуктов и овощей. Вероятность того, что покупатель купит консервы из мяса, равна 0,35; из рыбы – 0,23; морепродуктов – 0,22, овощей – 0,3. Какова вероятность того, что покупатель купит консервы из животного сырья?

**Задание 1.2.** На предприятии работают три копильные печи. Первая копильная печь вырабатывает  $X_1$  % готовой продукции от суточной производительности, вторая  $X_2$  %, третья –  $X_3$  %. При этом первая копильная печь дает  $Y_1$  % нестандартных единиц продукции, вторая –  $Y_2$  %, третья –  $Y_3$  %. Определить вероятность того, что наудачу взятая со склада единица продукции:

- не соответствует стандарту,
- соответствует стандарту.

Номер варианта	$X_1, X_2, X_3$	$Y_1, Y_2, Y_3$
1	2	3
1	50, 20, 30	2,5; 2; 1
2	10,60, 30	3,2; 1,4; 0,8
3	20, 20, 60	2; 1; 0,5
4	30, 30, 40	3,8; 0,4; 2
5	25, 25, 50	1; 2,3; 5
6	75, 5, 20	7; 1; 0,8
7	65, 15, 20	3,1; 2,7; 5
8	30, 40, 30	0,1; 5; 2
9	50, 10, 40	4; 2,3; 1,6

**Задание 1.3.** Найти: а) математическое ожидание, б) дисперсию, в) среднее квадратичное отклонение дискретной случайной величины  $X$  по известному закону ее распределения, заданному таблично. Использовать формулы как для генеральной совокупности, так и для выборки. Сравнить результаты.

X	8	4	6	5
P	0,2	0,5	0,2	0,1

**Задание 1.4.** Два завода производят хлебопекарные печи в объеме 3:7. Первый завод производит 95 % печей без дефекта, второй – 90 %. Производственная компания покупает пару печей. Построить ряд распределения вероятности покупки печей без дефекта.

**Задание 1.5.** Магазин получил 1000 бутылок минеральной воды. Вероятность того, что при перевозке бутылка окажется разбитой, равна 0,003. Найти вероятности того, что магазин получит две разбитые бутылки.

**Задание 1.6.** В партии из 10 упаковок пресервов имеется 8 стандартных. Наудачу отобраны две упаковки. Составить закон распределения числа стандартных упаковок среди отобранных.

### Контрольные вопросы

1. Что такое вероятность?
2. Какие события называют независимыми?
3. Как рассчитать полную вероятность события?
4. Что такое закон распределения случайной величины?
5. Какие бывают виды законов распределения?
6. Назовите правило трех сигм.
7. В каких случаях используется распределение Пуассона?
8. Чем отличается математическое ожидание от среднего арифметического?
9. Чем отличается дисперсия выборки и генеральной совокупности?

### Лабораторная работа № 2

#### Построение матрицы планирования полного факторного эксперимента

**Цель** – получить навыки планирования эксперимента, выбора факторов и составления матрицы планирования.

#### **Задание:**

1. Изучить теоретический материал и примеры построения параметрических схем технологического процесса или продукта.
2. Решить задачи априорного ранжирования и построения матрицы, используя Microsoft Office Excel (MS Excel).

## **Теоретический материал, методические указания и примеры выполнения заданий**

Любой метод научного познания основывается на *моделировании* – исследовании явлений, процессов, объектов в природе и обществе путем построения и изучения их моделей. Различают материальное и идеальное моделирование. Материальное (или физическое) моделирование – это экспериментальный метод. Реальному объекту ставится в соответствие его материальная копия – физическая модель, допускающая исследование в лабораторных условиях.

Идеальное моделирование – это теоретический метод. Реальный объект заменяется его знаковой, математической моделью (этот метод обычно называют математическим моделированием). Модель формируется в виде математических соотношений и проводится ее исследование – аналитическое или с применением компьютерных программ.

*Модель* – это объект-заменитель объекта-оригинала, обеспечивающий изучение некоторых интересующих исследователя свойств оригинала.

Математическими моделями называются комплексы математических зависимостей и знаковых логических выражений, отображающих существенные характеристики изучаемого явления. Примером являются алгебраические и дифференциальные уравнения, уравнения регрессии, задача симплекс-метода. Каждое применение математической модели должно быть обоснованным и осторожным: такая модель всегда является абстрактной идеализацией задачи, а значит, упрощающие предположения могут привести к тому, что модель не будет служить действительным представлением данной задачи.

Аналитическая математическая модель может исследоваться в технологии продуктов питания одним из следующих способов:

- 1) для получения в общем виде явных зависимостей между исходными параметрами, например, соотношения компонентов в рецептуре;
- 2) численно, когда получают числовые результаты при конкретных начальных данных;
- 3) качественно, когда можно найти некоторые качественные характеристики, например, оценить устойчивость органолептических показателей в процессе написания компьютерных программ, в которых воспроизводится технологический процесс во времени. При этом требуется воспроизводить входные воздействия в виде наборов чисел.

В ходе моделирования получается реализация выходного процесса, по которой в случае необходимости могут быть даны рекомендации по изменению технологических параметров.

Преимущества применения математического моделирования:

- Полностью или частично отпадает необходимость в длительном и трудоемком этапе изготовления лабораторного образца или установки; в



затратах сырья, материалов, необходимых для опытных образцов, макетов и установок; в измерительных приборах и оборудовании для исследований.

- Значительно сокращается время определения характеристик и время испытаний.

- Появляется возможность разработки продуктов, содержащих компоненты, характеристики которых известны, но самих компонентов у разработчика в настоящее время нет; имитировать воздействия, воспроизведение которых при натуральных испытаниях затруднено, требует сложного оборудования, сопряжено с опасностью для установки или экспериментатора, а иногда вообще невозможно; легко получать дополнительные характеристики объекта, которые сложно или невозможно получить с помощью измерительных приборов.

Метод математического моделирования, как любой численный метод, обладает существенным недостатком: решение всегда носит частный характер, соответствуя фиксированным значениям параметров системы и начальных условий. Поэтому для всестороннего анализа объекта (системы) приходится многократно моделировать его процесс функционирования, варьируя исходные данные.

В технологии продуктов питания математическое моделирование является одним из этапов исследования, который часто проводится вначале с целью выбора объектов, предмета исследований, отсеивания гипотез. Моделирование включает следующие стадии:

- выявление актуальной проблемы, поиску решения которой будет посвящено исследование;

- хранения.

Этапы моделирования:

- выбор и предварительное изучение объекта исследований, описание параметров, характеризующих объект;

- выбор параметров (факторов) объекта, которые будут изучаться в данном исследовании, путем проведения отсеивающего эксперимента или методом априорного ранжирования;

- выбор интервалов варьирования факторов;

- составление плана эксперимента (полного, дробного);

- постановка эксперимента и получение экспериментальных данных;

- проверка воспроизводимости опытов, обработка экспериментальных данных;

- составление математической регрессионно-корреляционной модели и анализ модели на адекватность;

- принятие решения об использовании модели;

- применение модели при прогнозировании зависимости характеристик выбранного объекта исследований от изменения технологических параметров.

В качестве объекта исследований может быть выбран процесс или продукт.

*Процесс* – последовательная смена состояний системы во времени. Каждое мгновенное состояние системы описывается набором чисел – параметров процесса и начальных условий. Математическая модель технологического процесса – это совокупность соотношений (формул, уравнений, неравенств), которые связывают характеристики процесса с параметрами системы, исходной информацией и внешними условиями. Процесс может быть разделен на ряд элементарных этапов, каждый из которых может быть описан своей моделью.

*Продукт* – это законченный результат производственной деятельности. Математическая модель может включать количественные и качественные характеристики продукта.

Принципы моделирования:

- Принцип информационной достаточности. При наличии полной информации об объекте исследования построение модели не имеет смысла, т.к. не принесет новой информации. Однако количество информации должно быть достаточным.

- Принцип осуществимости. Вероятность осуществимости моделируемого процесса должна быть больше нуля.

- Принцип множественности моделей. Объект исследования может быть описан с разных сторон множеством моделей. Исследователь выбирает только интересующую его сторону.

- Принцип параметризации. Модель описывается в виде системы, параметры которой неизвестны.

Перед исследователем могут стоять прямые и косвенные задачи моделирования. К прямым относят задачи, когда требуется определить, каковым будет проявление объекта, если известны его структура и свойства (задачи прогнозирования). К косвенным относят задачи, в ходе решения которых выясняют, каковы должны быть свойства объекта, чтобы получить оптимальный результат (задача оптимизации). При использовании методов моделирования следует помнить, что в основе моделирования лежит теория подобия, и любая модель не равна объекту-оригиналу, а значит, не может полностью описать исследуемый процесс или продукт. Модель будет учитывать и адекватно описывать лишь выбранные исследователем свойства объекта-оригинала.

*Эксперимент* - важнейшая часть научного исследования. Его результаты позволяют подтвердить или отклонить выдвинутую научную гипотезу. Эксперимент должен обеспечить точность, повторяемость, достоверность. Методика выполнения эксперимента требует хорошего теоретического обоснования.

*Планирование эксперимента* – это выбор плана эксперимента, удовлетворяющий заданным требованиям. Он состоит в выборе числа и условий протекания опытов, необходимых и достаточных для исследования объекта с заданной точностью. Планирование эксперимента должно обеспечить минимизацию числа опытов и ресурсов.

Виды экспериментов:

1. Элиминирующий эксперимент – это эксперимент, задача которого выявить или устранить влияние различных неоднородностей.

2. Сравнительный эксперимент - это эксперимент, заключающийся в оценке влияния каждого фактора на процесс / продукт, расположении их в ряд по степени влияния на исследуемый показатель процесса / продукта.

3. Отсеивающий эксперимент – это эксперимент, в ходе которого отбираются для исследования лишь те факторы, которые существенно влияют на процесс.

4. Экстремальный эксперимент направлен на выявление оптимальных режимов, оптимальных составов и оптимальных параметров.

5. Аппроксимирующий эксперимент – это составление математической модели процесса.

6. Экстраполирующий эксперимент – это эксперимент, задача которого – предсказать явление, оценить, как оно будет протекать в дальнейшем.

Любое исследование начинается с описания параметров и факторов, влияющих на изучаемый процесс или продукт. Для определения данных показателей составляют параметрическую схему исследуемого продукта или технологического процесса.

Все факторы делят на две группы:

– входные, которые, в свою очередь, разделяют на управляющие и возмущающие;

– выходные, которые могут быть управляемыми и наблюдаемыми.

Управляющие – это те факторы, которые непосредственно влияют на исследуемый процесс или продукт и с помощью целенаправленного изменения которых можно управлять технологическим процессом.

Возмущающие факторы также оказывают влияние на ход процесса, состояние продукта, но управлять ими невозможно. Часто они носят случайный характер.

Управляемые факторы – это те факторы, которые непосредственно контролируются в продукте или по ходу технологического процесса и по которым судят о надлежащем функционировании и эффективности процессов. Это и есть те факторы, которые изучаются в ходе исследований.

Наблюдаемые факторы также отражают исследуемый продукт или характер процесса, но либо они не интересуют исследователя в данный момент, либо на них никак невозможно повлиять в данной конкретной производственной ситуации. Все четыре группы факторов отражаются в параметрической схеме процесса или продукта. Для описания объекта исследования используют схему «черного» ящика (рисунок 2.1).

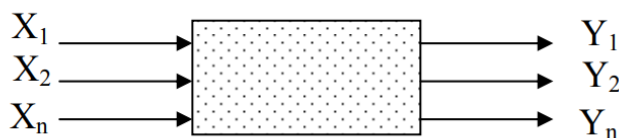


Рисунок 2.1 - Схема «черного» ящика

Стрелки справа изображают численные характеристики целей исследования, называемые еще критериями оптимизации, целевыми функциями, функциями отклика. Стрелки слева называют факторами, то есть способами воздействия на объект. Примеры схем представлены на рисунках 2.2-2.5.



Рисунок 2.2 – Параметрическая схема процесса ињектирования мяса

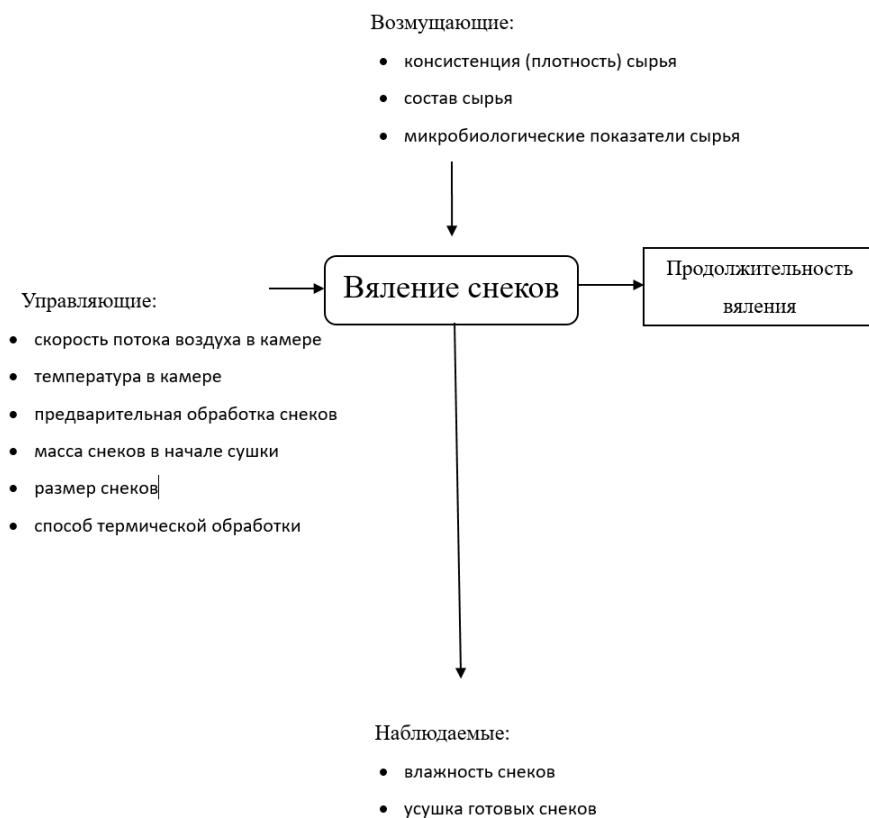


Рисунок 2.3 - Параметрическая схема процесса вяления снеков

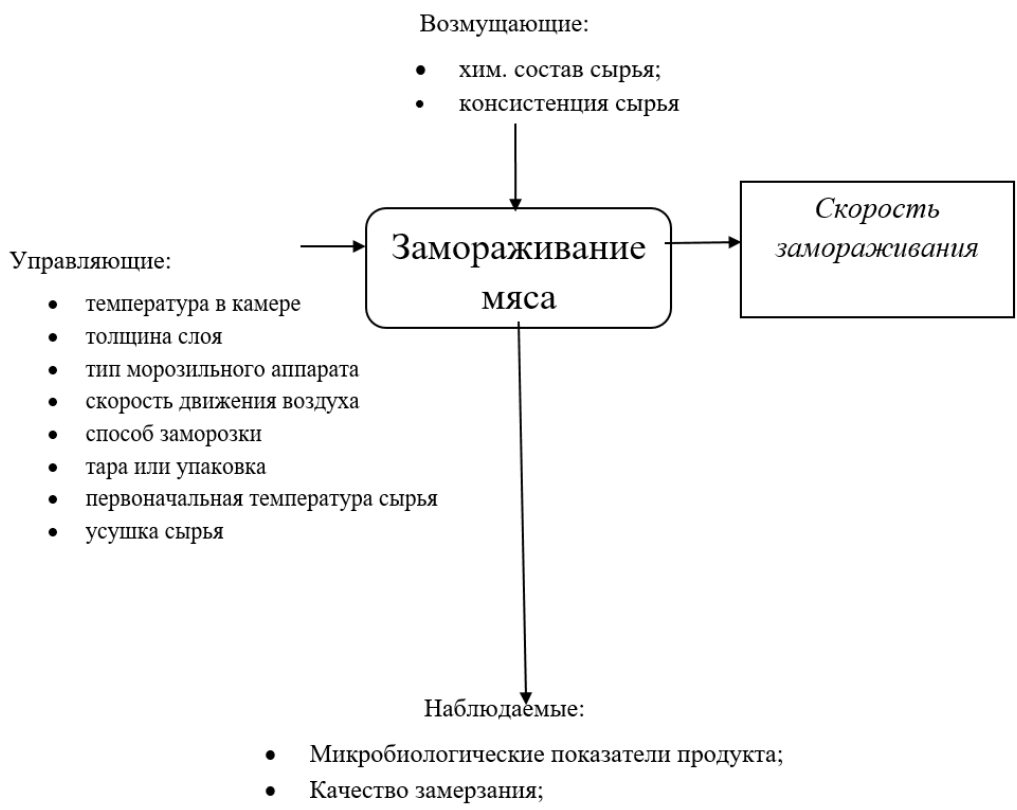


Рисунок 2.4 – Параметрическая схема процесса замораживания мяса



Рисунок 2.5 - Параметрическая схема процесса обжаривания картофеля для чипсов

В первую очередь необходимо проверить все описанные в параметрической схеме факторы на соответствие следующим требованиям:

– выбранные для исследования факторы должны быть из группы «управляющие», т. е. должна быть возможность установки и поддержания значения фактора на постоянном уровне в заданном интервале в течение всего опыта, а также возможность его измерения;

– факторы должны непосредственно воздействовать на объект исследования;

– точность измерения факторов должна отвечать поставленной задаче;

– факторы должны быть совместимыми. Это означает, что все их комбинации осуществимы и безопасны;

– факторы должны быть независимыми друг от друга.

Далее из факторов, отвечающих вышеуказанным требованиям, отбирают наиболее существенные. Для этого можно воспользоваться *методом априорного ранжирования факторов (методом экспертных оценок / ранговой корреляции)*. Особенность метода априорного ранжирования факторов заключается в том, что факторы, которые согласно априорной информации могут иметь существенное влияние, ранжируются в порядке убывания вносимого ими вклада. Вклад каждого фактора оценивается по величине ранга, который отведен ему исследователем при ранжировании всех факторов с учетом их предполагаемого влияния на параметры оптимизации.

