

КАЗАНСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

ИМЕНИ

---

А. Н. ТУПОЛЕВА – КАИ

Институт автоматизации и электронного приборостроения

Кафедра электронного приборостроения и менеджмента качества

**ОТЧЕТ ПО КУРСУ**  
**ТЕОРИЯ ИЗМЕРЕНИЙ**

Выполнила: Голубева М.К..

студентка группы 3334

Преподаватель: Бельский А.М.

# 1. Основы математической статистики

**Математическая статистика** – это раздел прикладной математики, в которой с помощью наблюдений и экспериментов стараются выяснить и сформировать законы и характеристики случайных величин. Математическая статистика, опираясь на теорию вероятности, помогает обрабатывать статистические данные во многих областях деятельности человека.

## **Основные задачи математической статистики:**

1. Создание методов сбора и группировки данных, полученных в результате наблюдений за случайными процессами.
2. Разработка методов анализа полученных статистических данных.
3. Получение выводов по данным наблюдений

Для анализа собранных данных нам необходимо их оценить и выразить вероятность события. Также в анализ входят функции распределения вероятностей или плотности вероятностей, оценка параметров известного распределения, оценка связей между случайными величинами.

Основные понятия математической статистики – генеральная совокупность и выборка.

**Генеральная совокупность** – это все множество исследования объектов вида, над которыми проводятся наблюдения. Она может быть бесконечной и конечной.

**Выборкой (выборочной совокупностью)** называется совокупность случайно отобранных объектов из генеральной совокупности, которая необходима для исследования.

Чтобы отчетливо отражать свойства генеральной совокупности, выборка должна быть представительной или репрезентативной. Выборочная совокупность бывает повторной, когда отобранный объект возвращают в генеральную совокупность, и бесповторной, когда объект не возвращают обратно.

Существует несколько способов получения выборки: простой, типический, серийный, механический.

- 1) **Простой отбор** – случайное извлечение объектов из генеральной совокупности с возвратом или без возврата.
- 2) **Типический отбор**, когда объекты отбираются не из всей генеральной совокупности, а из ее «типической» части.
- 3) **Серийный отбор** – объекты отбираются из генеральной совокупности не по одному, а сериями.
- 4) **Механический отбор** - генеральная совокупность «механически» делится на столько частей, сколько объектов должно войти в выборку и из каждой части выбирается один объект.

Для обработки данных наблюдаемые значения случайной величины располагают в порядке возрастания, то есть используют операцию **ранжирования**:

1. **Размах варьирования**  $R = x_{max} - x_{min}$

2. **Мода** ( $M_o^*$ ) – вариант, имеющий наибольшую частоту

3. **Медиана** ( $M_e^*$ ) – значение случайной величины, приходящееся на середину ряда.

Когда изучаемая случайная величина  $X$  является непрерывной или число значений её велико, то составляют **интервальный статистический ряд**.

**Эмпирической функцией** (функцией распределения выборки) называется функция  $F^*(x)$ ,

определяющая для каждого значения  $x$  частоту события  $F^*(x) = \frac{n_x}{n}$ ,

где  $n$  – число выборки,  $n_x$  – число наблюдений, меньших  $x \in \{X < x\}$  ( $x \in \mathbb{R}$ ). При увеличении объема выборки частота события приближается к вероятности этого события.

Эмпирическая функция  $F^*(x)$  является оценкой интегральной функции  $F(x)$  в теории вероятностей.

Функция  $F^*(x)$  обладает теми же свойствами, что и функция  $F(x)$ :

1.  $0 \leq F^*(x) \leq 1$
2.  $F^*(x)$  – неубывающая функция
3.  $F^*(-\infty) = 0, F^*(+\infty) = 1$ .

Для интегральной функции распределения  $F(x)$ :

$F(x + \Delta) - F(x) \approx f(x) * \Delta x$ , где  $f(x)$  – дифференциальная функция распределения (функция плотности вероятности). Под выборочным аналогом функции  $f(x)$  считают функцию:

$f^*(x) = \frac{F^*(x + \Delta) - F^*(x)}{\Delta x}$ , где  $F^*(x + \Delta) - F^*(x)$  – частота попадания наблюдаемых значений случайной величины  $X$  в интервал  $[x; x + \Delta x)$ . Таким образом, значение  $f^*(x)$  характеризует плотность частоты на этом интервале.

