

Министерство образования и науки Российской Федерации
Вологодский государственный технический университет

Кафедра высшей математики

Специальные разделы высшей математики

Учебно-методическое пособие
по изучению дисциплины

Факультет экологии

Направление магистратуры: 270800.68 «Строительство»

Профили: Водоснабжение городов и промышленных предприятий;
Водоотведение и очистка сточных вод

Вологда
2013

Специальные разделы высшей математики: учебно-методическое пособие по изучению дисциплины. – Вологда: ВоГТУ, 2013. – 39 с.

Методическое пособие подготовлено в помощь при самостоятельном изучении дисциплины «Специальные разделы высшей математики» студентам, обучающимся в магистратуре по направлению 270800.68 «Строительство» (профили: «Водоснабжение городов и промышленных предприятий», «Водоотведение и очистка сточных вод»). Пособие содержит необходимый теоретический материал, подробное описание практических занятий с примерами выполнения заданий для самостоятельной работы, лабораторный практикум.

Утверждено редакционно-издательским советом ВоГТУ

Составитель О.И. Микрюкова, канд. физ.-мат. наук, доцент

Рецензент Е.А. Лебедева, канд. техн. наук, и. о. заведующего кафедрой водоснабжения и водоотведения

Введение

Целью изучения дисциплины «Специальные разделы высшей математики» является формирование навыков математической обработки результатов экспериментальной работы. Для этого в курсе лекций изложен необходимый теоретический материал, касающийся данной темы. При изложении материала практических занятий рассмотрены отдельные задачи, возникающие при обработке экспериментальных данных, даны задания для самостоятельной работы. Задания лабораторного практикума позволяют студентам потренироваться в применении полученных навыков при описании результатов однофакторного и двухфакторного эксперимента.

Изложение материала дисциплины в данном пособии даёт возможность не тратить время на переписывание лекций, а после прочтения материала сосредоточиться на выполнении заданий для самостоятельной работы. Кроме того, пособие даёт возможность для самостоятельного изучения дисциплины.

Пособие разработано на основе рабочей программы соответствующей дисциплины, в которой указана литература, использованная при его составлении, и которую можно изучать самостоятельно.

Лекция 1. Виды экспериментов. Обработка результатов простейшего эксперимента

1.1. Эксперимент и его виды

Эксперимент – система операций, воздействий и (или) наблюдений, направленных на получение информации об исследуемом объекте.

Простейшим по содержанию экспериментом, но не всегда по выполнению, является измерение какой-либо характеристики изучаемого объекта. Например, плотности и модуля упругости твёрдых тел, плотности и коэффициента вязкости жидкостей.

При проведении эксперимента подобного типа измерение характеристики производится при определённых условиях: температуре, атмосферном давлении, влажности и т.д. Возникает вопрос: как будет меняться значение характеристики при изменении условий? Так мы приходим к эксперименту следующего типа, в котором изучаются зависимости одних величин от других.

Происходящее в таком эксперименте можно описать следующим образом: имеется одна или несколько величин, которые, по предположению, влияют на свойства изучаемого объекта и могут принимать различные значения. Такие величины называются *факторами* (в математике – независимыми переменными).

Если значения факторов могут изменяться по воле исследователя, то речь идёт об *активном* эксперименте. Если же они меняются самостоятельно и исследователь только регистрирует их значения в опыте, речь идёт о *пассивном* эксперименте. Под опытом понимается воспроизведение исследуемого явления в определённых условиях проведения эксперимента при возможности регистрации его результатов.

Под воздействием факторов могут меняться значения интересующей экспериментатора величины, характеризующей свойства объекта. Такая величина называется *откликом* (в математике – функцией).

Например, проводится изучение зависимости количества выделяемой проводником тепловой энергии от сопротивления проводника. Фактор – сопротивление проводника, величину которого экспериментатор может менять при помощи реостата. Эксперимент является активным. Отклик – количество тепла, выделяемое проводником при прохождении электрического тока.

Другой пример – изучение зависимости объёма стока воды в реке от количества осадков. Фактор – количество осадков, которое не зависит от наблюдателя и может им только регистрироваться. Отклик – объём стока воды, эксперимент, естественно, пассивный.

В большинстве случаев изучается зависимость между одной величиной, являющейся фактором, и другой величиной, являющейся откликом. Будем называть такие эксперименты *однофакторными*. При двух факторах и одном отклике будем говорить о *двухфакторном* эксперименте и т.д.