При сборе мнений путем опроса специалистов каждому из них предлагается заполнить анкету, в которой перечислены факторы, их размерность и предполагаемые интервалы варьирования. Заполняя анкету, специалист определяет место факторов в ранжированном ряду. Самому важному фактору, по мнению эксперта, выставляется ранг 1, и далее ранг увеличивается по мере уменьшения важности рассматриваемых факторов. Одновременно он может включить дополнительные факторы или высказать мнение об изменении интервалов варьирования. Результаты опроса специалистов (или ранжирования по литературным данным) обрабатываются следующим образом.

Определяют среднее значение суммы рангов по формуле (2.1):

$$a_{cp} = 0,5 \cdot m \cdot (k + 1), \quad (2.1)$$

где  $m$  – количество экспертов;  $k$  – количество факторов.

Далее определяют сумму рангов  $\sum a_i$  отклонение от среднего  $d_i$  – формула (2.2) – и квадрат среднего отклонения  $d_i^2$  по каждому из факторов.

$$d_i = a_{cp} - \sum a_i. \quad (2.2)$$

Согласованность мнений экспертов определяют по коэффициенту конкордации  $W_k$  (формула (2.3)):

$$W_k = \frac{12 \cdot \sum d_i^2}{m^2 \cdot k \cdot (k^2 - 1)}. \quad (2.3)$$

Чем ближе коэффициент конкордации к 1, тем более согласованы мнения экспертов. Если коэффициент оказывается меньше 0,25, результаты данного

опроса нельзя принимать для дальнейшего исследования из-за слишком низкой согласованности мнений специалистов.

Проверка достоверности результатов проводится по критерию Пирсона  $\chi^2$  (формула (2.4)):

$$\chi^2 = m \cdot (k - 1) \cdot W_k. \quad (2.4)$$

Табличное значение (см. таблицу Приложения А) определяется в зависимости от чисел степеней свободы ( $f = k - 1$ ) для определенного уровня значимости (обычно 5 %). Если расчетное значение больше табличного, результаты достоверны.

В основе построения моделей технологических процессов лежит *теория многофакторного эксперимента*, разработанная Р. Фишером, которая опирается на изучение состояния и поведения объекта при одновременном изменении нескольких входных факторов. Тогда эксперимент – это реализация всех или части состояний, в которых может находиться объект исследования, т. е. это совокупность опытов.

*Уровнем фактора* называют его фиксированное значение. Каждому опыту соответствует точка факторного пространства, определяемая упорядоченной совокупностью численных значений факторов, соответствующих условиям проведения опыта.

Область факторного пространства, доступная для исследования, называется *областью эксперимента*. Зависимость отклика от управляемых факторов – это *функция отклика*. *Поверхность отклика* – геометрический образ функции отклика. При выборе области эксперимента должно учитываться следующее. Прежде всего, надо оценить границы областей определения факторов. При этом должны учитываться принципиальные ограничения для значений факторов, которые не могут быть нарушены ни при каких обстоятельствах. Например, если фактор – температура, то нижним пределом будет абсолютный ноль. Кроме этого принимаются во внимание ограничения, связанные с технико-экономическими соображениями (стоимостью сырья, дефицитностью отдельных компонентов, временем процесса) и с конкретными условиями проведения процесса, например, существующей аппаратурой, технологией. Выбор экспериментальной области факторного пространства связан с тщательным анализом априорной информации, в результате которого определяется комбинация (или несколько комбинаций) уровней факторов. Каждая комбинация является многомерной точкой в факторном пространстве. Ее рассматривают как исходную точку для построения плана эксперимента и называют *основным (нулевым) уровнем ( $X_0$ )*. Построение плана эксперимента сводится к выбору экспериментальных точек, симметричных относительно нулевого уровня. Если координаты наилучшей точки неизвестны, но есть сведения о некоторой подобласти, в которой процесс идет достаточно хорошо, тогда основной уровень выбирается либо в центре, либо в случайной точке этой подобласти. Сведения о подобласти можно получить, анализируя изученные ранее подобные процессы, из теоретических соображений или из предыдущего эксперимента. Возможен случай с

несколькими эквивалентными по важности точками, координаты которых различны. Когда отсутствуют дополнительные данные (технологического, экономического характера), выбор произволен. Если эксперимент требует небольшого количества времени и денежных затрат, можно построить планы экспериментов вокруг нескольких точек.

Далее для каждого фактора выбирают два уровня, на которых он будет варьироваться в эксперименте. *Интервалом варьирования факторов* ( $\lambda$ ) называется некоторое число (свое для каждого фактора), прибавление которого к основному уровню дает верхний, а вычитание – нижний уровни фактора. Другими словами, интервал варьирования – это расстояние на координатной оси между основным и верхним (или нижним) уровнем. Для упрощения записи условий эксперимента и обработки экспериментальных данных вводятся кодированные значения уровней факторов: верхний уровень соответствует  $+1$  ( $X^+$ ), нижний:  $-1$  ( $X^-$ ), а основной – нулю.

На выбор интервалов варьирования накладывается ряд условий. Интервал варьирования не может быть меньше той ошибки, с которой экспериментатор фиксирует уровень фактора, иначе верхний и нижний уровни окажутся неразличимыми. С другой стороны, интервал не может быть настолько большим, чтобы верхний или нижний уровни оказались за пределами области определения. При выборе интервалов варьирования учитывают сведения о точности, с которой экспериментатор фиксирует значения факторов, о кривизне поверхности отклика и о диапазоне изменения параметра оптимизации. Обычно эта информация является ориентировочной (в некоторых случаях она может оказаться просто ошибочной), но это единственная разумная основа, на которой можно начинать планировать эксперимент. В ходе эксперимента ее часто приходится корректировать. Точность фиксирования факторов определяется точностью приборов и стабильностью, которую можно поддерживать в ходе опыта. Рассматривают широкий, средний и узкий интервалы варьирования, а также случай, когда трудно принять однозначное решение. Размер интервала варьирования составляет некоторую долю от области определения фактора. Обычно принимают, что если интервал составляет не более 10 % от области определения, его считают узким, не более 30 % – средним и в остальных случаях – широким.

Число опытов в эксперименте определяют по формуле (2.5):

$$N=p^m, \quad (2.5)$$

где  $m$  – число факторов;  $p$  – число уровней факторов.

В технологических исследованиях в основном применяют планы первого и второго порядков. Главное положение планирования эксперимента заключается в том, что его результат представляется в виде полинома, коэффициенты которого и подлежат определению. На первой стадии исследования обычно допускают, что зависимость между факторами определяется полиномом первой степени, и строят план первого порядка. Под *планами первого порядка* понимают такие планы, которые позволяют провести эксперимент для отыскания уравнения регрессии, содержащего только первые



степени факторов и их производные. Для эксперимента с одним фактором результат реализации плана первого порядка выглядит в виде уравнения регрессии следующим образом (2.6):

$$y = A_1 \cdot x_1, \quad (2.6)$$

где  $A_1$  - коэффициент пропорциональности.

Для линейной зависимости уравнение регрессии в общем виде выглядит следующим образом (2.7):

$$y = A_1 \cdot x_1 + A_2 \cdot x_2 + \dots + A_n \cdot x_n, \quad (2.7)$$

в данном случае говорят, что  $y$  линейно зависит от  $x_1, x_2, \dots, x_n$ .

Если существует взаимное влияние на отклик нескольких факторов, то регрессия перестает быть линейной, модель задается полиномом  $k$ -го порядка по факторам и называется *полиномиальной моделью  $k$ -го порядка*.

Зависимость типа (2.8) называется линейным взаимодействием факторов  $x_1$  и  $x_2$ .

$$y = A_{12} \cdot x_1 \cdot x_2. \quad (2.8)$$

Зависимость типа (2.9) называется квадратичной зависимостью, или зависимостью второго порядка.

$$y = A_{11} \cdot x_1^2. \quad (2.9)$$

Нелинейное взаимодействие величин  $x_1$  и  $x_2$  имеет вид (3.10):

$$y = x_1^n \cdot x_2^m. \quad (2.10)$$

Планы второго порядка описываются в общем виде как полный квадратичный полином (2.11):

$$y = A_0 + \sum_{i=1}^n b_i \cdot x_i + \sum_{i=1}^n \sum_{j=1}^n b_{ij} \cdot x_i \cdot x_j + \sum_{j=1}^n b_j \cdot x_j, \quad (2.11)$$

где  $i \neq j$ .

Первый этап планирования эксперимента для получения линейной модели основан на варьировании факторов на двух уровнях. В этом случае, если число факторов известно, можно сразу найти число опытов, необходимое для реализации всех возможных сочетаний уровней факторов (см. формулу (2.5)).

План эксперимента записывается в виде *матрицы плана*, представляющей собой прямоугольную таблицу, строки которой соответствуют опытам, а столбцы – факторам. Каждая строка матрицы плана определяет значения всех факторов в одном опыте эксперимента. Каждому опыту соответствует определенная точка факторного пространства. Опыты, проводящиеся в одной и той же точке факторного пространства, называются *параллельными опытами*.

Матрица планирования для двух факторов на двух уровнях будет иметь следующий вид (таблица 2.1).

Таблица 2.1 - Матрица планирования для двух факторов на двух уровнях

$x_1$	$x_2$	$y$
–	–	$y_1$
+	–	$y_2$
–	+	$y_3$
+	+	$y_4$

*Полный факторный эксперимент* – это эксперимент, в котором реализуются все возможные сочетания уровней факторов. Свойства факторного эксперимента:

- симметричность относительно центра эксперимента – алгебраическая сумма элементов каждого столбца равна нулю;
- условие нормировки – сумма квадратов элементов каждого столбца равна числу опытов;
- ортогональность – сумма почленных произведений любых двух столбцов матрицы равна нулю;
- ротатабельность – точки в матрице планирования подбираются так, что точность предсказания значений параметра оптимизации одинакова на равных расстояниях от центра эксперимента и не зависит от направления.

После проведения экспериментов в соответствии с выбранной матрицей получают значения параметра оптимизации (значения  $y$ ), представляющие собой среднеарифметические величины параллельных опытов. Следующим этапом будет вычисление коэффициентов линейной регрессии (уравнение линейной регрессии в общем виде представлено уравнением (2.7)) – формулы (2.12), (2.13):

$$A_i = \frac{\sum x_{iu} \cdot \bar{y}_u}{N} \quad (2.12)$$

$$A_0 = \frac{\sum \bar{y}_u}{N}, \quad (2.13)$$

где  $x_{iu}$  – значение переменной в соответствующем столбце плана эксперимента;  $\bar{y}_u$  – результат  $u$ -го опыта, среднее арифметическое значение;  $N$  – общее число опытов;  $u$  – номер варианта опыта;  $i$  – номер фактора;  $A_i$  – коэффициенты уравнения регрессии при факторе;  $A_0$  – свободный член уравнения регрессии.

Коэффициенты при независимых переменных указывают на силу влияния факторов. Чем больше численная величина коэффициента, тем большее влияние оказывает фактор. Если коэффициент имеет знак плюс, то с увеличением значения фактора исследуемый параметр ( $y$ ) увеличивается, а если минус, то уменьшается. Величина коэффициента соответствует степени влияния данного фактора на изменение величины исследуемого параметра.

Планируя эксперимент, на первом этапе стремятся получить линейную модель. Однако производство пищевых продуктов очень редко описывается линейной моделью. Один из часто встречающихся видов нелинейности связан с тем, что эффект одного фактора зависит от уровня, на котором находится другой фактор. В этом случае говорят, что имеет место эффект взаимодействия двух факторов. Полный факторный эксперимент позволяет количественно оценивать эффекты взаимодействия. Для этого необходимо, пользуясь правилом перемножения столбцов, получить столбец произведения двух факторов. При вычислении коэффициента, соответствующего эффекту взаимодействия, пользуются формулой (2.12). Для полного факторного эксперимента с двумя факторами матрица планирования с учетом эффекта взаимодействия будет иметь вид (таблица 2.2):

Таблица 2.2 - Матрица планирования для двух факторов с учетом эффекта взаимодействия

$x_1$	$x_2$	$x_1x_2$	$y$
-	-	+	$y_1$
+	-	-	$y_2$
-	+	-	$y_3$
+	+	+	$y_4$

**Пример 2.1.** Дана матрица планирования. Найдите коэффициенты линейной регрессии, запишите уравнение линейной регрессии, если известно, что есть взаимное влияние факторов.

$u$	Натуральные значения факторов, г			Отклик, %
	$X_1$	$X_2$	$X_3$	$\bar{y}_u$
1	12	30	43	50
2	8	22	41	45
3	12	30	41	40
4	12	22	41	70
5	8	22	43	80
6	8	30	43	75
7	12	22	43	64
8	8	30	41	85

Построим матрицу планирования в кодированных значениях:

Кодированные значения факторов							Отклик, %
$X_1$	$X_2$	$X_3$	$X_1X_2$	$X_1X_3$	$X_2X_3$	$X_1X_2X_3$	$\bar{y}_u$
+	+	+	+	+	+	+	50
-	-	-	+	+	+	-	45
+	+	-	+	-	-	-	40
+	-	-	-	-	+	+	70
-	-	+	+	-	-	+	80
-	+	+	-	-	+	-	75
+	-	+	-	+	-	-	64
-	+	-	-	+	-	+	85

По формулам (2.12) и (2.13) рассчитываем коэффициенты регрессии:

$$A_0 = \frac{\sum \bar{y}_u}{N} = \frac{50+45+40+70+80+75+64+85}{8} = 64;$$

$$A_1 = \frac{\sum x_{1u} \cdot \bar{y}_u}{N} = \frac{50-45+40+70-80-75+64-85}{8} = -7,6.$$

Аналогично  $A_2 = -1,1$ ;  $A_3 = 3,6$ .

Коэффициенты регрессии взаимодействий:

$$A_{12} = \frac{\sum x_{iu} \cdot x_{ju} \cdot \bar{y}_u}{N} = \frac{50+45+40-70+80-75-64-85}{8} = -9,9.$$

Аналогично рассчитываем  $A_{13} = -2,6$ ,  $A_{23} = -3,6$ ,  $A_{123} = 7,6$ .

Уравнение регрессии имеет следующий вид:

$$y = 64 - 7,6 x_1 - 1,1 x_2 + 3,6 x_3 - 9,9 x_1 x_2 - 3,6 x_2 x_3 - 2,6 x_1 x_3 + 7,6 x_1 x_2 x_3.$$

**Задание 2.1.** Составить параметрическую схему, указав управляемые, управляющие, возмущающие и наблюдаемые факторы процессов:

1. замеса теста на хлеб;
2. измельчения фарша в куттере для производства вареных колбас;
3. варки овощного бульона.

**Задание 2.2.** Провести процедуру априорного ранжирования следующих экспертных оценок (таблица 2.3):

Таблица 2.3 – Результаты опроса экспертов

Эксперт	Фактор										
	1	2	3	4	5	6	7	8	9	10	11
1	4	1	2	3	5	6	7	8	9	10	11
2	3	1	2	4	5	6	7	8	9	10	11
3	5	2	1	4	3	6	7	8	9	10	11
4	4	1	2	3	5	6	7	8	9	10	11
5	3	1	2	4	5	6	7	8	9	10	11
6	3	2	1	4	9	10	11	5	6	7	8
7	3	1	2	5	4	6	7	8	9	10	11
8	4	2	1	3	5	7	6	8	9	10	11
9	3	2	1	4	6	5	7	8	9	10	11
10	3	1	2	4	5	6	7	8	9	10	11

**Задание 2.3.** Даны две матрицы планирования. Какая из них удовлетворяет всем свойствам факторного эксперимента?

$X_1$	$X_2$
-	-
+	-
+	+
-	+

$X_1$	$X_2$
-	+
+	-
-	+
+	-

**Задание 2.4.** На качество пивного суслу, о котором судят по выходу экстракта, влияет концентрация трех ферментов: амилосубтилина ( $X_1$ ), амилоризина ( $X_2$ ) и цитороземина ( $X_3$ ). Основной фон:  $X_{10} = 0,04$  %;  $X_{20} = 1,0$  %;  $X_{30} = 1,5$  %. Интервалы варьирования для каждого фактора соответственно составляют:  $\lambda_1 = 0,02$  %;  $\lambda_2 = 0,5$  %;  $\lambda_3 = 0,5$  %. Постройте ортогональную матрицу планирования, факторы взаимно не влияют на качество пивного суслу. Найдите коэффициенты уравнения регрессии.

**Задание 2.5.** Постройте план эксперимента зависимости выхода биомассы от концентрации сульфата аммония  $X_1$ , диаммоний фосфата  $X_2$  и биотина  $X_3$  в питательной среде при выращивании сахаромикетов. На основании ранее выполненных исследований известен нулевой уровень

изучаемых факторов:  $X_{10} = 400$  мг/л;  $X_{20} = 30$  мг/л;  $X_{30} = 1$  мкг/л. Интервалы варьирования для каждого соответственно равны:  $\lambda_1 = 200$  мг/л;  $\lambda_2 = 20$  мг/л;  $\lambda_3 = 1$  мкг/л. Факторы взаимно влияют на выход биомассы. Найдите коэффициенты уравнения регрессии.

### Контрольные вопросы

1. Что такое модель?
2. Приведите этапы математического моделирования.
3. Какие бывают виды факторов?
4. Какие требования предъявляют к факторам?
5. Каковы этапы процедуры априорного ранжирования факторов?
6. Каковы свойства матрицы планирования?

### Лабораторная работа № 3

#### Построение матрицы планирования дробного факторного эксперимента

**Цель** – получить навыки выбора наиболее экономного по числу опытов плана эксперимента.

#### Задание:

1. Изучить теоретический материал и примеры построения матриц дробного факторного эксперимента.
2. Решить задачи сокращения числа опытов, обосновывая смешанные оценки влияния факторов.

#### Теоретический материал, методические указания и примеры выполнения заданий

Число опытов в полном факторном эксперименте превышает число коэффициентов линейной модели, причем тем больше, чем больше факторов. Можно сократить число необходимых опытов за счет той информации, которая не очень существенна при построении линейных моделей. При этом нужно стремиться, чтобы матрица планирования не лишилась своих свойств. Минимальное количество опытов для нахождения коэффициентов линейной регрессии рассчитывают по формуле (3.1):

$$N \geq k + 1, \quad (3.1)$$

где  $k$  – количество существенных переменных в уравнении регрессии.