**Чтобы изобразить статистическое распределение графически**, используют полигона и гистограммы.

**Полигоном частот** называют ломаную, отрезки которой соединяют точки с координатами  $(x_i; n_i)$ .

Полигон служит для изображения дискретного статистического ряда. Полигон частот является аналогом многоугольника распределения дискретной случайной величины в теории вероятностей.



**Гистограммой частот (частостей)** называют ступенчатую фигуру, состоящую из прямоугольников, основания которых расположены на оси  $Ox$  и длины их равны длинам частичных интервалов ( $h$ ), а высоты равны отношению:

$$\frac{n_i}{h} \text{ - для гистограммных частот;}$$

$$\frac{n_i}{h * n} \text{ - для гистограммы частостей.}$$

Гистограмма является графическим изображением интервального ряда. Площадь гистограммы частот равна  $n$ , а гистограммы частостей равна 1.

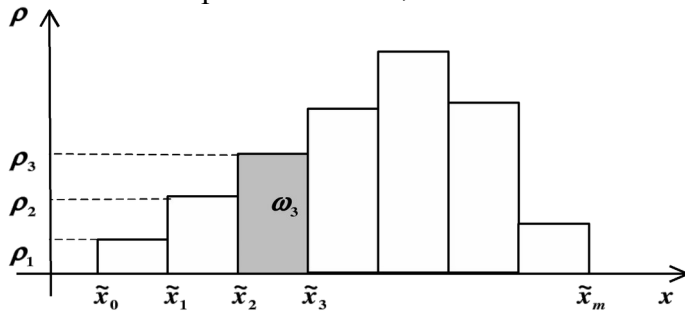


Рис 2. Гистограмма частот.

Можно построить полигон для интервального ряда, если преобразовать его в дискретный ряд. В этом случае интервалы заменяют их серединными значениями и ставят в соответствие интервальные частоты (частости).\_\_

## 2. **Однократные измерения**

Однократные измерения – совокупность операций, при которых необходимо воспроизвести единицу физической величины, сравнить с ней измеряемое значение, зафиксировать результаты сравнения и оценить погрешность измерения.

Однократные измерения можно считать частным случаем многократных измерений, потому что:

- 1) При совершении однократных измерений применяются средства измерений с метрологическими характеристиками, определенными при совершении многократных измерений;
- 2) Представление результатов измерений требует учета всех погрешностей и доверительной вероятности.

Однократные измерения требуют меньше времени, являются менее трудоемкими и более удобными для практического использования. Например, при совершении торговых операций. Необходимо

отметить, что однократные измерения не проще многократных, несмотря на меньшее количество измерений и отсутствие применений вероятностных статистических методов при обработке данных.

Когда проводят однократное измерение, за результат принимают полученное значение величины. Из-за того, что по одному измерению невозможно определить погрешность результата, нужно использовать информацию, которую мы получаем априори: важно правильно выбрать средство измерений для данных условий и определить диапазоны измерений. Например, если мы собираемся измерять массу каких-либо веществ, необходимо познакомиться с весоизмерительным оборудованием и провести предварительный анализ источников погрешностей.

Правильность оценивания погрешностей результата измерений зависит от полноты учета всех факторов, влияющих на результат измерений, и от метода нахождения этих погрешностей. Конечно, в реальной жизни мы не можем учесть все возможные факторы, но всегда требуется дать наиболее точный результат.

На совершение однократных измерений существует свой нормативный документ Госстандарта России – МИ 1552-86 «Государственная система обеспечения единства измерений. Измерения прямые однократные. Оценивание погрешностей результатов измерений». МИ 1552-86 распространяется на нормативную, конструкторскую, технологическую и другую документацию, регламентирующую требования к выполнению измерений, и устанавливает методы оценивания погрешностей результатов прямых однократных измерений.

Измерения проводятся при нормальных условиях (температуры, давления и т.д.). Дополнительные погрешности, зависящие от условий измерения, в этом случае равны нулю. При расширении диапазона измерений или при условиях измерений, отличающихся от нормальных, появляются дополнительные погрешности, которые приходится учитывать.