В общем случае и отклик может быть представлен несколькими величинами. Однако проведение эксперимента, описание и анализ его результатов в таком случае является очень непростой задачей, поэтому в дальнейшем будем рассматривать только эксперименты, в которых отклик представлен одной величиной.

Формула, описывающая зависимость значений отклика от значений факторов, называется функцией отклика.

1.2. Обработка результатов простейшего эксперимента

Напомним, что результатом простейшего эксперимента должно быть получение значения какой-либо характеристики объекта исследования. Это значение может быть найдено на основе прямых и косвенных измерений.

При прямых измерениях определяемая величина сравнивается непосредственно с единицей измерения при помощи измерительного инструмента (прибора), например, измерение длины при помощи линейки, массы при помощи рычажных весов, времени при помощи секундомера и т.п. При косвенных измерениях измеряемая величина вычисляется по формуле, в которую входят ре-

зультаты прямых измерений. Например, объём прямоугольного помещения находят как произведение размеров: длины, ширины и высоты. Плотность жидкости или твёрдого вещества находят как частное массы и объёма.

И при прямых, и при косвенных измерениях нельзя указать точного значения находимой величины.

При прямых измерениях это происходит в силу нескольких причин. Во-первых, измеряемая величина не равна целому количеству единиц измерения, пусть даже и очень маленьких. Во-вторых, измеряемая величина очень часто представляет собой некоторую идеализированную характеристику. Например, при измерении длины металлической проволоки предполагается, что она представляет собой цилиндр, основания которого (края проволоки) параллельны друг другу и перпендикулярны образующей цилиндра. Понятно, однако, что изготовить реально такой образец не получится. В-третьих, значения измеряемых величин зависят от условий измерения, которые могут меняться, и от действия различных случайных факторов. Это приводит к тому, что интересующая исследователя величина имеет не одно какое-то значение, а может принимать различные значения на некотором промежутке. Говорят, что величина распределена на некотором промежутке, или имеет распределение.

Поэтому в результате измерения нужно, по существу, указать промежуток, интервал, на котором она принимает свои значения, а также указать её приближенное значение, которое затем используется для нахождения других величин. Другими словами нужно дать точечную и интервальную оценку измеряемой величины.

При косвенных измерениях неточность в нахождении величины вызвана тем, что для её вычисления используются величины, найденные неточно. И в этом случае также искомая величина будет принимать значения на некотором интервале.

1.2.1. Нахождение точечной и интервальной оценки при прямых (непосредственных) измерениях

Обозначим X - измеряемую величину, а x - её точечную оценку, которую также называют приближённым значением X . Интервал, на котором величина X принимает свои значения, имеет вид: $[x - \Delta x, x + \Delta x]$, величина Δx называется абсолютной погрешностью измерения X .

Другими словами, значение величины X находится на промежутке $[x - \Delta x, x + \Delta x]$, что кратко обозначают таким образом:

$$X \in [x - \Delta x, x + \Delta x] \text{ или } X = x \pm \Delta x.$$

Относительной погрешностью δx называется отношение $\frac{\Delta x}{x}$, следовательно $\Delta x = \delta x \cdot x$.

Для величин, измеряемых однократно (вес, длина, время и т.д.), за абсолютную погрешность берут половину цены деления шкалы измерительного инструмента. При этом предполагается, что измеряемая величина существенно не изменяется в процессе измерения, а само измерение производят должным образом.

Если измеряемая величина может под действием разных случайных факторов заметно изменяться от измерения к измерению, то ее значение находят несколько раз. При этом получается набор измеренных значений $x_1, x_2, \dots, x_n$. Абсолютную погрешность и приближённое значение в таком случае можно вычислить так.

Находят наибольшее и наименьшее из всех измеренных значений $x_{\max}$ и $x_{\min}$. За приближённое значение берут $x = \frac{x_{\max} + x_{\min}}{2}$, за абсолютную погрешность $\Delta x = \frac{x_{\max} - x_{\min}}{2}$.

Пример 1. Даны 10 измерений величины X (диаметра цилиндра в мм): 14,85; 14,80; 14,84; 14,81; 14,79; 14,81; 14,80; 14,85; 14,84; 14,80.

Дать точечную и интервальную оценку диаметра цилиндра.