Формальное приравнивание произведения факторов фактору, не входящему в это произведение, носит название *метода дробного факторного эксперимента*. Если используется половина полного факторного эксперимента, то такой план дробного эксперимента называется *полурепликой*. Так, для трех факторов  $x_3$  можно приравнять к  $-x_1x_2$  или  $x_1x_2$ . При реализации обеих полуреплик можно получить отдельные оценки для линейных эффектов и эффектов взаимодействия. Объединение двух полуреплик и есть полный факторный эксперимент  $2^3$ . Матрица из восьми опытов для четырехфакторного

планирования будет полуреplikой от полного факторного эксперимента  $2^4$ , а для пятифакторного планирования – четвертьреplikой от  $2^5$ . В последнем случае два линейных эффекта приравниваются к эффектам взаимодействия. Для обозначения дробных реплик, в которых  $p$  линейных эффектов приравнены к эффектам взаимодействия, удобно пользоваться условным обозначением  $2^{k-p}$ . Так, полуреплика от  $2^3$  запишется в виде  $2^{3-1}$ , а четвертьреплика от  $2^5$  – в виде  $2^{5-2}$ .

Таблица 3.1 - Матрица планирования дробного факторного эксперимента для трех факторов (первая полуреплика)

$x_1$	$x_2$	$x_1x_2 = x_3$	$y$
-	-	+	$y_1$
+	-	-	$y_2$
-	+	-	$y_3$
+	+	+	$y_4$

Таблица 3.2 - Матрица планирования дробного факторного эксперимента для трех факторов (вторая полуреплика)

$x_1$	$x_2$	$-x_1x_2 = x_3$	$y$
-	-	-	$y_1$
+	-	+	$y_2$
-	+	+	$y_3$
+	+	-	$y_4$

Соотношение, показывающее, с каким из эффектов смешан данный эффект, называется *генерирующим соотношением*. Для первой матрицы (таблица 3.1) генерирующим является соотношение:  $x_3 = x_1 \cdot x_2$ , а для второй матрицы (таблица 3.2):  $x_3 = -x_1 \cdot x_2$ .

Для произведения трех столбцов первой матрицы (таблица 3.1) выполняется соотношение:  $1 = x_1 \cdot x_2 \cdot x_3$ , а для второй матрицы (таблица 3.2):  $-1 = x_1 \cdot x_2 \cdot x_3$ . Данное уравнение называют *определяющим контрастом*. Контраст необходим для определения смешанных эффектов. Умножив поочередно определяющий контраст на  $x_1$ ,  $x_2$  и  $x_3$  и принимая во внимание, что  $x_i^2 = 1$ , получим (формулы 3.2-3.4):

$$x_1 = x_1^2 \cdot x_2 \cdot x_3 = x_2 \cdot x_3. \quad (3.2)$$

$$x_2 = x_1 \cdot x_3. \quad (3.3)$$

$$x_3 = x_1 \cdot x_2. \quad (3.4)$$

Полученным соотношениям соответствуют смешанные оценки коэффициентов уравнения регрессии:  $b_1 = \beta_1 + \beta_{23}$ ,  $b_2 = \beta_2 + \beta_{13}$ ,  $b_3 = \beta_3 + \beta_{12}$ . При использовании методики дробного факторного эксперимента коэффициенты регрессии, которые определяются совместно, не могут в явном виде присутствовать в уравнении регрессии, т. к. можно определить лишь их сумму, а не каждый коэффициент по отдельности. Поэтому из уравнения вычеркивают те коэффициенты, которые входят в смешанные оценки. При отсутствии

априорной информации об эффектах взаимодействия необходимо выбрать реплику с наибольшей разрешающей способностью, так как тройные взаимодействия обычно менее важны, чем парные. Если существует информация об эффектах взаимодействия, она должна использоваться при выборе реплики. При построении главных полуреplik в определяющий контраст надо включать наибольшее число факторов.

**Пример 3.1.** Описать определяющий контраст и смешанные оценки коэффициентов уравнения регрессии для плана эксперимента с 4 факторами, если в качестве генерирующих соотношений выбраны:

А)  $x_4 = x_1x_2x_3$ .

Б)  $x_4 = -x_1x_2$ .

Случай А. Определяющий контраст:  $1 = x_4x_1x_2x_3$ .

Смешанные оценки:

$$x_1 = x_2x_3x_4 \quad b_1 = \beta_1 + \beta_{234},$$

$$x_2 = x_1x_3x_4 \quad b_2 = \beta_2 + \beta_{134},$$

$$x_3 = x_1x_2x_4 \quad b_3 = \beta_3 + \beta_{124},$$

$$x_4 = x_1x_2x_3 \quad b_4 = \beta_4 + \beta_{123},$$

$$x_1x_4 = x_2x_3 \quad b_{14} = \beta_{14} + \beta_{23},$$

$$x_1x_2 = x_3x_4 \quad b_{12} = \beta_{12} + \beta_{34},$$

$$x_1x_3 = x_2x_4 \quad b_{13} = \beta_{13} + \beta_{24}.$$

Случай Б. Определяющий контраст:  $1 = -x_4x_1x_2$ .

Смешанные оценки:

$$x_1 = -x_2x_4 \quad b_1 = \beta_1 - \beta_{24},$$

$$x_2 = -x_1x_4 \quad b_2 = \beta_2 - \beta_{14},$$

$$x_3 = -x_1x_2x_3x_4 \quad b_3 = \beta_3 - \beta_{1234},$$

$$x_4 = -x_1x_2 \quad b_4 = \beta_4 - \beta_{12},$$

$$x_2x_3 = -x_1x_3x_4 \quad b_{23} = \beta_{23} - \beta_{134},$$

$$x_1x_2 = -x_4 \quad b_{12} = \beta_{12} - \beta_4,$$

$$x_3x_4 = -x_1x_2x_3 \quad b_{34} = \beta_{34} - \beta_{123},$$

$$x_1x_4 = -x_2 \quad b_{14} = \beta_{14} - \beta_2.$$

**Задание 3.1.** Составить план наиболее экономного дробного факторного эксперимента при данном списке существенных переменных, записать матрицу плана:  $x_1, x_2, x_3, x_4, x_1x_2, x_2x_3, x_2x_4$ .

**Задание 3.2.** Исследуется скорость замерзания слоя соленой рыбы. В качестве управляющих переменных выбраны:  $x_1$  – насыпная плотность рыбы, г/см<sup>3</sup>;  $x_2$  – содержание соли, %;  $x_3$  – концентрация хладагента, %;  $x_4$  – температура, °С. Факторы изменяются в следующей области:  $0,72 \leq x_1 \leq 1,02$ ;  $5 \leq x_2 \leq 15$ ;  $0,75 \leq x_3 \leq 1,25$ ;  $-30 \leq x_4 \leq -20$ .

Составить план наиболее экономного по количеству опытов дробного факторного эксперимента, записать матрицу плана в кодированных и натуральных значениях переменных, уравнение регрессии в общем виде.

### **Контрольные вопросы**

1. Как выбрать минимальное количество опытов для проведения экспериментов, используя математическую модель в виде уравнения регрессии?
2. Что такое определяющий контраст?
3. Что такое генерирующее соотношение?
4. Как определить смешанные оценки коэффициентов уравнения регрессии?

### **Лабораторная работа № 4**

#### **Моделирование рецептов методом линейного программирования**

**Цель** - ознакомиться с симплекс-методом решения задач линейного программирования на примере моделирования рецептуры многокомпонентного продукта по критерию минимальной себестоимости.

#### **Задание:**

1. Изучить теоретический материал и примеры построения симплекс-таблиц.
2. Решить задачу моделирования рецептов методом линейного программирования, используя MS Excel.

#### **Теоретический материал, методические указания и примеры выполнения заданий**

Симплекс-метод – алгоритм решения оптимизационной задачи линейного программирования путём перебора вершин выпуклого многогранника в многомерном пространстве.

В технологии продуктов питания наиболее часто при помощи симплекс-метода решают два типа задач: о диете и технологическую.

Одним из основных направлений в области здорового питания является производство продуктов питания, имеющих поликомпонентный состав и включающих как основные нутриенты, так и микронутриенты, к которым относятся витамины, минеральные вещества. В организм человека пищевые нутриенты должны поступать в определенном количестве и в определенном соотношении, что чётко учитывается при моделировании пищевых продуктов сложного сырьевого состава.

*Задача о диете – задача минимума.* Пусть в диете присутствуют продукты  $F_1, F_2 \dots F_n$ . Для того, чтобы диета была сбалансированной, в ней должны присутствовать питательные вещества  $N_1, N_2, N_3$ . В каждом продукте  $F_i$  содержится  $a_{ij}$  количество питательного вещества  $N_j$ . Составим матрицу питательности (таблица 4.1).



Таблица 4.1 – Матрица питательности

Питательное вещество	Продукт		
	F <sub>1</sub>	F <sub>2</sub>	...F <sub>n</sub>
N <sub>1</sub>	a <sub>11</sub>	a <sub>12</sub>	...a <sub>1n</sub>
N <sub>2</sub>	a <sub>21</sub>	a <sub>22</sub>	...a <sub>2n</sub>
...N <sub>m</sub>	a <sub>31</sub>	a <sub>32</sub>	...a <sub>3n</sub>

Пусть  $b_1, b_2 \dots b_m$  – минимальное количество питательного вещества  $N_1, N_2 \dots N_m$  соответственно, необходимое для удовлетворения потребности человека, а  $x_1, x_2 \dots x_n$  – количество продукта  $F_1, F_2 \dots F_n$  соответственно, которое необходимо съесть, чтобы удовлетворить потребность в питательных веществах.

Тогда рацион должен удовлетворять следующим ограничениям (4.1):

$$\begin{cases} a_{11} \cdot x_1 + a_{12} \cdot x_2 + a_{1n} \cdot x_n \geq b_1. \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + a_{2n} \cdot x_n \geq b_2. \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + a_{mn} \cdot x_n \geq b_m. \\ x_1 \geq 0, x_2 \geq 0, x_n \geq 0. \end{cases} \quad (4.1)$$

Задача, в которой фигурируют ограничения в форме неравенств, называется *основной задачей линейного программирования*.

Пусть  $C_1, C_2 \dots C_n$  – цена продуктов  $F_1, F_2 \dots F_n$  соответственно. Необходимо, чтобы на поддержание диеты человек тратил как можно меньше денег. Тогда целевая функция будет функцией минимума и примет вид (4.2):

$$C_1 \times x_1 + C_2 \times x_2 + \dots C_n \times x_n \rightarrow \min. \quad (4.2)$$

Задача нахождения минимума линейной целевой функции (линейной формы) называется *общей (стандартной) задачей* линейного программирования.

*Технологическая задача – задача максимума.* Пусть на производстве есть ассортимент продукции  $G_1, G_2 \dots G_n$  и запас сырья  $R_1, R_2 \dots R_m$ . Для производства каждой ассортиментной единицы  $G_i$  необходимо  $a_{ij}$  количество сырья  $R_j$ . Составим технологическую матрицу (таблица 4.2).

Пусть  $b_1, b_2 \dots b_m$  – количество сырья  $R_1, R_2, R_m$  соответственно, имеющееся на складе и которое может быть использовано для производства продукции, а  $x_1, x_2 \dots x_n$  – количество продукта  $G_1, G_2 \dots G_n$  соответственно, которое необходимо произвести.

Таблица 4.2 – Технологическая матрица

Вид сырья	Ассортимент		
	G <sub>1</sub>	G <sub>2</sub>	...G <sub>n</sub>
R <sub>1</sub>	a <sub>11</sub>	a <sub>12</sub>	...a <sub>1n</sub>
R <sub>2</sub>	a <sub>21</sub>	a <sub>22</sub>	...a <sub>2n</sub>
...R <sub>m</sub>	a <sub>31</sub>	a <sub>32</sub>	...a <sub>3n</sub>

Тогда бесперебойность производства должна удовлетворять следующим ограничениям (4.3):

$$\begin{cases} a_{11} \cdot x_1 + a_{12} \cdot x_2 + a_{1n} \cdot x_n \leq b_1 \\ a_{21} \cdot x_1 + a_{22} \cdot x_2 + a_{2n} \cdot x_n \leq b_2 \\ a_{m1} \cdot x_1 + a_{m2} \cdot x_2 + a_{mn} \cdot x_n \leq b_m \\ x_1 \geq 0, x_2 \geq 0, x_n \geq 0 \end{cases} \quad (4.3)$$

Пусть  $C_1, C_2 \dots C_n$  – количество продуктов  $G_1, G_2 \dots G_n$  соответственно. Необходимо, чтобы предприятие выпускало как можно больше продукции. Тогда целевая функция будет функцией максимума и примет вид (4.4):

$$C_1 \cdot x_1 + C_2 \cdot x_2 + \dots C_n \cdot x_n \rightarrow \max. \quad (4.4)$$

Задачу линейного программирования можно решить при помощи надстройки «Поиск решений» в MS Excel.

**Пример 4.1.** Рассмотрим методику проектирования рецептурной смеси на примере составления рецептуры рыба-растительного мусса. Надстройка «Поиск решений» по умолчанию отключена. Чтобы включить ее, необходимо перейти по вкладке Файл в группу Параметры. В появившемся диалоговом окне Параметры, выберите Надстройки, далее Управление: Надстройки Excel, далее Перейти. В окне Надстройки устанавливаем галочку напротив поля «Поиск решения», нажимаем ОК (рисунок 4.1).

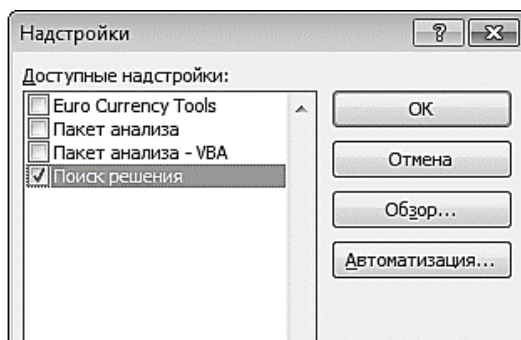


Рисунок 4.1 – Окно установки пакета «Поиск решений»

Теперь во вкладке «Данные» появилась новая группа «Анализ» с кнопкой «Поиск решения» (рисунок 4.2).

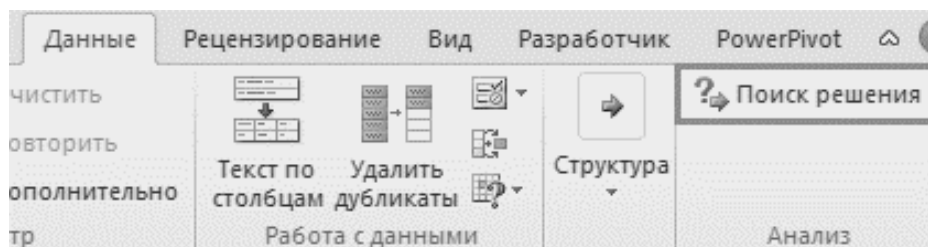


Рисунок 4.2 – Панель «Данные»

Прежде всего, необходимо сформировать информационную таблицу данных для расчета рецептуры, в которую должна входить информация по содержанию жира, белка, углеводов и воды в каждом компоненте рецептуры, а

также средняя оптовая цена за 1 кг продукта. В таблице 4.3 представлена информационная матрица данных для расчета рецептур рыборастворительного мусса с добавлением цветной капусты. Информация из данной таблицы вносится в Microsoft Excel.

Таблица 4.3 – Информационная матрица данных для расчета рецептур рыборастворительного мусса с добавлением цветной капусты

Ингредиенты	X <sub>i</sub>	Массовая доля, %				Цена, руб./кг
		жира	белка	углеводов	воды	
Судак	X <sub>1</sub>	1,1	18,4	0	79,2	250
Рис	X <sub>2</sub>	1,0	7,0	74,0	14,0	90
Цветная капуста	X <sub>3</sub>	0,3	2,5	4,2	90,0	110
Соль поваренная пищевая	X <sub>4</sub>	0	0	0	0,2	9
Яичный порошок	X <sub>5</sub>	7,3	46,0	4,5	7,3	67
Желатин	X <sub>6</sub>	0,4	87,2	0,7	10,0	85
Молоко	X <sub>7</sub>	0,05	3,0	4,9	91,2	25
Перец душистый молотый	X <sub>8</sub>	3,26	10,95	38,31	12,0	140
Копченая паприка	X <sub>9</sub>	17,0	15,0	13,0	11,0	220

На основании информационной матрицы данных (таблица 4.3) формируем систему линейных алгебраических балансовых уравнений по белку, жирам, углеводам, воде и массе рыборастворительного мусса (таблица 4.4). В качестве образца составим линейное алгебраическое неравенство для показателя жира. Каждый из девяти компонентов рецептуры, обозначенный X с соответствующим порядковому номеру индексом, умножается на соответствующее значение массовой доли жира, содержащейся в данном компоненте. Сумма множителей согласно требованиям должна быть больше или равна 24 г, что ограничивает пределы варьирования компонентов с целью получения оптимального варианта рецептуры при заданных условиях. Такие компоненты, как молоко, яичный порошок, желатин, поваренная соль пищевая, перец душистый молотый и копченая паприка, имеют фиксированную массу, предварительно рассчитанную по рецептуре, варьированию подлежат только основные компоненты, такие как судак и цветная капуста.

Таблица 4.4 – Система балансовых линейных алгебраических уравнений и неравенств в рецептуре рыборастворительного мусса

Показатели	Уравнения и ограничения
1	2
Баланс:	
Жир	$0,011 \cdot X_1 + 0,01 \cdot X_2 + 0,003 \cdot X_3 + 0,073 \cdot X_5 + 0,004 \cdot X_6 + 0,0005 \cdot X_7 + 0,0326 \cdot X_8 + 0,17 \cdot X_9 \geq 24$

1	2
Белок	$0,184 \cdot X_1 + 0,07 \cdot X_2 + 0,025 \cdot X_3 + 0,46 \cdot X_5 + 0,872 \cdot X_6 + 0,03 \cdot X_7 + 0,1095 \cdot X_8 + 0,15 X_9 \geq 11$
Углеводы	$0,74 \cdot X_2 + 0,042 \cdot X_3 + 0,045 \cdot X_5 + 0,007 \cdot X_6 + 0,049 \cdot X_7 + 0,3831 \cdot X_8 + 0,13 \cdot X_9 \geq 3$
Вода	$0,792 \cdot X_1 + 0,14 \cdot X_2 + 0,9 \cdot X_3 + 0,002 X_4 + 0,073 X_5 + 0,1 X_6 + 0,91 X_7 + 0,12 X_8 + 0,11 X_9$
Молоко	$X_4 = 0,3$
Яичный порошок	$X_5 = 0,3$
Желатин	$X_6 = 0,075$
Соль поваренная пищевая	$X_7 = 0,4$
Перец душистый молотый	$X_8 = 0,006$
Копченая паприка	$X_9 = 0,025$
Масса	$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7 + X_8 + X_9 = 100$

В качестве критерия оптимизации (функционала) выбрана себестоимость рыборастворительного мусса и энергетическая ценность. Выражение функции цели - минимальная себестоимость рыборастворительного мусса запишется в виде (4.5):

$$250X_1 + 90X_2 + 110X_3 + 9X_4 + 67X_5 + 85X_6 + 25X_7 + 140X_8 + 220X_9 \rightarrow \min. \quad (4.5)$$

Максимизация энергетической ценности рыборастворительного мусса имеет вид (4.6):

$$\sum_{i=1}^n E_i * x_i \rightarrow \max, \quad (4.6)$$

где  $x_i$  – вектор неизвестных искомым,  $E_i$  – энергетическая ценность  $i$ -го ингредиента, ккал.