### **3. Многократные наблюдения**

Если необходима большая уверенность в получаемом результате, то проводятся многократные измерения. Если говорить коротко, то измеряют величину одного и того же размера, результат которого состоит из множества однократных измерений. Все проходит в одинаковых условиях с одними и теми же средствами измерения. Прежде чем приступить к обобщению результатов измерений, определяют, нет ли в полученных результатах грубых погрешностей. За результат многократного измерения обычно принимают среднее арифметическое значение из результатов однократных измерений, входящих в ряд. Такие измерения характерны при выполнении метрологических работ, а также находят широкое применение в научных исследованиях. Важной чертой многократных измерений является получение и использование большого объема измерительной информации. Применение многократных измерений позволяет повысить точность измерения до определенного предела, но недостаток полученной информации не позволяет получить точное значение поправок, значений составляющих погрешностей и т.п. В связи с этим устанавливают необходимое число измерений, которое позволяет получить результат измерений, в котором случайная погрешность

пренебрежимо мала по сравнению с неисключенной систематической погрешностью.

## Лабораторная

## работа

№1

Оценка результатов прямых равноточных многократных измерений

Согласно ГОСТ 8.207-76 обработка результатов наблюдений производится в последовательности:

1. Исключаются известные систематические погрешности из результатов наблюдений.
2. Определяют границы неисключенной систематической погрешности (остатка) результата измерений.
3. Вычисляется среднее арифметическое исправленных результатов наблюдений, которое принимается за результат измерения.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

1. Вычисляется оценка СКО результатов наблюдений.

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}.$$

2. Проверяется наличие в группе наблюдений грубых погрешностей. Если они есть, то их исключают из группы и вновь повторяют вычисление  $\bar{x}$  и  $S_x$ .

$$v_{ext} = \frac{x_{iext} - \bar{x}}{S_x},$$

где  $x_{iext}$  —экстремальные результаты наблюдений.

Полученные результаты  $v_{ext}$  сравним с приложением 1.

$q$	$n = 4$	$n = 6$	$n = 8$	$n = 10$	$n = 12$	$n = 15$	$n = 20$
0,01	1,73	2,16	2,43	2,62	2,75	2,90	3,08
0,02	1,72	2,13	2,37	2,54	2,66	2,80	2,96
0,05	1,71	2,10	2,27	2,41	2,52	2,64	2,78
0,10	1,69	2,00	2,17	2,29	2,39	2,49	2,62

6. Вычисляют оценку СКО среднего арифметического.

7. Проверяют гипотезу о принадлежности результатов наблюдений к нормальному закону распределения.

Проверку нормальности, согласно ГОСТ 11.006–74, при  $n > 50$ , проводят по критериям Пирсона или Мизеса–Смирнова.

При  $15 < n < 50$  используют двойной составной  $q$ -критерий.

8. Вычисляют доверительные границы случайной погрешности(случайной составляющей погрешности).  $\Delta_{сл} = t \cdot S_x$ ,

где  $t$  - коэффициент Стьюдента.

9. Вычисляют доверительные границы погрешностей результата измерения.

Доверительную вероятность  $P_{дов}$  принимают такой же, что при вычислении границ  $\Delta_{сл}$ .

Результаты измерений:

0,43	0,59	0,57	0,46	0,37	0,46	0,45	0,66	0,92	1,15	
------	------	------	------	------	------	------	------	------	------	--

Систематические погрешности – 0,92 и 1,15.

$n = 8, \bar{x} = 0,498 \approx 0,5,$

$3,71 * 0,097 = 0,36$

$$S_x = \sqrt{\square}, \Delta_{сл} =$$

$$(P_{дов} = 0,99, t = 3,71).$$

$$\bar{x} = 0,5 \pm 0,36$$

## Лабораторная работа №2

## Определение класса точности осциллографа

Класс точности – обобщенная характеристика средств измерений, определяемая пределами допускаемых основных и дополнительных погрешностей, а также рядом других свойств, влияющих на точность осуществляемых с их помощью измерений.

Имеется два подхода к нормированию погрешностей средств измерений.

1. Единые правила установления пределов допускаемых погрешностей показаний по классам точности регламентирует ГОСТ 8.401-80. Под классом точности СИ понимают их обобщенные характеристики, определяемые пределами допускаемой основной и дополнительной погрешности. При этом нет деления на погрешность систематическую и случайную. ГОСТ 8.401-80 не устанавливает классы точности СИ, для которых предусмотрены нормы отдельно для систематической и случайной составляющих погрешностей, а так же если необходимо их учитывать динамические характеристики. Классы точности устанавливаются в тех случаях, когда погрешности СИ могут быть выражены числом или сравнительно простой формулой. Второй метод сформулирован ГОСТ 8.009-84. “Нормирование и использование метрологических характеристик средств измерений” Данный стандарт устанавливает комплекс метрологических характеристик, которые должны быть известны при выпуске СИ.