Выбираем наибольшее из указанных значений: $x_{\max} = 14,85$, наименьшее из всех значений: $x_{\min} = 14,79$. Тогда точечная оценка (приближённое значение) равно $x = \frac{14,85 + 14,79}{2} = 14,82$, абсолютная погрешность $\Delta x = \frac{14,85 - 14,79}{2} = 0,03$. Значит, диаметр цилиндра $X = 14,82 \pm 0,03$.

Замечание.

Если величина погрешности Δx , найденная из разброса значений, окажется сравнимой с величиной погрешности измерительного инструмента Δx_u , то общую оценку абсолютной погрешности Δx_o можно найти по формуле: $\Delta x_o = \sqrt{(\Delta x)^2 + (\Delta x_u)^2}$.

1.2.2. Нахождение точечной и интервальной оценки измеряемой величины при косвенных измерениях

Величины, найденные при прямых измерениях, обычно используются в дальнейших вычислениях (косвенных измерениях) для определения других

величин. Естественно, найденные таким образом величины будут также принимать значения на некотором интервале.

Пусть величина Y вычисляется с использованием значений величин $X_1, X_2, \dots, X_n$, т.е. Y можно считать функцией этих величин $Y = f(X_1, \dots, X_n)$. Напомним, что в конкретных задачах для обозначения и функции и переменных могут быть использованы другие символы. Значения $X_1, X_2, \dots, X_n$ найдены с некоторой погрешностью. Т.е. $X_1 = x_1 \pm \Delta x_1$; $X_2 = x_2 \pm \Delta x_2$; . . . $X_n = x_n \pm \Delta x_n$.

Распространённым способом нахождения интервала $[y - \Delta y; y + \Delta y]$, на котором принимает свои значения Y , является способ, основанный на следующей идее: величина абсолютной погрешности функции Δy определяется величиной её полного дифференциала. При этом величина дифференциала изменяется в большую сторону для повышения надёжности.

Приближенное значение (точечную оценку) y находят как $y = f(x_1, \dots, x_n)$.

Формулу для абсолютной погрешности величины Y можно составить следующим образом.

Абсолютная погрешность Δy вычисляется по формуле полного дифференциала функции Y , в котором все частные производные берутся по модулю, а символы дифференциалов независимых величин заменяются на символы их абсолютных погрешностей, сами величины представлены их точечными оценками (приближёнными значениями).

Напомним, что формула полного дифференциала функции Y , обозначаемого dY , имеет вид: $dY = Y'_{x_1} \cdot dx_1 + Y'_{x_2} \cdot dx_2 + \dots + Y'_{x_n} \cdot dx_n$.

Пример. Составим формулу для нахождения абсолютной погрешности суммы и разности двух величин.

Пусть $Y = X_1 + X_2$. Тогда $dY = dX_1 + dX_2$, а $\Delta Y = \Delta x_1 + \Delta x_2$.

Пусть $Y = X_1 - X_2$. Тогда $dY = dX_1 - dX_2$, а $\Delta Y = \Delta x_1 + \Delta x_2$.

Полученный результат можно сформулировать в виде правила: абсолютная погрешность суммы или разности двух величин равна сумме абсолютных погрешностей этих величин.

В некоторых случаях абсолютную погрешность бывает проще найти через относительную погрешность.

Относительная погрешность δY вычисляется по формуле полного дифференциала $\ln Y$, в которой все частные производные берутся по модулю, а символы дифференциалов независимых переменных заменены на символы абсолютных погрешностей, величины представлены их точечными оценками (приближёнными значениями).

Формула полного дифференциала получается аналогично предыдущей:

$$d(\ln Y) = \frac{dY}{Y} = \ln Y'_{x_1} \cdot dx_1 + \ln Y'_{x_2} \cdot dx_2 + \dots + \ln Y'_{x_n} \cdot dx_n.$$

Пример. Найти относительную погрешность произведения и частного двух величин.

Пусть $Y = X_1 \cdot X_2$. Тогда

$$\ln Y = \ln X_1 + \ln X_2, \quad d(\ln Y) = \frac{dX_1}{X_1} + \frac{dX_2}{X_2}, \quad \delta Y = \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} = \delta x_1 + \delta x_2.$$

Пусть $Y = \frac{X_1}{X_2}$. Тогда

$$\ln Y = \ln X_1 - \ln X_2, \quad d(\ln Y) = \frac{dX_1}{X_1} - \frac{dX_2}{X_2}, \quad \delta Y = \frac{\Delta x_1}{x_1} + \frac{\Delta x_2}{x_2} = \delta x_1 + \delta x_2.$$

Результат запомним в виде правила: относительная погрешность произведения и частного двух величин равна сумме их относительных погрешностей.