	A	B	C
1			
2		X	Масса, кг
3	Ингредиенты		
4	Судак	X1	31,00
5	Рис	X2	35,00
6	Цветная капуста	X3	22,94
7	Соль	X4	3,00
8	Яичный порошок	X5	3,00
9	Желатин	X6	0,75
10	Молоко	X7	4,00
11	Перец молотый	X8	0,06
12	Копченая паприка	X9	0,25
13	масса, кг		100,00
14	стандарт молока		

Рисунок 4.3 - Сумма всех компонентов мусса

Систему линейных балансовых уравнений решаем в компьютерной системе Excel. Методология проведения оптимизации заключается в выполнении следующих последовательных операций. В ячейке C13 вычисляем сумму всех компонентов мусса по формуле =СУММ (C4:C12), как указано на рисунке 4.3.



продукции A7:A12) и установить её равной минимальному значению. Пустое диалоговое окно выглядит следующим образом (рисунок 4.6):

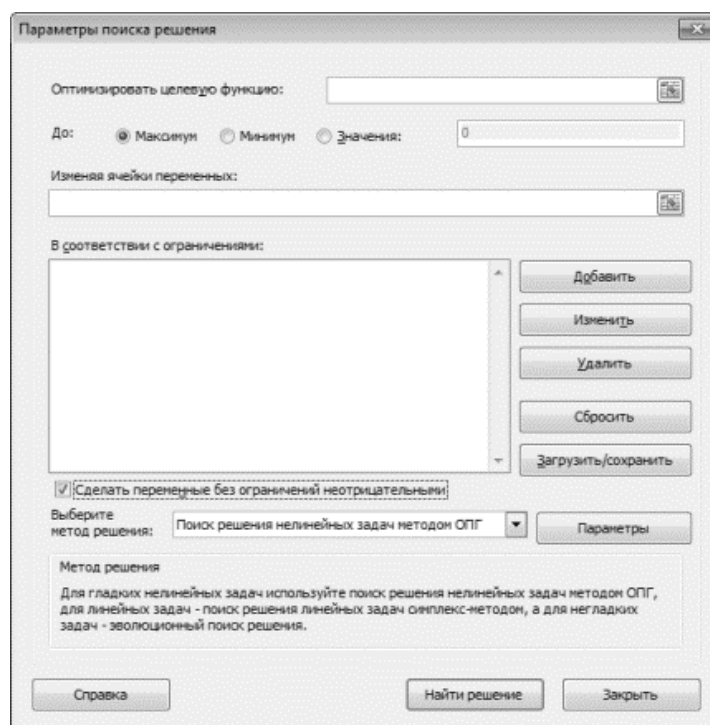


Рисунок 4.6 - Пустое диалоговое окно поиска решения

Далее выбираем изменяющиеся ячейки - это ячейки, содержащие массы отдельных ингредиентов, т. е. в данном случае C4:C6 (рисунок 4.7).

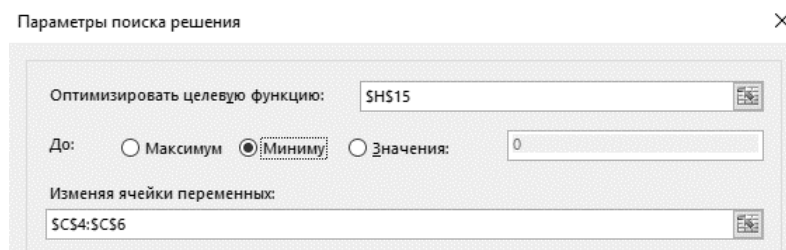


Рисунок 4.7 – Заполнение окна поиска решения

Затем добавляем ограничение: все ячейки, содержащие массу отдельных видов ингредиентов, которые должны быть больше или равны нулю (если таковые имеются), массовые доли жира, белка, углеводов, воды в 100 кг готового продукта, должны быть равны соответствующим ячейкам. Ограничения – это правила, которые лимитируют возможные решения проблемы. После ввода всех параметров нажимаем кнопку «выполнить» в окне поиска решений (рисунок 4.8).

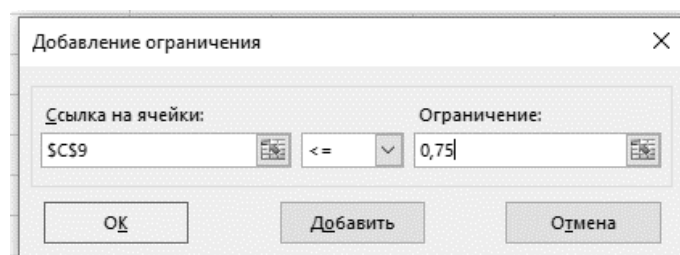


Рисунок 4.8 – Добавление ограничений

На рисунке 4.9 приведена рецептура с минимальной себестоимостью. Искомое значение появится в ячейке, куда был направлен результат оптимизации целевой функции в верхней строчке параметров поиска решения, в данном случае в ячейке Н15.

	A	B	C	D	E	F	G	H
1								
2		X	Масса, кг		Массовая доля, %			Оптовая цена, руб./кг
3	Ингредиенты			жира	белка	углеводов	воды	
4	Судак	X1	31,00	1,1	18,4	0	79,2	250
5	Рис	X2	35,00	1	7	74	14	90
6	Цветная капуста	X3	22,94	0,3	2,5	4,2	90	110
7	Соль	X4	3,00	0	0	0	0,2	9
8	Яичный порошок	X5	3,00	7,3	46	4,5	7,3	67
9	Желатин	X6	0,75	0,4	87,2	0,7	10	85
10	Молоко	X7	4	0,05	3	4,9	91,2	25
11	Перец молотый	X8	0,06	3,26	10,95	38,31	12	140
12	Копченая паприка	X9	0,25	17	15	13	11	220
13	масса, кг		100,00					
14	стандарт молока			1,028276	10,92557	27,25522	54,0807	
15	функция цели, руб							13878,55
16	ввод балансовых уравнений			1,028276	10,92557	27,25522	54,0807	

Рисунок 4.9 – Фрагмент оптимизации рецептуры по максимальной энергетической ценности в рыборастворительном муссе в Excel

Из расчёта, проведенного симплекс-методом, следует, что минимальная себестоимость продукции составляет 13878,55 руб. за 88,94 кг основного сырья и 14854, 37 руб. за 100 кг готового продукта с учетом дополнительных компонентов при зафиксированных выше пропорциях компонентов, обуславливающих требуемые энергетическую ценность и содержание минеральных веществ.

**Задание 4.1.** Решить задачу линейного программирования при помощи надстройки «Поиск решений» в MS Excel. Для изготовления хлеба используются два вида муки: ржаная и пшеничная. Для изготовления одного вида хлеба используется 0,2 т ржаной и 0,1 т пшеничной муки. Для изготовления второго вида хлеба используется 0,15 т ржаной и 0,2 т пшеничной муки. Доход пекарни от производства одного хлеба по первой рецептуре составляет 12 руб., от производства одного хлеба по второй рецептуре – 15 руб. Определить, сколько хлеба каждого вида должна изготовить пекарня, чтобы обеспечить наибольшую рентабельность их производства, если в распоряжении имеется 60 т ржаной и 40 т пшеничной муки.

### **Контрольные вопросы**

1. В каких ситуациях применяют матрицу питательности?
2. Когда применяют технологическую матрицу?
3. Как выглядит основная задача линейного программирования?

### **Лабораторная работа № 5**

#### **Обработка результатов прямых и косвенных измерений, проверка однородности независимых выборок и воспроизводимости опытов в MS Excel**

**Цель** – получить навыки статистической обработки результатов эксперимента и проверки данных опытов на однородность и независимость.

**Задание:**

1. Изучить теоретический материал и примеры обработки измерений.
2. Решить задачи на определение ошибок опыта и проверку гипотез, используя MS Excel.

#### **Теоретический материал, методические указания и примеры выполнения заданий**

Все измерения делят на прямые, косвенные и совместные измерения.

*Прямыми* называются измерения, в результате которых значение исследуемой величины непосредственно показывается измерительным прибором, например, измерение массы с помощью весов.

В *косвенных* измерениях искомая величина получается из расчета по формуле, в которую входят величины, получаемые при прямых измерениях.

*Совместными* измерениями называются такие измерения, при которых одновременно измеряют две или более неоднородных величины для нахождения зависимости между ними или для определения параметров этой зависимости.

Любые измерения всегда производятся с погрешностями, связанными с ограниченной точностью измерительных приборов, метода измерений, квалификацией экспериментатора, изменением условий измерения, поэтому в задачу измерения входит нахождение не только самой величины, но и погрешности (неопределенности). В рамках ISO 9000 было разработано «Руководство по вычислению неопределенности в измерении», в котором описаны понятие неопределенности измерений и способы ее вычисления. В рекомендациях по межгосударственной стандартизации дано разъяснение терминов «погрешность» и «неопределенность».

*Погрешность измерения* – это отклонение измеренного значения величины от ее «истинного» значения. По своей природе или характеру проявления погрешность может быть «случайной» и «систематической». Метод выражения погрешности измерений:



$$a \pm \Delta a,$$

где  $a$  – измеренная величина,  $\Delta a$  – суммарная абсолютная погрешность, определяемая методикой выполнения измерений.

*Неопределенность измерения* – это «сомнения в истинности полученного результата», т.е. в основе концепции неопределенности лежит неполное знание значения измеряемой величины, которое представлено оператору в виде ряда величин, полученных в результате измерительного эксперимента и каким-то образом характеризующих измеряемую величину. Метод выражения неопределенности:

$$a \pm Ua,$$

где  $a$  – измеренная величина;  $Ua$  – расширенная неопределенность, определяемая измерителем.

*Расширенная неопределенность (доверительные границы) результата измерения* – интервал вокруг результата измерения, в который, как ожидается, попадает большая часть значений, приписанных к измеряемой величине.

*Вероятность охвата, вероятность покрытия (доверительная вероятность)* – вероятность, с которой большая часть значений, приписанных к измеряемой величине, попадает в заданную область расширенной неопределенности результата измерений. При выборе вероятности охвата необходимо иметь представление о виде закона распределения неопределенности.

Основное различие в концепциях неопределенности и погрешности состоит в том, к какой величине относят дисперсию: к действительному значению измеряемой величины или к результату измерения.

Расширенную неопределенность (доверительные границы) рассчитывают, зная значения неопределенностей типов А и Б (случайной и систематической погрешностей).

*Неопределенность типа А (случайная погрешность)* – составляющая неопределенности (погрешности) измерения, изменяющаяся случайным образом в серии измерений и не носящая систематического характера. Случайная погрешность многократных измерений определяется по формуле (5.1):

$$\Delta X = t_{\alpha} S_{\langle X \rangle}, \quad (5.1)$$

где  $t_{\alpha}$  – коэффициент Стьюдента для выбранного уровня надежности  $\alpha$  (Приложение Б);  $S_{\langle X \rangle}$  – выборочное среднее квадратичное отклонение среднего значения, вычисляемое по формуле (5.2):

$$S_{\langle X \rangle} = \frac{S_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n(n-1)}}, \quad (5.2)$$

где  $S_x$  – стандартная неопределенность (стандартное отклонение погрешности) результата измерения – неопределенность, представленная в виде стандартного отклонения для выборки.

В MsExcel можно автоматически вычислить среднее значение величины, набрав в ячейке =СРЗНАЧ и выделив те ячейки, значения которых нужно посчитать (рисунок 5.1).

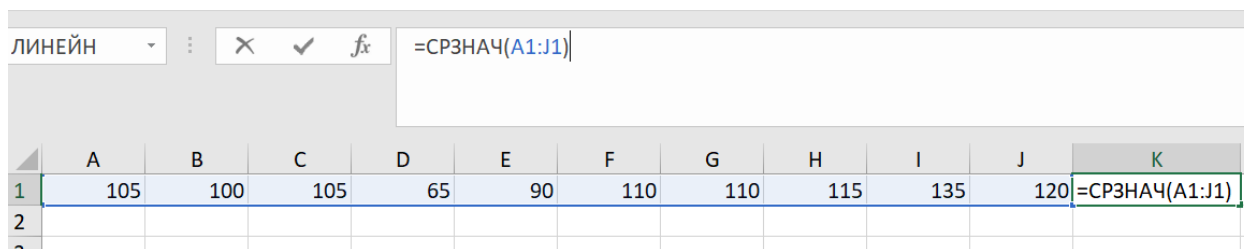


Рисунок 5.1 – Вычисление среднего значения в Ms Excel

*Неопределенность типа Б (систематическая погрешность)* – составляющая неопределенности измерений, постоянная или изменяющаяся по известному закону (например, погрешность прибора, калибровки, методики измерения).

Для прямых измерений систематической будет погрешность прибора измерения. Величину систематической погрешности считают равной половине цены деления шкалы измерительного прибора, либо она записана в паспорте прибора.

На неопределенность типа Б часто влияет округление чисел. Существуют следующие правила округления:

- все вычисления окончательного результата следует производить с числом значащих цифр, превышающих на единицу число значащих цифр, полученных при измерениях;
- если первая из отбрасываемых цифр больше 5, то последняя из сохраняемых цифр усиливается, т. е. увеличивается на единицу;
- если первая из отбрасываемых цифр меньше 5, то последняя из сохраняемых цифр не усиливается, т. е. остается без изменения;
- если первая слева из отбрасываемых цифр равна 5 и за ней следуют цифры, отличные от нуля, то последняя из сохраняемых цифр усиливается, т. е. увеличивается на единицу;
- если первая слева из отбрасываемых цифр равна 5 и за ней не следуют отличные от нуля цифры, то последняя оставшаяся цифра усиливается, если она нечетная, и остается без изменения, если она четная, – это банковское округление «к ближайшему четному».

Если необходимо объединить результаты нескольких серий прямых измерений одной и той же величины ( $n_1 \neq n_2$ ), то вместо среднего арифметического определяют взвешенное среднее  $\langle X \rangle$  по формуле (5.3):

$$\langle X \rangle = \frac{n_1 \cdot \bar{X}_1 + n_2 \cdot \bar{X}_2}{n_1 + n_2}. \quad (5.3)$$

*Полная абсолютная погрешность* прямого измерения  $\Delta X$  равна квадратному корню из суммы квадратов случайной и систематической погрешностей и определяется формулой (5.4):

$$\Delta X = \sqrt{\Delta X_{ct}^2 + \Delta X_{np}^2}. \quad (5.4)$$

Значение полной абсолютной погрешности косвенно измеренной величины определяется формулами (5.5-5.6):

$$\Delta x = \sqrt{(\Delta_x)^2 + (\Delta_y)^2 + (\Delta_z)^2 + \dots + (\Delta_t)^2}, \quad (5.5)$$

где  $\Delta_x = \left(\frac{\partial f}{\partial x}\right)\Delta x$ ;  $\Delta_y = \left(\frac{\partial f}{\partial y}\right)\Delta y$ ;  $\Delta_z = \left(\frac{\partial f}{\partial z}\right)\Delta z$ ; ...;  $\Delta_t = \left(\frac{\partial f}{\partial t}\right)\Delta t$ . (5.6)

Относительная погрешность как прямых, так и косвенных измерений определяется по формуле (5.7):

$$\delta x = \frac{\Delta x}{\bar{x}} \cdot 100\%. \quad (5.7)$$

Результат обработки измерений записывается в форме:  $x = (\bar{X} \pm \Delta X)$  ед. изм.,  $\delta x$ , %.

*Промахи, грубые ошибки* - это измерения, имеющие отклонение от среднего значения, существенно превосходящее ожидаемую погрешность при заданных условиях измерения. Промахи должны быть устранены путем контрольных измерений. Для проверки выбросов (грубых ошибок) используются различные критерии, один из них - *критерий Шовене*, который определяется по формуле (5.8):

$$Z = \frac{|X_k - \bar{X}|}{S_x}, \quad (5.8)$$

где  $X_k$  - аномальный результат.

По таблице критерия Шовене (Приложение В), зная Z, находят M. Если  $M \geq n$ , то результат является промахом и отбрасывается.

**Пример 5.1.** Пусть при измерениях проведено 10 отсчетов массы продукта. Результаты измерения представлены в таблице 5.1. Определить грубые промахи (выбросы) и абсолютную погрешность (расширенную неопределенность) измерений.

Таблица 5.1 - Результаты измерения массы T, г

Вариант	1	2	3	4	5	6	7	8	9	10
1	105	100	105	65	90	110	110	115	135	120

Зададим доверительную вероятность  $\alpha = 0,95$ .

Вычислим среднее значение массы (формула (5.9)):

$$\bar{T} = \frac{\sum_{i=1}^n T_i}{n}; \quad (5.9)$$

$$\bar{T} = \frac{105 + 100 + 105 + 65 + 90 + 110 + 110 + 115 + 135 + 120}{10} = 105,5 \text{ г.}$$

Округляем полученное число до целого значения:  $\bar{T} = 106 \text{ г.}$

Вычисляем среднее квадратичное отклонение отдельных отсчетов по формуле (5.10).

$$S_T = \sqrt{\frac{\sum_{i=1}^n (T_i - \bar{T})^2}{n-1}} ; \quad (5.10)$$

$$S_T = \sqrt{\frac{(-1)^2 + (-6)^2 + (-1)^2 + (-41)^2 + (-16)^2 + 4^2 + 9^2 + 29^2 + 14^2 + 4^2}{9}} = 18,62.$$

Округляем полученное число до целого согласно точности измерительного прибора:  $S_T = 19$  г.