Формула выражения основной погрешности	Пределы допускаемой основной погрешности	Обозначение класса точности	
		в документации	на приборе
Абсолютная $\Delta = \pm a;$ $\Delta = \pm (a + bx)$	$\pm a;$ $\pm (a + bx)$	L M	L M
Приведенная $\gamma = \frac{\Delta}{X_N} 100\% = \pm p$	$\gamma = \pm 1,5$	1,5	1,5
Относительная $\delta = \frac{\Delta}{x} 100\% = \pm q$	$\delta = \pm 0,5$	0,5	(0,5)
Относительная $\delta = \pm \left[ c + d \left( \left  \frac{X_K}{x} \right  - 1 \right) \right]$	$\delta = \pm 0,02 / 0,01$	$c/d = 0,02 / 0,01$	0,02 / 0,01

Результаты измерений:

3	2,9	2,9	2,9	3	3	3	3	3	3
---	-----	-----	-----	---	---	---	---	---	---

$$n = 8, \bar{x} = \frac{3 \cdot 7 + 2,9 \cdot 3}{10} = 2,97,$$

$$S_x = \sqrt{\square}, \Delta_{\text{ср}} =$$

$$3,3 \cdot 0,48 = 0,159.$$

$$\Delta = \sum_{\square} \Delta_i = \frac{7 \cdot (3 - 3) + 3 \cdot (3 - 2,9)}{10} = 0,03$$

$$K = \frac{\Delta}{x_n} \cdot 100\% = \frac{0,03}{3} \cdot 100 = 1\%$$

Выполнив вычисления можно сделать, что осциллограф относится к 6 классу точности..

#### 4. Косвенные измерения

Косвенными называют такие измерения, при которых числовое значение измеряемой величины определяется по известной функциональной зависимости через другие величины, которые можно прямо измерить. При косвенных измерениях числовое значение измеряемой величины получают с участием оператора на основе прямых измерений – решением одного уравнения. К косвенным измерениям прибегают в тех случаях, когда неудобно или невозможно осуществить автоматическое вычисление известной зависимости между одной или несколькими входными величинами и измеряемой величиной. Например, мощность в цепях постоянного тока определяет оператор, умножая напряжение на ток, измеренные прямым измерением с помощью амперметра и вольтметра.

В этом случае искомое значение величины находят на основе измерения других величин в соответствии с зависимостью

Результат находят из решения уравнения, выражающего эту зависимость:

$Q = f(X, Y, Z, \dots, W)$ . где  $Q$  – измеряемая величина;  $X, Y, Z, \dots, W$  – величины, размер которых определяется из прямых измерений.

Например, требуется измерить удельное электрическое сопротивление некоторого материала. Так как приборов для прямых измерений удельного сопротивления нет, его можно измерить только косвенно. Для этого воспользуемся уравнением

$$\rho = \frac{RS}{L}$$

где  $\rho$  – удельное сопротивление;  $R$  – электрическое сопротивление;  $S$  – площадь поперечного сечения;  $L$  – длина образца.

Если измерить длину  $L$ , площадь поперечного сечения  $S$  и электрическое сопротивление  $R$ , то можно вычислить и его удельное сопротивление.

Косвенные измерения достаточно часто встречаются в метрологии, где ими пользуются при воспроизведении единиц. Такие измерения позволяют получать более точный результат, чем прямые. Особенно велика роль косвенных измерений в естественных науках, когда реализация



прямых измерений при изучении явлений затруднительна. Например, явления, изучаемые в астрономии, молекулярной и атомной физике и т. д.

Примеры косвенных измерений: определение эффективной мощности двигателя при его испытании на основании прямых измерений крутящего момента и частоты вращения вала двигателя; определение площади фигур или объема тел по прямым измерениям их геометрических размеров.

Методика обработки результатов косвенных измерений приведена в МИ2083-90 «ГСИ. Измерения косвенные. Определение результатов измерений и оценивание их погрешностей».

### **Совместные и совокупные измерения**

Эти измерения характеризуются тем, что значения искомых величин находят из решения системы уравнений, которые связывают искомые величины с некоторыми другими, получаемыми в результате эксперимента посредством прямых или косвенных измерений.