Аналогично можно получить ещё одно правило: относительная погрешность степени величины равна произведению показателя степени на относительную погрешность величины.

Пользуясь указанными правилами, во многих случаях можно составить формулу для нахождения абсолютной погрешности, не прибегая к дифференцированию функции нескольких переменных.

Пример 2. Для определения плотности металла взвешивается кусок проволоки из этого металла длиной l и радиусом поперечного сечения r . В этом случае плотность ρ определяется по формуле:

$$\rho = \frac{m}{\pi r^2 l}, \text{ где } m - \text{масса взвешиваемого куска проволоки.}$$

Найти точечную и интервальную оценки величины плотности при следующих данных:

$$m = (67,5 \pm 0,1) \text{ г}; \quad r = (0,1 \pm 0,005) \text{ см}; \quad l = (100 \pm 0,1) \text{ см}.$$

Составим формулу относительной погрешности плотности:

$$\delta \rho = \delta m + \delta r^2 + \delta l = \delta m + 2\delta r + \delta l = \frac{\Delta m}{m} + 2\frac{\Delta r}{r} + \frac{\Delta l}{l}.$$

Тогда абсолютную погрешность можно найти по формуле:

$$\Delta \rho = \delta \rho \cdot \rho = \left(\frac{\Delta m}{m} + 2\frac{\Delta r}{r} + \frac{\Delta l}{l} \right) \cdot \frac{m}{\pi r^2 l}.$$

Подставляем данные в полученную формулу и округляем значение погрешности до одной значащей цифры:

$$\Delta\rho = \left(\frac{0,1}{67,5} + 2 \frac{0,005}{0,1} + \frac{0,1}{100} \right) \cdot \frac{67,5}{\pi(0,1)^2 100} = 2,2019... \approx 2.$$

Находим точечную оценку плотности:

$$\rho_{\tau} = \frac{m}{\pi r^2 l} = \frac{67,5}{\pi(0,1)^2 100} = 21,4859... \approx 21.$$

Итак, плотность металла $\rho = (21 \pm 2) \text{ г/см}^3$. Т.е. $\rho \in [19; 23] \text{ г/см}^3$.

В более сложных случаях, когда правила не применимы, формулу абсолютной и относительной погрешности получают, используя навыки дифференцирования функций нескольких переменных.

Лекция 2. Нахождение формулы функциональной зависимости методом наименьших квадратов

2.1. Основные положения метода наименьших квадратов

При проведении исследований и измерений в различных областях деятельности человека очень часто зависимости между двумя величинами x и y получаются в табличном виде:

x	x_1	x_2	$\dots$	x_n
y	y_1	y_2	$\dots$	y_n

Чтобы получить зависимость y от x в аналитическом виде, т.е. в виде формулы, широко применяется *метод наименьших квадратов* (МНК).

Он основан на том, что из множества функций $y = f(x)$ наилучшей является та, для которой сумма S квадратов отклонений наблюдаемых значений от вычисленных является наименьшей; значение S вычисляется по формуле

$$S = \sum_{i=1}^n (y_i - f(x_i))^2.$$

Здесь n – число измерений, y_i – полученные в опыте табличные значения функции, $f(x_i)$ – значения функции, вычисленные по выбранной формуле.

Первым этапом применения МНК является выбор типа или класса функций $f(x)$. Часто используются

а) линейная функция $f(x) = ax + b$;

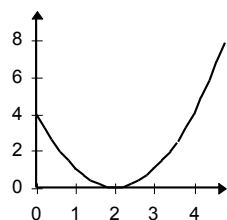
б) квадратичная функция $f(x) = ax^2 + bx + c$;

в) показательная или экспоненциальная функция $f(x) = be^{ax}$;

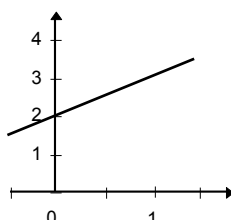
г) логарифмическая функция $f(x) = a \ln x + b$;

д) гиперболическая функция $f(x) = \frac{a}{x} + b$.