Проверяем отсчеты на наличие промахов. Наибольшее отклонение от среднего значения имеет отсчет под номером 4:  $T_4 = 65$  г. Вычисляем нормированное отклонение  $T_4$  от среднего значения:

$$Z = \frac{|T_4 - \bar{T}|}{S_T} = \frac{|65 - 106|}{19} = 2,16.$$

По Прилож. В определяем количество отсчетов (опытов), при котором рассмотренное значение нельзя считать промахом:  $M = 17$ . В рассматриваемом опыте 10 отсчетов, поэтому значение  $T = 65$  г является промахом, и его нужно исключить из обрабатываемого ряда.

Теперь получаем новый набор измерений (таблица 5.2) и проводим их обработку.

Таблица 5.2 - Результаты измерения массы после удаления первого промаха

$n$ п/п	1	2	3	4	5	6	7	8	9
$T$ , г	105	100	105	90	110	110	115	135	120

Теперь  $n = 9$ . Вычисляем новое среднее значение:

$$\bar{T} = \frac{105 + 100 + 105 + 90 + 110 + 110 + 115 + 135 + 120}{9} = 110,2.$$

Вычисляем среднее квадратичное отклонение от среднего значения:

$$S_T = \sqrt{\frac{(-5)^2 + (-10)^2 + (-5)^2 + (-20)^2 + 0^2 + 0^2 + (-25)^2 + (-10)^2 + 5^2}{8}} = 12,72.$$

Округляем это значение до целого числа и получаем  $\Delta S_T = 13$  г.

Проверим, не является ли промахом результат  $T_8 = 135$  г. Для этого найдем значение  $Z$ :

$$Z = \frac{|T_8 - \bar{T}|}{S_T} = \frac{|110 - 135|}{13} = 1,9.$$

По Прилож. В значению  $z = 1,9$  соответствует значение  $M = 9$ . Это значит, что рассмотренное значение измерения массы также является промахом, и его нужно отбросить. Теперь таблица состоит из результатов восьми измерений (таблица 5.3).

Таблица 5.3 - Результаты измерения массы после удаления второго промаха

$n$ п/п	1	2	3	4	5	6	7	8
$T$ , г	105	100	105	90	110	110	115	120

Вычисляем новое среднее значение массы:

$$\bar{T} = \frac{105+100+105+90+110+115+120+110}{8} = 106,9\text{г}.$$

Округляем эту величину до целого значения, получаем  $\bar{T} = 107\text{г}$ . Наибольшее отклонение от этого среднего значения имеет результат измерения  $T_4 = 90\text{г}$ . Проверим, является ли этот результат измерения промахом.

Вычисляем среднее квадратичное отклонение от среднего значения:

$$S_T = \sqrt{\frac{(-2)^2 + (-7)^2 + (-2)^2 + 17^2 + 3^2 + 3^2 + 2^2 + 13^2}{7}} = 8,6\text{г} \approx 9\text{г}.$$

Вычислим z:

$$Z = \frac{|T_4 - \bar{T}|}{S_T} = \frac{|90 - 107|}{9} = 1,9.$$

По Приложению В  $z = 1,9$  соответствует значение  $M = 9$ , а в рассматриваемом случае измерений всего восемь. Это означает, что данный результат измерения также является промахом, и его надо отбросить. Теперь таблица состоит из семи результатов измерений (таблица 5.4).

Таблица 5.4 - Результаты измерения массы после удаления третьего промаха

n п/п	1	2	3	4	5	6	7
T, г	105	100	105	110	110	115	120

Тогда:

$$\bar{T} = \frac{105+100+105+110+110+115+120}{7} = 109,3\text{г} \approx 109\text{г}.$$

$$S_T = \sqrt{\frac{(-4)^2 + (-9)^2 + (-4)^2 + 1^2 + 6^2 + 11^2}{6}} = 6,7\text{г} \approx 7\text{г}.$$

Проверим, не является ли промахом значение  $T_7 = 120\text{г}$ .

$$Z = \frac{|120 - 109|}{7} = \frac{11}{7} = 1,6.$$

Полученному значению  $z = 1,6$  соответствует значение  $M = 5$ , а в данном случае измерений 7, поэтому данное значение нельзя считать промахом. Очевидно, что остальные значения проверять не имеет смысла, так как им будет соответствовать еще меньшее значение  $M$ , а значит, они не будут промахами.

Надежности  $\alpha = 0,95$  и количеству измерений  $n = 7$  соответствует коэффициент Стьюдента  $t_\alpha = 2,4$  (Прилож. Б). Тогда находим  $\Delta T_{\text{ст}} = \frac{6,7}{\sqrt{7}} \cdot 2,4 \approx 6\text{г}$ .

Результат измерения массы имеет вид:  $T = (109 \pm 6)\text{г}; \alpha = 0,95$ .

В ходе исследования выдвигается *научная гипотеза* – утверждение, которое можно подтвердить или опровергнуть на основании результатов исследования. Ей соответствует *статистическая гипотеза* – представление

научной гипотезы в форме, приемлемой для проверки методами статистического анализа данных.

$H_0$  – *нулевая гипотеза*: предположение о том, что различия между сравниваемыми выборками отсутствуют. Её математический смысл состоит в том, что  $\bar{X}_1 - \bar{X}_2 \rightarrow 0$ . В действительности различия могут отклоняться от 0, но быть недостоверными или недоказанными.  $H_1(\bar{H})$  – *альтернативная гипотеза* (противостоящая нулевой гипотезе) – различия между выборками есть и они достоверны.

В ходе проверки гипотез могут возникнуть два типа ошибок:

– *Ошибка I рода* – вероятность отклонения нулевой гипотезы, когда она на самом деле верна (ее также называют  $\alpha$ -риском, ложноотрицательным результатом, риском производителя).

– *Ошибка II рода* – вероятность остаться в рамках нулевой гипотезы, когда на самом деле она неверна (ее также называют  $\beta$ -риском, ложноположительным результатом, риском потребителя).

*Критерий для проверки гипотезы* – это решающее правило (метод), отвергающее или принимающее нулевую гипотезу на основе выборочных наблюдений.

*Односторонний критерий* – статистический критерий, учитывающий априорные знания о направлении (увеличении либо уменьшении) значения исследуемого параметра.

*Двусторонний критерий* – условие проверки гипотез, определяемое до сбора данных. Двусторонний тест не предполагает, что заранее известно направление смещения значения анализируемого параметра.

Установлено пять стадий проверки гипотез:

– Определение нулевой ( $H_0$ ) и альтернативной гипотезы ( $H_1$ ) при исследовании. Определение уровня значимости  $\alpha$  (максимально допустимой исследователем вероятности ошибочного отклонения нулевой гипотезы). Обычно уровень значимости принимают равным 0,05 или 0,01.

– Отбор необходимых данных из выборки.

– Вычисление значения статистики критерия, отвечающей  $H_0$ .

– Вычисление критической области, проверка статистики критерия на предмет попадания в критическую область.

– Интерпретация результатов.

*t-критерий Стьюдента* для независимых выборок применяется для сравнения средних значений двух независимых между собой выборок (формула (5.11)). При условии, что сравниваемые значения не составляют пару коррелирующих значений, распределение признаков в каждой выборке соответствует нормальному распределению, дисперсии признака в выборках примерно равны, т. е. дисперсии не признаны однородными (это проверяется с помощью критерия F Фишера):

$$t_{\alpha;f} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad (5.11)$$

где  $\bar{x}_1$  – среднее арифметическое первой выборки;  $\bar{x}_2$  – среднее арифметическое второй выборки;  $S_1$  – стандартное отклонение первой выборки;  $S_2$  – стандартное отклонение второй выборки;  $n_1$  – объем первой выборки;  $n_2$  – объем второй выборки;  $f$  – число степеней свободы – параметр статистического распределения, равный числу независимых связей оцениваемой статистической выборки,  $f = n_1 + n_2 - 2$ .

Если после проверки гипотез на однородность по критерию F Фишера дисперсии признаны однородными, *t-критерий Стьюдента* вычисляется как:

$$t_{2\alpha;f} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{S_0^2 \times (n_1 + n_2)}{n_1 \times n_2}}}, \quad (5.12)$$

где  $S_0^2 = \frac{(n_1 - 1) \times S_1^2 + (n_2 - 1) \times S_2^2}{n_1 + n_2 - 2}$  – общая выборочная дисперсия;  $f = n_1 + n_2 - 2$ .

Для проверки гипотезы отличия исследуемого образца от теоретического (эталонного значения) применяют формулу (5.13):

$$t_{2\alpha;f} = \frac{|\bar{x}_1 - x_{\text{теор}}|}{\sqrt{\frac{S^2}{n}}}, \quad (5.13)$$

где  $f = n - 1$ .

Если  $t_\alpha$  расчетное меньше  $t_\alpha$  табличного (см. Приложение Б), то выборки однородны.

При поиске табличного значения для одностороннего критерия берут удвоенный уровень значимости:  $2\alpha$ ; для двустороннего:  $\alpha$ .

Без таблицы можно найти теоретические значения критерия Стьюдента в Excel. Для этого нужно после знака равно выбрать статистическую функцию:

**СТЮДЕНТ.РАСП** – левостороннее *t*-распределение Стьюдента. Оно используется чаще всего. Вводится расчетное значение *t*-критерия, количество степеней свободы и опция (0 или 1), определяющая, что нужно рассчитать: плотность или значение функции.

**СТЮДЕНТ.РАСП.2X** – двусторонне распределение. В качестве аргумента выбирается абсолютное значение (по модулю) *t*-критерия и количество степеней свободы.

**СТЮДЕНТ.ТЕСТ** – функция для проверки гипотезы о равенстве математических ожиданий в двух выборках.

*Критерий Фишера* (формула (5.14)) применяется для проверки равенства дисперсий двух выборок. Его относят к критериям рассеяния. При проверке гипотезы о равенстве средних значений в двух выборках с использованием критерия Стьюдента имеет смысл предварительно проверить гипотезу о равенстве дисперсий при помощи критерия Фишера. Он также применяется для проверки точности приборов, методов, при оценке значимости коэффициентов уравнения регрессии. Статистика теста сводится к отношению выборочных дисперсий. Чтобы применить критерий Фишера, необходимо, чтобы числитель и знаменатель были независимыми случайными величинами и имели нормальное распределение. Кроме того, предполагается, что дисперсия случайных величин, квадраты которых суммируются, одинакова.

$$F_{\alpha; f_1; f_2} = \frac{S_{max}^2}{S_{min}^2}, \quad (5.14)$$

где  $f_1 = n_{\max} \text{ дисперсии} - 1$  (число степеней свободы числителя);  $f_2 = n_{\min} \text{ дисперсии} - 1$  (число степеней свободы знаменателя).

Если  $F$  расчетное меньше  $F$  табличного (см. Приложение Г), то выборки однородны.

Без таблицы можно найти теоретические значения критерия Фишера в Excel. Для этого нужно после знака равно выбрать статистическую функцию: ФРАСПОБР (вероятность; степень свободы 1; степень свободы 2).

*G-критерий Кохрена* (формула (5.15)) используют при сравнении трёх и более выборок одинакового объёма:

$$G_{\alpha; f_1; f_2} = \frac{S_{max}^2}{\sum_{i=1}^N S_i^2}, \quad (5.15)$$

где  $f_2$  – количество сравниваемых дисперсий;  $f_1 = n - 1$  (число степеней свободы числителя).

Если  $G$  расчетное меньше  $G$  табличного (см. Приложение Д), то выборки однородны.

**Задание 5.1.** Необходимо определить молярность раствора хлористого натрия. При этом с помощью прямых измерений получены значения числа молей и объема:  $v = (3,5 \pm 0,5)$  моль,  $\delta v = 5 \%$ ;  $V = (150 \pm 10)$  л,  $\delta V = 6 \%$ .

**Задание 5.2.** Пусть при измерениях проведено 10 отсчетов массы продукта. Результаты измерения представлены в таблице 5.5. Определить грубые промахи (выбросы) и абсолютную погрешность (расширенную неопределенность) измерений.

Таблица 5.5 – Результаты измерения массы Т, г

Вариант	1	2	3	4	5	6	7	8	9	10
1	120	125	112	143	134	129	127	131	129	128
2	324	327	320	325	330	319	322	326	326	329
3	542	541	536	539	550	527	538	544	542	540
4	126	128	123	128	124	120	117	135	127	125
5	103	98	95	97	99	93	95	105	101	100
6	623	627	629	617	620	619	622	629	628	626
7	283	298	285	282	284	285	289	280	278	285
8	445	447	440	437	442	445	444	446	442	440
9	160	146	163	162	161	159	149	170	165	162

**Задание 5.3.** Пяти лабораториям поручено исследовать содержание золы в мышечной ткани образца говядины. Образец был разделен на 40 частей, в каждую лабораторию отправлено по 8 кусков. Дисперсии результатов измерений в разных лабораториях получились следующими: 3,86; 4,27; 1,35; 3,90; 1,64. Проверить гипотезу межлабораторной сходимости результатов.

**Задание 5.4.** Согласно паспортным данным реактив едкий натрий ЧДА должен содержать не менее 99 % основного вещества. Требуется проверить гипотезу соответствия данного реактива паспортными данными по результатам,



полученным в ходе трех определений содержания NaOH в реактиве: 98,3 %; 97,3 %; 97,8 %.

**Задание 5.5.** Разрабатывается рецептура фруктового желе. Необходимо установить, существенно ли уменьшится время застывания желе при добавлении в рецептуру дорогостоящего компонента, т.е. насколько оправдано введение данного компонента в состав рецептуры. В таблице 5.6 приведены значения времени застывания желе без и с дополнительным компонентом. Можно ли сказать, что время застывания значительно уменьшилось?

Таблица 5.6 – Рецептуры желе

№ п/п	Первоначальная рецептура, время, мин	Новая рецептура, время, мин
1	19	21
2	28	15
3	14	11
4	29	12
5	15	21
$\bar{X}$	21	16
$S^2$	50,5	23

**Задание 5.6.** Проводятся испытания шести образцов копченой сельди на соленость (таблица 5.7). От каждого образца взято по одной пробе, которую исследуют параллельно два лаборанта. Необходимо определить, наблюдаются ли различия между результатами испытания двух лаборантов.

Таблица 5.7 – Результат испытаний копченой сельди на соленость, %

Образец	1	2	3	4	5	6
Лаборант 1	5,250	4,975	5,050	5,075	4,795	5,190
Лаборант 2	5,230	4,980	5,020	5,085	4,750	5,120

### Контрольные вопросы

1. Чем отличаются понятия «погрешность» и «неопределённость»?
2. Как вычислить неопределённость прямых и косвенных измерений?
3. Как определить грубые ошибки?
4. Когда применяют критерий Стьюдента?
5. Назовите условия применения критериев Фишера и Кохрена.

### Лабораторная работа № 6

#### Построение и анализ регрессии в MS Excel

**Цель** – получить навыки построения регрессии графически в MS Excel, а также ее анализа на адекватность.

#### Задание:

1. Изучить теоретический материал, примеры построения и анализа регрессии в MS Excel.

2. Решить задачи построения и анализа регрессии, используя MS Excel.

### Теоретический материал, методические указания и примеры выполнения заданий

Термины «регрессия» и «корреляция» – основные в статистике. Основная задача корреляционного анализа – выявление значимости связи между значениями различных случайных величин. Необходимо разделять понятия регрессионной и функциональной зависимостей. Зависимость между величинами, при которой одному значению одной величины (аргумента) отвечает одно или несколько определенных значений другой величины (функции), называется *функциональной*. Зависимость между величинами, при которой каждому значению одной величины отвечает с соответствующей вероятностью множество возможных значений другой, называют *вероятностной* (статистической, регрессионной).

*Корреляция* – статистическая взаимосвязь двух или нескольких случайных величин. Она не является причинно-следственной.

*Аппроксимация* – приближение – замена одних объектов другими, более простыми, но близкими к исходным.

Для вычисления тесноты связи между фактором и откликом для однофакторной линейной регрессии используют *коэффициент корреляции Пирсона*. Для вычисления коэффициента корреляции  $r$  применяют формулу (6.1).

$$r(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6.1)$$

Значения коэффициента корреляции изменяются от -1 до 1. Чем ближе по модулю значение к 1, тем сильнее связь, чем ближе к 0, тем связь слабее. Отрицательный знак показывает на наличие обратно пропорциональной зависимости.

Для многофакторных регрессий используют множественный коэффициент корреляции  $R^2$  (коэффициент детерминации для линейных регрессий и индекс детерминации для нелинейных регрессий), рассчитываемый по упрощенной формуле (6.2).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{\sum_{i=1}^n (y_i - \bar{y})} \quad (6.2)$$

где  $\hat{y}_i$  – рассчитанное по уравнению регрессии значение отклика.

$R^2$  показывает, насколько предсказание по уравнению регрессии лучше, чем предсказание по среднему значению отклика  $y$ .  $R^2$  характеризует долю разброса отклика, описываемую регрессией ( $1 \geq R^2 \geq 0$ ). Чем ближе  $R^2$  к единице, тем лучше уравнение регрессии описывает экспериментальные данные.

Если исследуется однофакторная регрессия, уравнение можно рассчитать при помощи программного пакета MS Excel. Для этого в столбец А вводят данные переменного параметра  $x$ , в столбец В – данные  $y$ . Выделив данные в столбцах А и Б, нажимают «Вставка» - «Диаграмма» - «Точечная» (рисунок 6.1).

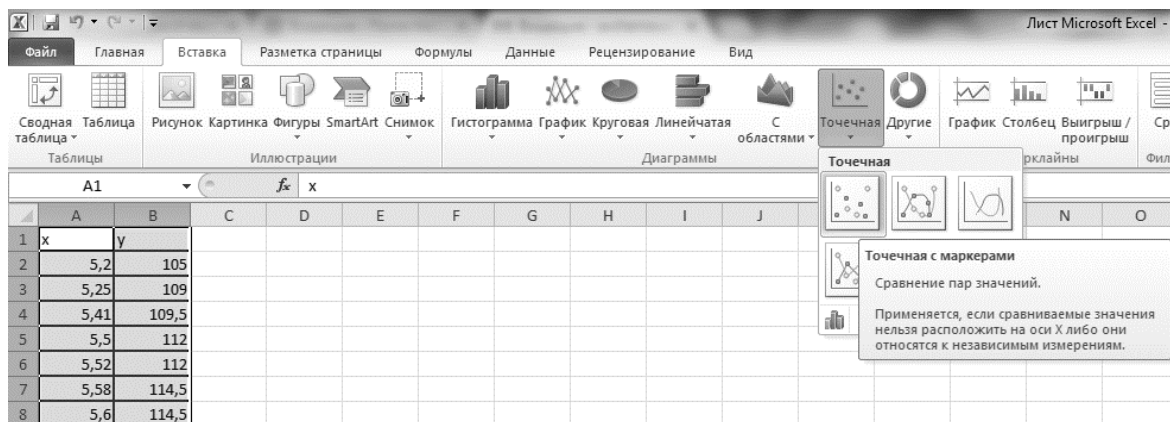


Рисунок 6.1 – Построение точечной диаграммы

Такая диаграмма называется *диаграммой рассеяния*. Двойным щелчком по осям можно вызвать вкладку «Формат оси» и откорректировать внешний вид представления данных (рисунок 6.2).