Если измеряемые величины - одноименные, измерения называются совокупными, если неоднородные – совместными.

Целью таких измерений обычно является установление функциональной зависимости между измеряемыми величинами. Это возможно путем замены экспериментальных данных значениями некоторой теоретической функции. Если замена справедлива для всех значений аргумента в заданном интервале, то заменяющую функцию называют аппроксимирующей. Если замена справедлива только для некоторых, дискретных значений аргумента – функцию называют интерполирующей.

Очевидно, что подобные задачи должны решаться в два этапа:

- выбор структуры предполагаемой зависимости, т. е. выбор вида математической модели (линейная, полиномиальная, экспоненциальная и т. п.);
- вычисление параметров (коэффициентов) этой модели.

В связи с этим, при выполнении совместных измерений, во-первых, возникает задача выбора аппроксимирующей зависимости  $y = f(x)$  так, чтобы она наилучшим образом описывала истинную зависимость, во-вторых, необходимо ответить на вопрос – действительно ли выбранная функция наилучшим образом приближается к искомой зависимости и какой мерой можно оценить близость экспериментальной зависимости к истинной.

Одним из возможных вариантов решения подобных задач это применение метода наименьших квадратов (МНК).

В этом методе оценки параметров выбранной модели определяют из условия, что сумма квадратов отклонений расчетных значений аппроксимирующей функции от экспериментальных значений должна быть минимальна. При этом полагают, что результаты измерений  $x_i, y_i$  ( $i = 1, 2, \dots, n$ ) удовлетворяет условиям:

- значения аргументов  $x_i$  известны точно;
- систематические погрешности из  $y_i$  исключены, а результаты измерений  $y_i$  содержат только случайные составляющие погрешности, которые независимы, имеют одинаковые дисперсии и распределены по нормальному закону.

При выполнении этих условий МНК дает несмещенные оценки параметров модели, имеющие минимальные дисперсии.

Пусть искомые значения величин  $a_j$  ( $j = 1, 2, \dots, m$ ) находятся из решения системы линейных уравнений:

$$\begin{aligned}
 y_1 &= a_1 x_{11} + a_2 x_{12} + \dots + a_m x_{1m}, \\
 y_2 &= a_1 x_{21} + a_2 x_{22} + \dots + a_m x_{2m}, \\
 &\dots\dots\dots \\
 y_n &= a_1 x_{n1} + a_2 x_{n2} + \dots + a_m x_{nm},
 \end{aligned}
 \tag{35}$$

здесь  $y_i$  ( $i = 1, 2, \dots, n$ ) – измеряемые значения величины  $y$ ,  $x_{ij}$  – известные значения аргумента.

Систему уравнений (35) перепишем в виде

$$y_i = \sum_{j=1}^m a_j x_{ij} \quad (i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m)
 \tag{36}$$

Так как  $y_i$  измерены с некоторой погрешностью  $\Delta_i$  то действительный результат измерения  $l_i$  будет иметь вид

$$l_i = \sum_{j=1}^m a_j x_{ij} + \Delta_i,
 \tag{37}$$

или

$$\Delta_i = l_i - \sum_{j=1}^m a_j x_{ij}$$

В соответствии с МНК наилучшие оценки для  $a_j$  должны соответствовать условию

$$\sum_{i=1}^n \Delta_i^2 = \min,$$

или

$$\sum_{i=1}^n \left( l_i - \sum_{j=1}^m a_j x_{ij} \right)^2 = \min
 \tag{38}$$

Это условие выполняется, если все частные производные от (2.38) по искомым параметрам  $a_j$  будут равны нулю, т. е.

$$\sum_{i=1}^n \left( l_i - \sum_{j=1}^m a_j x_{ij} \right) x_{ij} = 0 \quad (39)$$

Система уравнений (39) линейна относительно  $a_j$ , решая ее известными методами определяем искомые параметры постулируемой модели (35). Оценки параметров  $\hat{a}_j$  являются состоятельными, несмещенными и эффективными.

Оценка дисперсии случайной погрешности согласно [2.8] будет равна

$$\sigma^2 = \frac{1}{n-m} \sum_{i=1}^n \left( l_i - \sum_{j=1}^m a_j x_{ij} \right)^2, \quad (40)$$

здесь  $l_i$  – измеренные величины, соответствующие  $y_i$ ;  $a_j$  – коэффициенты, найденные по (39).

При обработке экспериментального материала с целью определения функции преобразования в качестве математической модели часто выбирают полином вида:

$$y_1 = a_1 + a_2 x + a_3 x^2 \dots + a_m x^{m-1},$$

а задачей самого измерения является определение коэффициентов  $a_j$  ( $j = 1, 2, \dots, m$ )

Пусть в результате эксперимента получено  $n$  пар чисел  $x_i \rightarrow y_i$ , тогда результат измерений можно представить в виде

$$\begin{aligned} l_1 &= \sum_{j=1}^m a_j x_1^{j-1} + \Delta_1, \\ l_2 &= \sum_{j=1}^m a_j x_2^{j-1} + \Delta_2, \\ &\dots\dots\dots \\ l_n &= \sum_{j=1}^m a_j x_n^{j-1} + \Delta_n. \end{aligned} \quad (41)$$

Тогда в соответствии с МНК и учетом (39) получим:

$$\sum_{i=1}^n \left( l_i - \sum_{j=1}^m a_j x_{iu}^{j-1} \right) x_i^{j-1} = 0. \quad (42)$$

Решение этой системы, линейной относительно  $a_j$ , дает значение оценок этих коэффициентов.

## 5. Пассивное измерение

Если в ходе эксперимента регистрация контролируемых параметров параметров  $Y$  осуществляется в

процессе нормального функционирования, без внесения каких-либо специальных воздействий, то эксперимент называют пассивным.

Особенностью пассивного эксперимента является то, что экспериментатор не может активно воздействовать на исследуемый объект, а сам эксперимент сводится к сбору измерительной информации и ее оптимальной обработке.

Задачей пассивного эксперимента чаще всего является построение математической модели объекта. В зависимости от того, как зависят друг от друга факторы и отклики, модель может быть представлена в виде функциональной зависимости (детерминированная модель) или в виде некоторой диффузной модели, параметры которой могут быть только оценены, т.к. вычисляются на основе статистического материала.

В результате пассивного эксперимента экспериментатор получает пары чисел  $x_i \rightarrow y_i$  ( $i = 1 \dots n$ ), которым на плоскости  $xOy$  соответствуют точки (рис.10). Совокупность всех экспериментальных точек образует поле корреляции.

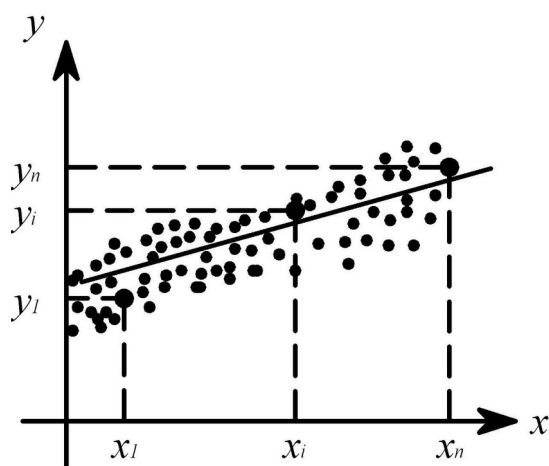


Рис. 10. Графическое изображение уравнения регрессии  $y=f(x)$

Учитывая вероятностный характер полученных данных, предполагаемую зависимость  $y = f(x)$  можно рассматривать только как зависимость математического ожидания  $M[y]$  от  $x$ .

При исследовании статистически связанных зависимостей сталкиваются с двумя случаями.

В первом случае экспериментатор задает (или просто фиксирует) вполне определенным, детерминированные значения независимой переменной  $x_i$ , для которых наблюдается соответствующие случайные значения  $y_i$ , определяемые своими статистическими характеристиками: оценкой математического ожидания и СКО. В этом случае модель называется регрессионной.

Во втором случае наблюдаемые значения  $x_i$  и  $y_i$  представляют собой элементы выборки, т. е. независимая переменная  $x$  сама является случайной величиной и может быть охарактеризована соответствующими оценками. Модель такого вида связи называется корреляционной.

Пусть случайная величина  $y$  зависит от одной или нескольких неслучайных величин  $x_j (j = 1, 2, \dots, k)$ .