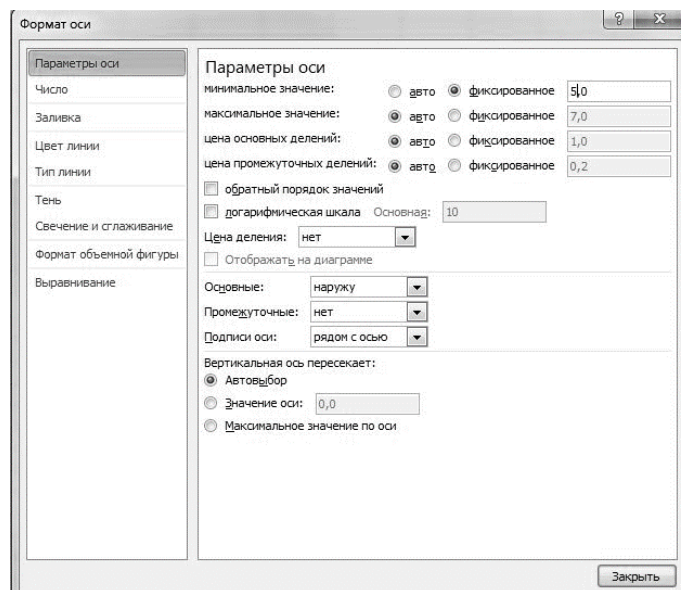


Рисунок 6.2 – Вкладка корректировки формата оси

Щелкая правой кнопкой по любой точке диаграммы, выбирают строку «Добавить линию тренда» (рисунок 6.3).

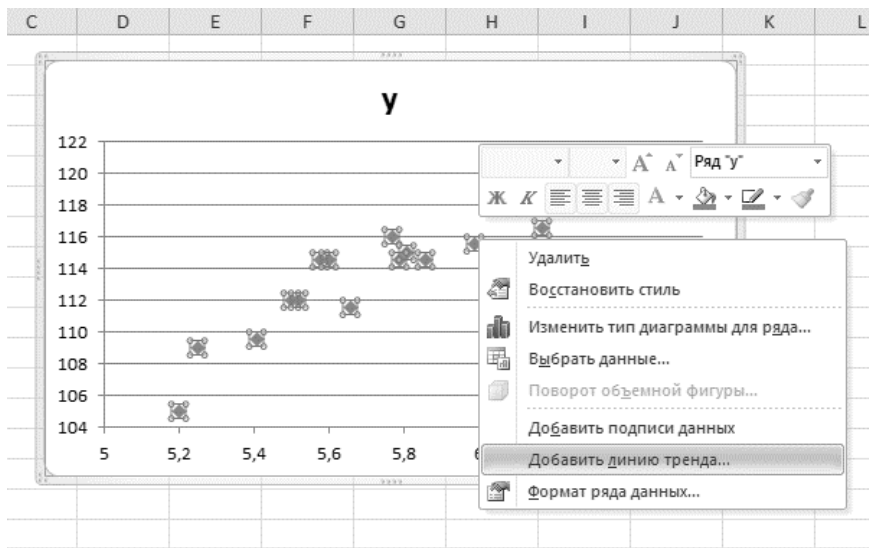


Рисунок 6.3 – Добавление линии тренда на диаграмму

Откроется окно «Формат линии тренда», в котором выбирается вид уравнения (линейное, логарифмическое, степенное, экспоненциальное, полиномиальное, линейная фильтрация). Необходимо поставить галочки в окнах «Показать уравнение на диаграмме» и «Поместить на диаграмме величину достоверности аппроксимации» (рисунок 6.4).

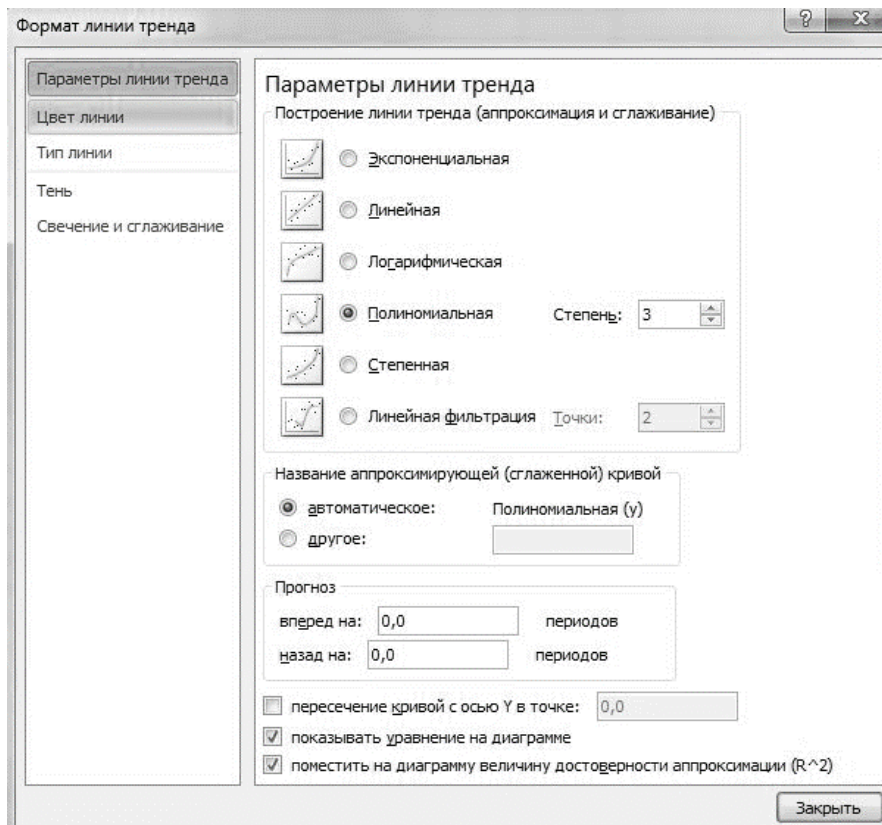


Рисунок 6.4 – Формат линии тренда

При этом на графике будет рассчитано и приведено уравнение, показана величина  $R^2$  и линия графика – *кривая регрессии* (рисунок 6.5).

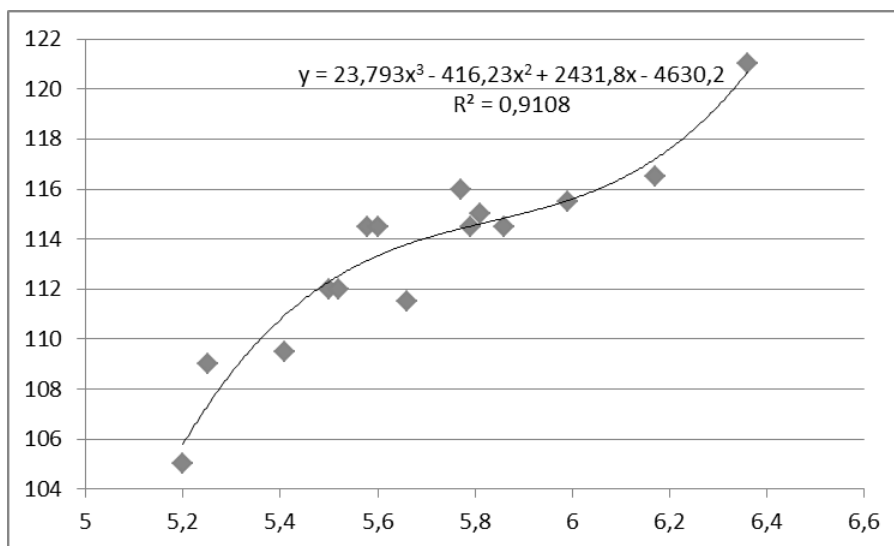


Рисунок 6.5 – Построение кривой регрессии

В случае недостаточной достоверности аппроксимации можно задаться другим видом уравнения и повторить расчеты.

Построить линейное уравнение регрессии для любого количества факторов можно и при помощи надстройки «Анализ данных», выбрав пункт «Регрессия» (рисунок 6.6).

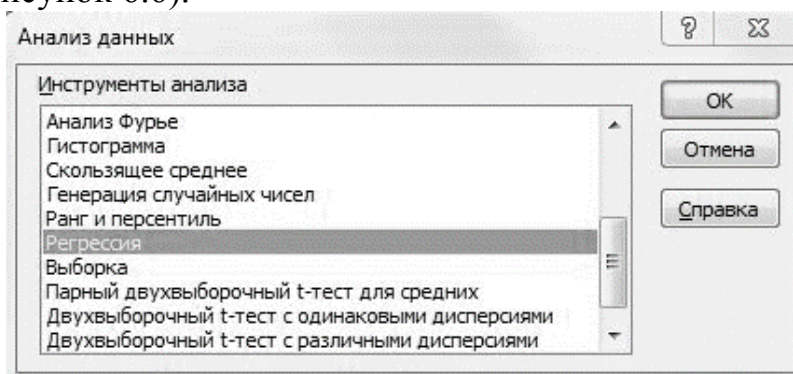


Рисунок 6.6 – Надстройка «Анализ данных»

Поставив курсор в окне «Входной интервал y», необходимо выделить значения столбца с данными y (столбец B) – рисунок 6.7.

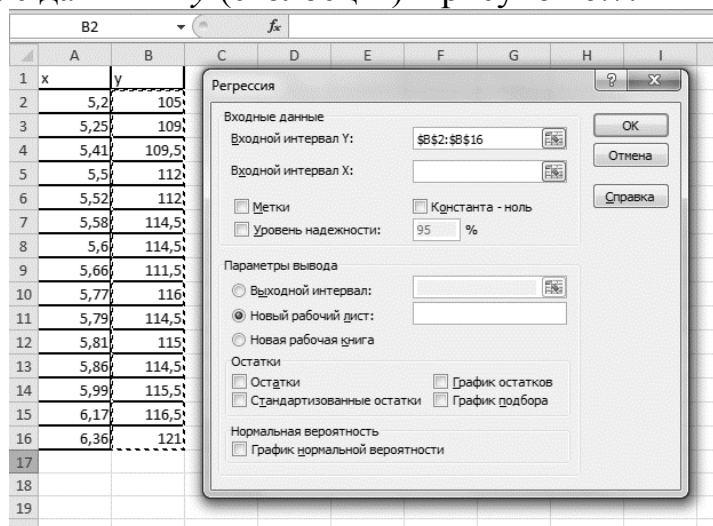


Рисунок 6.7 – Выбор входного интервала y

Аналогично поступают с «входным интервалом x». Кроме этого, необходимо поставить галочки в пунктах «Остатки», «График остатков», «График нормальной вероятности», которые нужны для проверки уравнения регрессии на адекватность (рисунок 6.8).

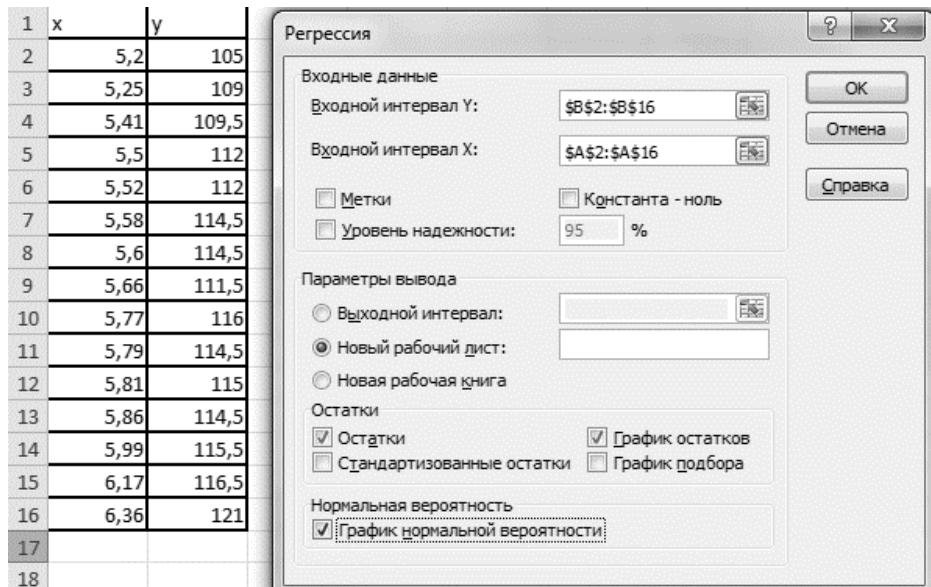


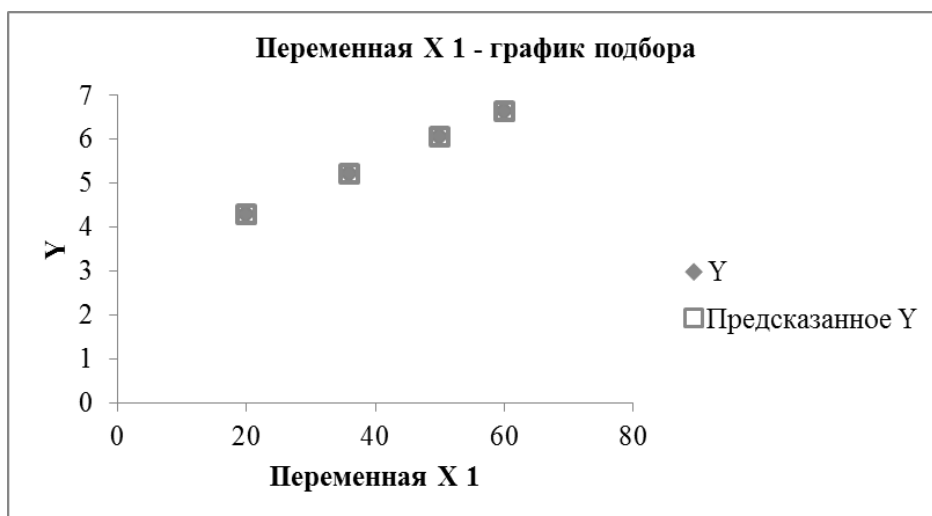
Рисунок 6.8 – Выбор входного интервала x

На листе появится таблица «Вывод итогов», где особое внимание необходимо уделить графам  $R$ -квадрат и «у-пересечение», «переменная  $x_1$ ». «У-пересечение» показывает коэффициент при свободном члене уравнения линейной регрессии, «переменная  $x_1$ » - при  $x_1$  (рисунок 6.9).

ВЫВОД ИТОГОВ		
<i>Регрессионная статистика</i>		
Множественный R	0,914554	
<b>R-квадрат</b>	<b>0,836408</b>	
Нормированный R-квадрат	0,823824	
Стандартная ошибка	1,58088	
Наблюдения	15	
<i>Дисперсионный анализ</i>		
	<i>df</i>	<i>SS</i>
Регрессия	1	166,1107
Остаток	13	32,48934
Итого	14	198,6
<i>Коэффициенты</i>		
У-пересечение	51,90711	7,553703
Переменная X 1	10,79201	1,323739

Рисунок 6.9 – Вывод итогов

Программа производит вычисление остатков, т. е. разность между наблюдаемым и предсказанным значениями зависимой переменной  $Y$  при заданном значении независимой переменной  $X$ . На графике подбора можно визуально наблюдать данную разность (рисунок 6.10).



ВЫВОД ОСТАТКА			
<i>Наблюдение</i>	<i>Предсказанное Y</i>	<i>Остатки</i>	<i>Стандартные остатки</i>
1	6,639448732	0,000551268	0,281974814
2	6,049206174	0,000793826	0,406043732
3	5,222866593	-0,002866593	-1,466269032
4	4,278478501	0,001521499	0,778250486

Рисунок 6.10 – График подбора

Для оценки пригодности эмпирической модели регрессии остатки откладываются по вертикальной оси, а значения  $X$  - по горизонтальной. Если эмпирическая модель пригодна, график не должен иметь ярко выраженной закономерности. Если же модель регрессии непригодна, на рисунке проявится зависимость между значениями  $X$  и остатками (рисунок 6.11)

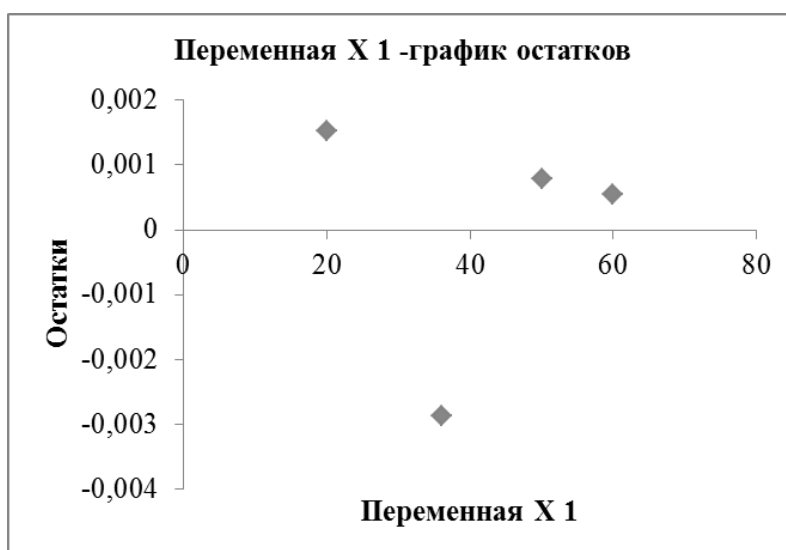


Рисунок 6.11 – График остатков

Как видно из рисунка 6.11, явно выраженной зависимости нет, значит, уравнение регрессии адекватно. Чтобы проверить предположение о нормальном распределении ошибок, строят график нормального распределения на основе точечного графика, на вертикальной оси которого отложены значения остатков, а на горизонтальной оси - соответствующие квантили стандартизованного нормального распределения (рисунок 6.12).