Уравнением регрессии называют функциональную зависимость математического ожидания  $M[y]$  от  $x_j$ , т. е.

$$M[y] = \varphi(x_1, x_2, \dots, x_k),$$

а регрессивным анализом - процедуру построения уравнения регрессии и анализ его с помощью аппарата математической статистики. Он включает операции оценивания независимых параметров модели, проверки их статистической значимости, проверки адекватности полученной модели исходному статистическому материалу и т. п.

Видом (структурой) уравнения задаются заранее, исходя из физической сущности изучаемого явления, анализа накопленного статистического материала, на основании априорной информации. Поэтому непосредственной задачей регрессионного анализа является статистическое оценивание неизвестных параметров постулированной модели.

Структура модели определяется видом функции  $\varphi(x_1, \dots, x_k)$ . Если эта функция линейна относительно параметров, то говорят о линейной модели регрессии, если - нет, то модель называется нелинейной.

Простая регрессия предполагает наличие одной независимой и одной зависимой переменной т. е.

$$M[y] = \varphi(x).$$

Если число независимых переменных несколько, то регрессия называется множественной.

Методами регрессионного анализа можно решать следующие задачи:

- описание и установление возможной причинной связи между переменными;
- предсказание (экстраполяцию) зависимой переменной по значениям независимой переменной;

Рассмотрим простую регрессионную связь.

Пусть в результате однофакторного эксперимента получен статистический материал, т. е. выполнено  $n$  пар измерений единственного входного сигнала  $x_i (i = 1, 2, \dots, k)$  и соответствующих значений  $y_i$ .

Будем искать математическую модель регрессионной связи в виде степенного полинома

$$y = b_1 + b_2x + \dots + b_mx^{m-1}. \quad (43)$$

где  $b_1, \dots, b_m$  - искомые параметры модели.

Известно [8], что при правильном выборе степени полинома  $m$  такая модель позволяет достаточно точно аппроксимировать истинную зависимость.

В соответствии с (2.43) для  $i$ -го результата (точки) можно записать

$$y_i = b_1 + b_2 x_i + \dots + b_m x_i^{m-1}. \quad (44)$$

Поскольку измерение  $y_i$  выполнены с погрешностями, то результат  $i$ -го измерения можно записать в виде

$$l_i = y_i + \Delta_i,$$

или

$$\Delta_i = l_i - \sum_{j=1}^m b_j x_i^{j-1}.$$

Используя МНК, т. е. выполняя условие

$$\sum_{i=1}^n \Delta_i^2 \rightarrow \min,$$

получим систему уравнений, линейную относительно искомым параметров  $j$

$$\sum_{i=1}^n \left( l_i - \sum_{j=1}^m b_j x_i^{j-1} \right) x_i^{j-1} = 0, \quad (45)$$

решая которую, получим оценки параметров  $b_j$  постулированной регрессионной модели.

Очевидно, что полученная модель

$$\hat{y} = \hat{b}_1 + \hat{b}_2 x + \dots + \hat{b}_m x^{m-1} \quad (46)$$

может лишь с той или иной точностью соответствовать экспериментальной совокупности  $x_i \rightarrow y_i$ , т. к. МНК не гарантирует точного совпадения отклика, вычисленного по (2.46) и полученного

экспериментально, а только утверждает, что  $\sum_{i=1}^n \Delta_i^2 \rightarrow \min$ . Поэтому заключительным этапом анализа является установление адекватности полученной регрессионной модели экспериментальным данным.

В качестве меры адекватности может быть использован коэффициент корреляции [3].

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (47)$$

чем ближе  $r_{xy}$  к единице, тем ближе экспериментальные точки к (46). Другой мерой адекватности модели является оценка дисперсии ошибки,  $F$  – критерий и др. Более подробно анализ качества полученной модели будет рассмотрен в теме «Активные эксперименты».

## Расчёт ТПП.

Особенностью пассивного эксперимента является то, что экспериментатор не может активно воздействовать на исследуемый объект, а сам эксперимент сводится к сбору измерительной информации и ее оптимальной обработке.

Задачей пассивного эксперимента чаще всего является построение математической модели объекта.

В результате пассивного эксперимента экспериментатор получает пары чисел  $x_i \rightarrow y_i$  ( $i = 1 \dots n$ ), которым на плоскости соответствуют точки. Совокупность всех экспериментальных точек образует поле корреляции.

При исследовании статистически связанных зависимостей сталкиваются с двумя случаями.

В первом случае экспериментатор задает (или просто фиксирует) вполне определенным, детерминированным значения независимой переменной, для которых наблюдается соответствующие случайные значения, определяемые своими статистическими характеристиками: оценкой математического ожидания и СКО. В этом случае модель называется регрессионной.

Уравнением регрессии называют функциональную зависимость математического ожидания  $M[y]$ , а регрессивным анализом - процедуру построения уравнения регрессии и анализ его с помощью аппарата математической статистики.

Результаты измерений:

							$x_{cp}$
	0,28	0,29	0,25	0,22	0,26	0,27	
0,36	0,32	0,39	0,35	0,34	0,31	0,35	
0,43	0,44	0,42	0,45	0,46	0,41	0,45	
0,51	0,52	0,55		0,51	0,53	0,53	

Где  $x_{cp}$  определяется по формуле:

$$x_{cp} = \frac{\sum_{i=1}^n x_i}{n}, \text{ где } n=6$$

Составим систему уравнений вида:  $y=a+b*x$ . Получим:

$$\{ 0,27=a+50*b;$$

$$\{ 0,35=a+60*b;$$

$$\{ 0,45=a+70*b;$$

$$\{ 0,53=a+80*b;$$

Решая эту систему получим что

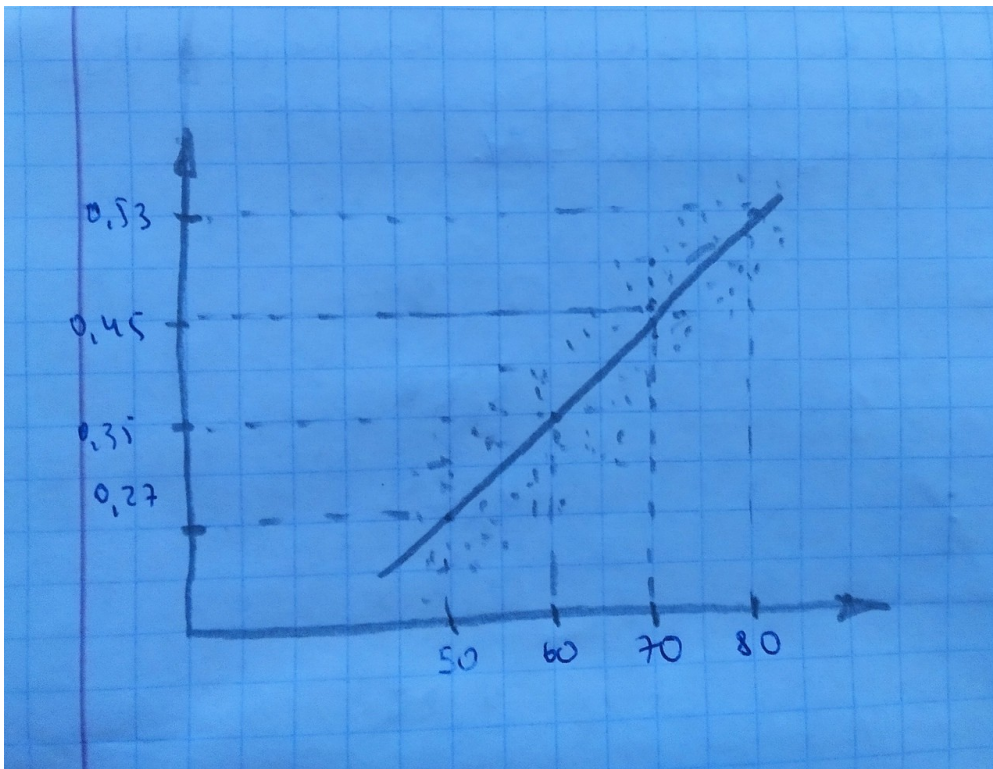
$$b=0,008=\text{const};$$

$$a_1=-0,13$$

$$a_2=-0,13$$

$$a_3=-0,11$$

$$a_4=-0,11$$



6. Активный многофакторный эксперимент (постановка задачи, смысл, статистическая обработка)  
Активный эксперимент подразумевает использование специальных, целенаправленных возмущений (воздействий)  $X$  по заранее определенной программе.  
Каждый способ получения информации имеет свои недостатки и преимущества. Так, например, активный эксперимент не всегда можно осуществить на любом объекте, т. к. это может привести к аварийным ситуациям