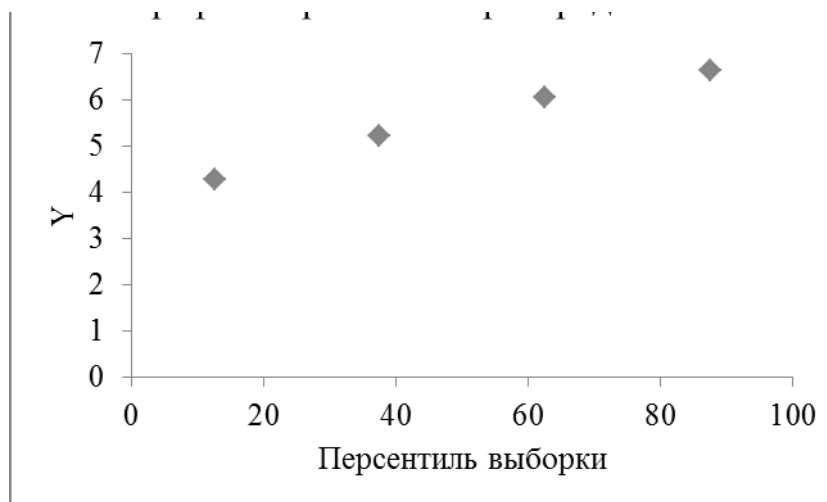


Рисунок 6.12 – График нормального распределения

После построения уравнения регрессии проводят *интерпретационную* оценку, насколько результаты соответствуют здравому смыслу и существующей информации о поведении и свойствах объекта исследования. Интерпретация включает анализ значений коэффициентов регрессии и знаков перед ними. Технологические процессы пищевой промышленности не относятся к объектам макромира, поэтому, если значения коэффициентов регрессии подозрительно велики, это может быть следствием ошибки в расчетах. Знак перед коэффициентом регрессии описывает прямое или обратное пропорциональное влияние входных переменных на отклик.

После интерпретации необходимо проверить уравнение регрессии на адекватность. Для проверки значимости коэффициентов регрессии проводят параллельные эксперименты для любой точки матрицы планирования. Находят их среднее значение и среднеквадратичное отклонение  $S_y$ . Далее определяют дисперсию коэффициентов уравнения регрессии по следующей формуле (6.3):

$$S_{b_i} = \frac{S_y}{m}, \quad (6.3)$$

где  $m$  – количество членов в уравнении регрессии (кроме  $b_0$ ).

Для оценки значимости коэффициентов по доверительному интервалу вычисляют доверительный интервал для коэффициента по формуле (6.4).

$$\Delta b_i = \pm t S_{b_i}, \quad (6.4)$$

где  $t$  – критерий Стьюдента (табличное значение критерия при 5 %-ном уровне значимости и при количестве степеней свободы  $f = n - 1$ ).



Доверительный интервал одинаков для всех коэффициентов. Коэффициент значим, если его абсолютная величина больше доверительного интервала:  $\Delta b_i < b_{ji}$

Незначимые коэффициенты исключаются из модели. При этом, если коэффициенты модели не коррелированы между собой, исключение незначимых коэффициентов не скажется на остальных коэффициентах. В обратном случае оставшиеся коэффициенты пересчитываются заново.

Адекватность модели определяют по критерию Фишера. Экспериментальное значение F-критерия (критерия Фишера) вычисляют по формуле (6.5):

$$F_{расч} = \frac{S_{ад}^2}{S_y^2}, \quad (6.5)$$

где  $S_{ад}^2$  – дисперсия адекватности модели (остаточная дисперсия);  $S_y^2$  – дисперсия нового опыта (эксперимента).

Дисперсию адекватности вычисляют по формуле (6.6):

$$S_{ад}^2 = \frac{\sum_{i=1}^N \Delta Y^2}{f} = \frac{\sum_{i=1}^N (Y_{i\text{ расч}} - Y_{i\text{ эксп}})^2}{N - (m + 1)}, \quad (6.6)$$

где  $f$  – число степеней свободы;  $m$  – количество значимых коэффициентов модели (в уравнении регрессии, кроме  $b_0$ );  $N$  – количество опытов.

Табличное значение F-критерия для уровня значимости  $p = 0,05$  зависит от  $f_1$  (число степеней свободы большей дисперсии),  $f_2$  (число степеней свободы меньшей дисперсии). Сравнивают расчетные и табличные значения критериев Фишера: если  $F_{расч} \leq F_{табл}$  – модель адекватна.

**Задание 6.1.** В таблице 6.1 дана матрица эксперимента. При помощи программы MS Excel составить уравнение регрессии, определить коэффициент  $R^2$ , составить и проанализировать график остатков.

Таблица 6.1 – Матрица эксперимента

Х	У
6,64	3,29
6,55	3,24
6,37	3,14
6,05	3,01
Х	У
5,22	2,07
4,28	1,8

**Задание 6.2.** Дана матрица планирования полного факторного эксперимента – таблица 6.2:

Таблица 6.2 – Матрица планирования эксперимента

№ опыта	$x_1$	$x_2$	$x_3$	$x_1x_2$	$x_1x_3$	$x_2x_3$	$Y_{\text{эксп}}$	$Y_{\text{расч}}$
1	+	+	+	+	+	+	120	121
2	-	+	+	-	-	+	75	74
3	+	-	+	-	+	-	54	53
4	-	-	+	+	-	-	47	46
5	+	+	-	+	-	-	81	82
6	-	+	-	-	+	-	77	78
7	+	-	-	-	-	+	112	111
8	-	-	-	+	+	+	50	51

Были найдены коэффициенты уравнения регрессии, и уравнение приобрело вид:

$$Y = 130 + 24x_1 + 18x_2 + 15x_3 - 7x_1x_2 - 27x_1x_3 - 9x_2x_3.$$

Для проверки значимости коэффициентов регрессии проведены параллельные эксперименты для точки 5:  $Y_{21} = 80$ ;  $Y_{21} = 77$ ;  $Y_{21} = 83$ . Определить значимость коэффициентов уравнения регрессии. Проверить уравнение регрессии на адекватность.

### Контрольные вопросы

1. Что такое регрессия?
2. Что такое корреляция?
3. Чем отличается функциональная зависимость от статистической?
4. На что указывает коэффициент корреляции?
5. Как проверить уравнение регрессии на адекватность?

### Лабораторная работа № 7

#### Нахождение оптимальных параметров технологических процессов методом крутого восхождения по поверхности отклика

**Цель** – получить навыки определения оптимума параметров технологического процесса или рецептуры на основе математической модели методом крутого восхождения или спуска по поверхности отклика.

#### Задание

1. Изучить теоретический материал, описывающий процедуру метода.
2. Провести нахождение оптимальных параметров, используя MS Excel.

#### Теоретический материал, методические указания и примеры выполнения заданий

Метод крутого восхождения по поверхности отклика используется при наличии математической модели процесса или продукта в виде уравнения регрессии второго порядка.

Сначала выделяется область в факторном пространстве и определяется нулевая точка, где ставится эксперимент, по результатам которого делается

линейное приближение. Далее осуществляется движение по поверхности отклика в направлении градиента линейного приближения. Если необходимо, делается еще одно линейное приближение, и так продолжается до тех пор, пока исследователь не попадет в область, где линейное приближение поверхности отклика оказывается непригодным. В этой области независимых переменных ставится эксперимент для описания локальной области, близкой к экстремуму, как правило, оценивается квадратичная модель поверхности отклика второго порядка. Об эффективности движения по градиенту можно судить по величине параметра оптимизации. Движение по градиенту считается эффективным, если реализация мысленных опытов, рассчитанных на стадии крутого восхождения, приводит к улучшению значения параметра оптимизации по сравнению с самым хорошим результатом в матрице. При эффективном крутом восхождении возможны два исхода: область оптимума достигнута или область оптимума не достигнута.

Область оптимума достигнута - можно закончить исследование.

Область оптимума не достигнута. В этом случае ставится линейный план следующего цикла, и исследование продолжается.

Область оптимума близка. Если при реализации матрицы планирования удалось получить достаточно высокие значения параметра оптимизации и при крутом восхождении улучшить их не удалось, то наиболее типичными являются решения: 1) окончание исследования (выбирается лучший опыт); 2) построение плана второго порядка для описания области оптимума. Если линейная модель была неадекватной, то возможно третье решение – выяснение причины неадекватности линейной модели.

Если область оптимума далека, и линейная модель адекватна, то возможное объяснение - в характере поверхности отклика: она может иметь несколько экстремумов. В таких случаях целесообразно передвинуться в другую область факторного пространства и построить линейный план второго цикла крутого восхождения.

Область оптимума далека, и линейная модель неадекватна. Необходимо выяснить причины неадекватности линейной модели: интервалы варьирования выбраны неудачно; был выбран дробный факторный эксперимент. Целесообразно использовать метод «перевала», т. е. построить матрицу второй серии опытов, изменив все знаки на обратные. Это даст возможность освободить линейные эффекты от совместных оценок с парными взаимодействиями. В случае нелинейности исходной модели можно попытаться преобразовать параметр оптимизации - снизить степень полинома.

Крутое восхождение неэффективно. Положение оптимума неопределенное. Если нет информации о положении оптимума, и на стадии крутого восхождения не удалось улучшить значение параметра оптимизации, можно рекомендовать поставить опыты в центре эксперимента, с тем чтобы оценить вклад квадратичных членов. При значимом вкладе можно приступить к достройке линейного плана до плана второго порядка, так как наличие

квадратичных членов свидетельствует о близости к почти стационарной области.

*Алгоритм крутого восхождения / спуска.* Сначала для каждого фактора находят произведение коэффициента регрессии  $b_i$  на интервал варьирования  $\lambda_i$ , т. е.  $b_i\lambda_i$ . Эти произведения, как и коэффициенты регрессии, могут быть со знаками плюс и минус. Фактор с положительным коэффициентом регрессии следует увеличивать, фактор с отрицательным коэффициентом – уменьшать. Далее определяют запас для движения по каждому фактору  $\Delta_i$  от основного уровня до практически целесообразной величины в направлении ее увеличения или уменьшения ( $\Delta_i = X_{i_{\max}} - X_{i_0}$ ). Затем вычисляют абсолютные значения отношений  $\Delta_i / b_i\lambda_i$  для каждого фактора. Фактор, у которого это отношение окажется наименьшим, берется в качестве базового. Для этого фактора и выбирается шаг крутого восхождения. Обычно берут шаг в интервале  $\Delta/5 - \Delta/8$ .

Шаги для всех остальных факторов выбираются пропорционально их произведениям  $b_i\lambda_i$ . Так, если шаг базового фактора обозначить через  $S$ , а его произведение – через  $b\lambda$ , то шаг любого  $i$ -го фактора рассчитывается по формуле (7.1):

$$S_i = S \frac{b_i\lambda_i}{b\lambda}. \quad (7.1)$$

После этого намечают программу крутого восхождения, в которой в каждом следующем варианте изменяют величину каждого фактора на шаг. В соответствии со знаком коэффициентов регрессии факторы при этом будут либо уменьшаться, либо увеличиваться. Если при математической обработке матрицы планирования некоторые факторы оказались незначимыми, то при расчете программы крутого восхождения их можно брать на том же основном уровне.

**Пример 7.1.** Найти оптимальную рецептуру соуса по плану, приведенному в таблице 7.1. Изменяют концентрацию крахмала  $X_1$ , черного перца  $X_2$ , томата  $X_3$  и муки  $X_4$ .

Таблица 7.1 - План опыта для четырех факторов на двух уровнях

Номер опыта	Натуральные значения факторов, %				Кодированные значения факторов			
	$X_1$	$X_2$	$X_3$	$X_4$	$x_1$	$x_2$	$x_3$	$x_4$
1	5,3	0,5	0,6	1,7	-	-	-	-
2	6,7	0,5	0,6	2,3	+	-	-	+
3	5,3	0,7	0,6	2,3	-	+	-	+
4	6,7	0,7	0,6	1,7	+	+	-	-
5	5,3	0,5	0,9	2,3	-	-	+	+
6	6,7	0,5	0,9	1,7	+	-	+	-
7	5,3	0,7	0,9	1,7	-	+	+	-
8	6,7	0,7	0,9	2,3	+	+	+	+

На основании проведенных исследований и математической обработки полученных данных составили линейное уравнение регрессии:

$$y = 6603 + 283x_1 + 30x_2 + 222x_3 + 515x_4.$$

Исходные данные для осуществления этой программы приведены в таблице 7.2.

Таблица 7.2 - Исходные данные для составления программы крутого восхождения

Показатель	$X_1$	$X_2$	$X_3$	$X_4$
Основной уровень («фон») $x_0$	6,0	0,6	0,75	2,0
Интервал варьирования $\lambda_i$	0,7	0,1	0,15	0,3
Коэффициент регрессии $b_i$	+283	+30	+222	+515
Максимальный уровень $x_{i\max}$	10,0	2,0	2,0	5,0
Минимальный уровень $x_{i\min}$	0,0	0,0	0,0	0,0

Далее определяем произведение  $b_i\lambda_i$ :

$$b_1\lambda_1 = +283 \cdot 0,7 = +198;$$

$$b_2\lambda_2 = +30 \cdot 0,1 = +3;$$

$$b_3\lambda_3 = +222 \cdot 0,15 = +33;$$

$$b_4\lambda_4 = +515 \cdot 0,3 = +155.$$

Для нахождения запаса  $\Delta$  по каждому фактору находим разницу между максимальным значением данного фактора и основным уровнем. Если бы коэффициенты регрессии были отрицательными, то  $\Delta = x_{i\min} - x_0$ :

$$\Delta_1 = 10,0 - 6,0 = 4,0 \text{ \%};$$

$$\Delta_2 = 2,0 - 0,6 = 1,4 \text{ \%};$$

$$\Delta_3 = 2,0 - 0,75 = 1,25 \text{ \%};$$

$$\Delta_4 = 5,0 - 2,0 = 3,0 \text{ \%}.$$

Находим абсолютные значения отношений  $\Delta_i / b_i\lambda_i$ :

$$\frac{\Delta_1}{b_1\lambda_1} = \frac{4}{198} = 0,020;$$

$$\frac{\Delta_2}{b_2\lambda_2} = \frac{1,4}{3} = 0,500;$$

$$\frac{\Delta_3}{b_3\lambda_3} = \frac{1,25}{33} = 0,033;$$

$$\frac{\Delta_4}{b_4\lambda_4} = \frac{3,0}{155} = 0,019.$$

Поскольку наименьшим оказалось отношение  $S = \Delta_4 / b_4\lambda_4$  для муки, этот фактор выбирают в качестве базового при крутом восхождении. Шаг для фактора  $X_4$  принимают равным 0,5 %, тогда шаги для других факторов соответственно:

$$S_1 = 0,5 \frac{198}{155} = 0,64 (0,6\%); S_3 = 0,5 \frac{33}{155} = 0,11 (0,1\%).$$

Для незначимого фактора  $X_3$  шаг при крутом восхождении равен нулю.

Программа крутого восхождения из пяти вариантов представлена в таблице 7.3.

Таблица 7.3 - Программа крутого восхождения

№	$X_1, \%$	$X_2, \%$	$X_3, \%$	$X_4, \%$
1	$6,0 + 0,6 = 6,6$	0,6	$0,75 + 0,11 = 0,86$	$2,0 + 0,5 = 2,5$
2	$6,6 + 0,6 = 7,2$	0,6	$0,86 + 0,11 = 0,97$	$2,5 + 0,5 = 3,0$
3	$7,2 + 0,6 = 7,8$	0,6	$0,97 + 0,11 = 1,08$	$3,0 + 0,5 = 3,5$
4	$7,8 + 0,6 = 8,4$	0,6	$1,08 + 0,11 = 1,19$	$3,5 + 0,5 = 4,0$
5	$8,4 + 0,6 = 9,0$	0,6	$1,19 + 0,11 = 1,30$	$4,0 + 0,5 = 4,5$

После постановки опытов по этой программе за нулевой уровень принимают оптимальную по составу рецептуру и вновь ставят серию экспериментов по ортогональной матрице.

**Задание 7.1.** Составить программу крутого восхождения / спуска для следующих исходных данных (таблица 7.4).

Таблица 7.4 – Данные для составления программы крутого восхождения / спуска

Показатель	$X_1$	$X_2$	$X_3$
Основной уровень $x_0$	150,0	12,0	0,5
Интервал варьирования $\lambda_i$	20,0	1,0	0,1
Коэффициент регрессии $b_i$	12,0	-54,0	-11
Максимальный уровень $x_{imax}$	500,0	36,0	3,0
Минимальный уровень $x_{imin}$	50,0	0,0	0,0

### Контрольные вопросы

1. Приведите алгоритм программы крутого восхождения или спуска.
2. Как определяют запас для движения по каждому фактору?
3. Как определяют шаг восхождения?

### Лабораторная работа № 8

#### Прогнозирование сроков хранения при помощи обобщенной функции желательности Харрингтона

**Цель** – получить навыки прогнозирования сроков хранения продуктов питания при помощи обобщенной функции желательности Харрингтона.

#### Задание:

1. Изучить теоретический материал, описывающий процедуру метода.
2. Провести прогнозирование, используя MS Excel.

#### Теоретический материал, методические указания и примеры выполнения заданий

Множество оптимизационных задач в пищевой технологии многокритериальные, например, когда требуется определить технологию,

обеспечивающую производство продукта с заданными физико-химическими, органолептическими свойствами, составом и сроком хранения. Точная оптимизация по всем этим критериям одновременно практически невозможна, встает проблема принятия компромиссного решения. В пространстве области определения всех критериев существует *область Парето* и *Парето-оптимальные решения*.

Построение области Парето – трудоемкая задача, проще всего ее решить путем сведения к задаче с одним обобщенным критерием. В основе построения обобщенного критерия лежит идея преобразования натуральных значений частных откликов в безразмерную шкалу желательности или предпочтительности. Значение частного отклика, переведенное в безразмерную шкалу желательности, обозначается через  $d_i$  ( $i=1, 2, \dots, n$ ) и называется *частной желательностью*. Шкала желательности имеет интервал от нуля до единицы (таблица 8.1). Значение  $d_i = 0$  соответствует абсолютно неприемлемому уровню данного свойства, а значение  $d_i = 1$  – самому лучшему значению свойства. Значение  $d_i = 0,37$  обычно соответствует границе допустимых значений.

Таблица 8.1 – Стандартные оценки по шкале желательности

Желательность	Оценка по шкале желательности
Отлично	$0,80 \leq D < 1,00$
Хорошо	$0,63 \leq D < 0,80$
Удовлетворительно	$0,37 \leq D < 0,63$
Плохо	$0,20 \leq D < 0,37$
Очень плохо	$0,00 \leq D < 0,20$

Обобщенный показатель желательности (D) рассчитывается как среднее геометрическое частных желательностей отдельных факторов по формулам (8.1-8.2):

- без учета коэффициентов весомости:

$$D = \sqrt[n]{\prod_{i=1}^n d_i} = \sqrt[n]{d_1 \times d_2 \times d_3 \times d_4 \times \dots \times d_n}; \quad (8.1)$$

- с учетом коэффициентов весомости:

$$D = \sqrt[n]{\prod_{i=1}^n d_i^{\beta_i}} = \sqrt[n]{d_1^{\beta_1} \times d_2^{\beta_2} \times d_3^{\beta_3} \times \dots \times d_n^{\beta_n}}, \quad (8.2)$$

где  $\beta$  – коэффициент весомости фактора.

Если к параметру (фактору) предъявляются односторонние ограничения (не более, не менее), то частная желательность определяется по формуле (8.3), а график будет выглядеть как, например, S-функция (рисунок 8.1):

$$d_i = \exp(-e(-y'_i)), \quad (8.3)$$

где  $y'_i$  – безразмерная величина, определяемая, в зависимости от того, ограничение справа или слева, по формуле (8.4):

$$y'_i = b_0 + b_1 \cdot y', \quad (8.4)$$

где  $b_0$  и  $b_1$  можно определить, если задать для двух значений  $y$  соответствующие значения желательности, предпочтительно в интервале от 0,2 до 0,8.

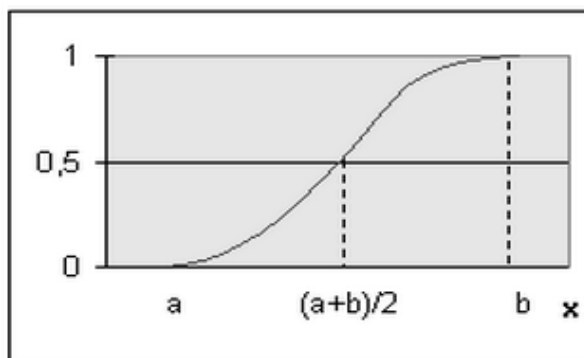


Рисунок 8.1 – График S-функции

Если к параметру (фактору) предъявляются двусторонние ограничения (не более и не менее), то частная желательность определяется по формуле (8.5), а график будет выглядеть как сигмоидная функция (рисунок 8.2).

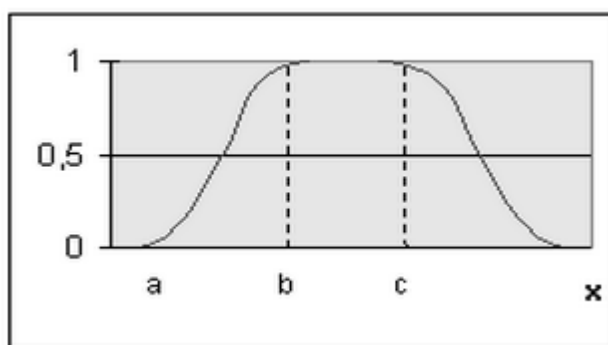


Рисунок 8.2 – График сигмоидной функции

$$d_i = \exp\left[-(|y'_i|)^n\right], \quad (8.5)$$

где  $i$  - текущий номер параметра;

$n$  - показатель степени, который вычисляют, задавшись числом  $d$  в интервале от 0,6 до 0,9 с расчетом  $y'_i$  - безразмерной величины, по формуле (8.6):

$$y'_i = \frac{2y_i - y_{\text{лев}} - y_{\text{прав}}}{y_{\text{прав}} - y_{\text{лев}}}, \quad (8.6)$$

где  $i = 1, 2, 3 \dots n$  - текущий номер параметра;  $y_{\text{лев}}$  - значение  $y$ , принятое в качестве верхнего ограничения;  $y_{\text{прав}}$  - значение  $y$ , принятое в качестве нижнего ограничения.

Затем для расчета  $n$  используют формулу (8.7):

$$n = \frac{\ln \ln \frac{1}{d}}{\ln |y'_i|}. \quad (8.7)$$

**Пример 8.1.** Исследуются сроки сохранения качества соленого полуфабриката из разделанной сельди. В качестве доминантных выбраны следующие характеристики:

- характеристика безопасности по показателю количества мезофильных аэробных и факультативно-анаэробных микроорганизмов (КМАФАнМ);



- характеристика сенсорных свойств по показателям средней органолептической оценки и консистенции;
- характеристика степени проявления эффектов созревания по тенденции накопления азота концевых аминокрупп (АКА);
- характеристика степени порчи соленого полуфабриката по накоплению небелковых азотистых веществ, выраженной в количестве летучих оснований (АЛО).

Значение  $i$ -го частного параметра оптимизации, обозначенное через  $d_i$ , является частным откликом, где  $i = 1, 2, 3, \dots, i$  – текущий номер параметра;  $n$  – количество частных параметров.

$$d_1 = d_2 = \exp(-e(-y'_i)), \quad (8.8)$$

где  $d_1$  – значение частного отклика показателя средней органолептической оценки;  $d_2$  – значение частного отклика показателя консистенции.

Значение показателя  $y$ , принятое в качестве оптимального, – 5 баллов. Значение показателя, принятое в качестве ограничения, – 3 балла.

$$d_3 = \exp(-e(-y'_i)), \quad (8.9)$$

где  $d_3$  – значение частного отклика показателя АЛО.

Значение показателя  $y$ , принятое в качестве оптимального, – 35 мг/100 г. Значение показателя, принятое в качестве ограничения, – 55 мг/100 г.

$$d_4 = \exp(-e(-y'_i)), \quad (8.10)$$

где  $d_4$  – значение частного отклика показателя КМАФАнМ.

Значение показателя  $y$ , принятое в качестве оптимального, – 0 кл/г. Значение показателя, принятое в качестве ограничения, –  $5 \cdot 10^4$  кл/г.

$$d_5 = e^{-\left(\frac{2y_i - y_{\text{лев}} - y_{\text{прав}}}{y_{\text{прав}} - y_{\text{лев}}}\right)^2}, \quad (8.11)$$

где  $d_5$  – значение частного отклика показателя АКА;  $y_{\text{прав}}$  – значение показателя, принятое в качестве верхнего ограничения (120 мг/100 г);  $y_{\text{лев}}$  – значение показателя, принятое в качестве нижнего ограничения (300 мг/100 г);  $y_i$  –  $i$ -е значение  $y$ .

Обобщённая функция желательности:

$$D = \sqrt[n]{\prod_{i=1}^n d_i} = \sqrt[5]{d_1 \times d_2 \times d_3 \times d_4 \times d_5}. \quad (8.12)$$

Значение обобщенной функции желательности  $D = 0,37$  соответствовало границе допустимых значений;  $D = 0$  – абсолютно неприемлемому уровню;  $D = 1$  – самому лучшему значению. Для каждой контрольной точки на протяжении срока хранения был рассчитан обобщенный показатель желательности и построен график его изменения (рисунок 8.3). Момент времени, когда значение данного показателя достигало порогового в 0,37, считался точкой потери желательной характеристики качества.

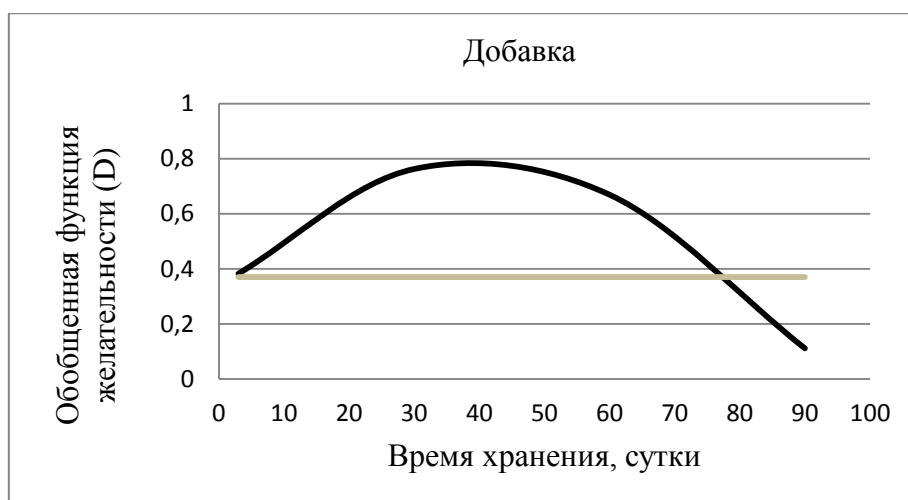


Рисунок 8.3 - График изменения функции желательности в процессе хранения для доминантных характеристик соленого полуфабриката

**Задание 8.1.** С помощью MS Excel рассчитать оптимальные сроки хранения сахарного печенья, если в качестве доминантных выбраны следующие характеристики:

- характеристика безопасности по показателю «количество плесеней» (оптимально 0 кл/г, допустимо 100 кл/г);
- характеристика сенсорных свойств по показателю консистенции (оптимально 5 баллов, допустимо 3 балла);
- характеристика массовой доли влаги (оптимально 7 %, допустимо 10 %).

Результаты экспериментов во время хранения в течение 180 суток представлены в таблице 8.2.

Таблица 8.2 - Результаты эксперимента по хранению сахарного печенья

Сутки	Количество плесеней, кл/г	Консистенция, балл	Массовая доля влаги, %
0	2	4,8	8,4
30	12	4,8	8,5
60	37	4,7	8,5
90	50	4,1	9,6
120	140	4,0	10,5
150	280	3,4	12,1
180	340	3,0	12,2

### Контрольные вопросы

1. Для чего используется функция желательности Харрингтона?
2. Как изменяются оценки желательности?
3. Как вычислить обобщенный показатель желательности?
4. Как вычислить частные отклики желательности?

## Список литературы

1. Аюпов, В. В. Математическое моделирование технических систем: учеб. пособие / В. В. Аюпов. – Пермь: ИПЦ «Прокрость», 2017. – 242 с.
2. Голованов, А. Н. Планирование эксперимента: учеб. пособие / А. Н. Голованов. – Томск: изд-во Томского гос. ун-та, 2011. – 76 с.
3. Горлач, Б. А. Математическое моделирование. Построение моделей и численная реализация: учеб. пособие / Б. А. Горлач, В. Г. Шахов. – Санкт-Петербург: Изд-во «Лань», 2015. – 292 с.
4. Григорьев, Ю. Д. Методы оптимального планирования эксперимента: линейные модели: учеб. пособие / Ю. Д. Григорьев. – Санкт-Петербург: Изд-во «Лань», 2015. – 320 с.
5. Макаричев Ю.А. Методы планирования эксперимента и обработки данных: учеб. пособие / Ю. А. Макаричев., Ю.Н. Иванников. – Самара: Самарский гос. технич. ун-т, 2016. – 131 с.
6. Моисеев, Н. Г. Теория планирования и обработки эксперимента: учеб. пособие / Н. Г. Моисеев, Ю. В. Захаров. – Йошкар-Ола: Поволжский гос. технол. ун-т, 2018. – 124 с.
7. Мусина, О.Н. Планирование и постановка научного эксперимента: учебно-метод. пособие / О. Н. Мусина. – Москва; Берлин: Директ-Медиа, 2015. – 88 с.
8. Юдин, Ю. В. Организация и математическое планирование эксперимента: учеб. пособие / Ю. В. Юдин, И. В. Майсурадзе, Ф. В. Водолазский. – Екатеринбург: Изд-во Уральского ун-та, 2018. – 128 с.

## ПРИЛОЖЕНИЯ

### Приложение А

#### Квантили распределения $\chi^2$ Пирсона

Число степеней свободы	Уровень значимости $\alpha$		
	0,01	0,025	0,05
1	6,6	5,0	3,84
2	9,2	7,4	5,99
3	11,3	9,4	7,82
4	13,3	11,1	9,49
5	15,1	12,8	11,1
6	16,8	14,4	12,6
7	18,5	16,0	14,1
8	20,1	17,5	15,5
9	21,7	19,0	16,9
10	23,2	20,5	18,3
11	24,7	21,9	19,7
12	26,2	23,3	21,0
13	27,7	24,7	22,4
14	29,1	26,1	23,7
15	30,6	27,5	25,0
16	32,0	28,8	26,3
17	33,1	30,2	27,6
18	34,8	31,3	28,9
19	36,2	32,9	30,1
20	37,6	34,2	31,4
21	38,9	35,5	32,7
22	40,3	36,8	33,9
23	41,6	38,1	35,2
24	43,0	39,4	36,4
25	44,3	40,6	37,6
26	45,6	41,9	38,9
27	47,0	43,2	40,1
28	48,3	44,5	41,3
29	49,6	45,7	42,6
30	50,9	47,0	43,8

## Приложение Б

### Квантили распределения Стьюдента

Число степеней свободы	Уровень значимости, $\alpha$		
	<b>0,10</b>	<b>0,05</b>	<b>0,01</b>
<b>f</b>			
1	6,3	12,7	63,7
2	2,9	4,3	9,9
3	2,4	3,2	5,8
4	2,1	2,8	4,6
5	2,0	2,6	4,0
6	1,9	2,4	3,7
7	1,9	2,4	3,5
8	1,9	2,3	3,4
9	1,8	2,3	3,2
10	1,8	2,2	3,2
11	1,8	2,2	3,1
12	1,8	2,2	3,1
13	1,8	2,2	3,0
14	1,7	2,1	3,0
15	1,7	2,1	2,9
16	1,7	2,1	2,9
17	1,7	2,1	2,9

## Приложение В

### Отбор промахов по критерию Шовене

<b>Z</b>	<b>M</b>	<b>Z</b>	<b>M</b>	<b>Z</b>	<b>M</b>	<b>Z</b>	<b>M</b>	<b>Z</b>	<b>M</b>
1,00	2	1,40	3	1,80	7	2,20	18	2,60	54
1,02	2	1,42	3	1,82	7	2,22	19	2,62	57
1,04	2	1,44	3	1,84	8	2,24	20	2,64	60
1,06	2	1,46	3	1,86	8	2,26	21	2,66	64
1,08	2	1,48	4	1,88	8	2,28	22	2,68	68
1,10	2	1,50	4	1,90	9	2,30	23	2,70	72
1,12	2	1,52	4	1,92	9	2,32	25	2,72	77
1,14	2	1,54	4	1,94	10	2,34	26	2,74	81
1,16	2	1,56	4	1,96	10	2,36	27	2,76	87
1,18	2	1,58	4	1,98	10	2,38	29	2,78	92
1,20	2	1,60	5	2,00	11	2,40	30	2,80	98
1,22	2	1,62	5	2,02	12	2,42	32	2,82	104
1,24	2	1,64	5	2,04	12	2,44	34	2,84	111
1,26	2	1,66	5	2,06	13	2,46	36	2,86	118
1,28	2	1,68	5	2,08	13	2,48	38	2,88	126
1,30	3	1,70	6	2,10	14	2,50	40	2,90	134
1,32	3	1,72	6	2,12	15	2,52	43	2,92	143
1,34	3	1,74	6	2,14	16	2,54	45	2,94	152
1,36	3	1,76	6	2,16	16	2,56	48	2,96	163
1,38	3	1,78	7	2,18	17	2,58	51	2,98	173

## Приложение Г

Квантили распределения Фишера при уровне значимости  $\alpha = 0,05$

f2 \ f1	1	2	3	4	5	6	8	12	24	$\infty$
1	161,5	199,5	215,7	224,6	230,2	233,9	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

## Приложение Д

Квантили распределения Кохрена при уровне значимости  $\alpha = 0,05$

Число дисперсий	Число степеней свободы										
	1	2	3	4	5	6	8	10	16	36	$\infty$
2	9985	9750	9392	9057	8772	8534	8159	7880	7341	6602	5000
3	9669	8709	7970	7454	7071	6771	6333	6025	5466	4748	3333
4	9065	7679	6841	6287	5895	5598	5175	4884	4366	3720	2500
5	8412	6838	5981	5441	5065	4783	4387	4118	3645	3066	2000
6	7898	6161	5321	4803	4447	4184	3817	3568	3135	2612	1667
7	7271	5612	4800	4307	3974	3726	3384	3154	2756	2278	1429
8	6798	5157	4377	3910	3595	3362	3043	2829	2462	2022	1250
9	6385	4775	4027	3584	3276	3067	2768	2568	2226	1820	1111
10	6020	4450	3733	3311	3029	2823	2541	2353	2032	1655	1000
12	5410	3924	3264	2880	2624	2439	2187	2020	1737	1403	833
15	4709	3346	2758	2419	2195	2034	1815	1671	1429	1144	0667
20	3894	2705	2205	1921	1735	1602	1422	1303	1108	0879	0500
24	3434	2354	1907	1656	1493	1374	1216	1113	0942	0743	0417
30	2929	1980	1593	1377	1237	1137	1001	0921	0771	0604	0333
40	2370	1576	1259	1082	0968	0887	0795	0713	0595	0462	0250
60	1737	1131	0895	0765	0682	0623	0552	0497	0411	0316	0167
120	0998	0632	0495	0419	0371	0337	0292	0266	0218	0165	0083

Все значения G-критерия меньше единицы, поэтому в таблице приведены лишь десятичные знаки, следующие после запятой, перед которой при пользовании таблицей нужно ставить ноль целых.

Учебное издание

Чернова Анастасия Валерьевна  
Белова Марина Павловна

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПРОЕКТИРОВАНИЯ  
ПРОДУКТОВ ПИТАНИЯ

Редактор Э.С. Круглова

Подписано в печать 11.10.2021 г. Формат 60 x 84/16.  
Уч.-изд. л. 4,6. Печ. л. 4,4. Тираж 20 экз. Заказ № 76

Издательство федерального государственного бюджетного  
образовательного учреждения  
высшего образования  
«Калининградский государственный технический университет»  
230006, Калининград, Советский проспект, 1