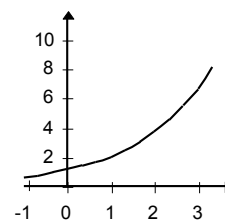
Чтобы выбрать вид функции, табличные значения наносят на график и соотносят расположение точек с формой графика одной из пяти предлагаемых функций:



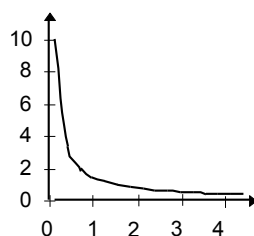
$$f(x) = ax^2 + bx + c$$



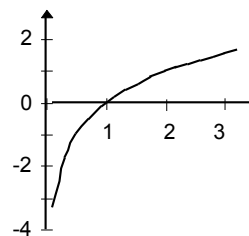
$$f(x) = ax + b$$



$$f(x) = be^{ax}$$



$$f(x) = \frac{a}{x} + b$$



$$f(x) = a \ln x + b$$

Для описания табличной зависимости выбирают ту функцию, к графику которой ближе всего расположены точки. Вид функции можно выбрать и более строго: наилучшим образом описывает (аппроксимирует) полученную в опыте зависимость та функция, для которой сумма квадратов отклонений S меньше, чем для других. Но этот способ достаточно трудоемкий.

Второй этап применения МНК – определение числовых параметров выбранной функции.

2.2. Определение параметров функциональных зависимостей

1) Определение параметров линейной зависимости $f(x) = ax + b$.

Согласно основному положению МНК параметры a и b должны быть

такими, чтобы величина $\sum_{i=1}^n (y_i - (ax_i + b))^2 = S(a, b)$ была минимальной. Со-

ставленная сумма является функцией двух переменных. Необходимым условием минимума такой функции $S(a, b)$ является равенство 0 её частных производных:

$$\begin{cases} S'_a = -2 \sum_{i=1}^n (y_i - (ax_i + b))x_i = 0 \\ S'_b = -2 \sum_{i=1}^n (y_i - (ax_i + b)) = 0 \end{cases}.$$

Преобразование написанных условий приводит к следующей системе для определения a и b :

$$\begin{cases} a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i x_i \\ a \sum_{i=1}^n x_i + b \cdot n = \sum_{i=1}^n y_i \end{cases}.$$

Решая эту систему, получаем формулы:

$$a = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}; \quad (1) \quad b = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}; \quad (2)$$

2) Определение параметров экспоненциальной зависимости $f(x) = be^{ax}$.

Прологарифмируем функцию f : $\ln f = ax + \ln b$

Зависимость $\ln f$ от x линейная, поэтому заменим в формулах (1) и (2) y_i на $\ln y_i$, а b на $\ln b$:

$$a = \frac{n \sum_{i=1}^n x_i \ln y_i - \sum_{i=1}^n \ln y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}; \quad (3); \quad \ln b = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n \ln y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i \ln y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad (4).$$

Окончательно находим $b = e^{\ln b}$.

3) Определение параметров логарифмической зависимости $f(x) = a \ln x + b$.

Отличие данной формулы от линейной состоит в замене x на $\ln x$, поэтому:

$$a = \frac{n \sum_{i=1}^n y_i \ln x_i - \sum_{i=1}^n y_i \sum_{i=1}^n \ln x_i}{n \sum_{i=1}^n (\ln x_i)^2 - \left(\sum_{i=1}^n \ln x_i \right)^2}; \quad (5); \quad b = \frac{\sum_{i=1}^n (\ln x_i)^2 \sum_{i=1}^n y_i - \sum_{i=1}^n \ln x_i \sum_{i=1}^n y_i \ln x_i}{n \sum_{i=1}^n (\ln x_i)^2 - \left(\sum_{i=1}^n \ln x_i \right)^2} \quad (6).$$

4) Определение параметров гиперболической зависимости $f(x) = \frac{a}{x} + b$.

В данной формуле, в отличие от линейной, x заменен на $\frac{1}{x}$, поэтому

$$a = \frac{n \sum_{i=1}^n \frac{y_i}{x_i} - \sum_{i=1}^n y_i \sum_{i=1}^n \frac{1}{x_i}}{n \sum_{i=1}^n \frac{1}{x_i^2} - \left(\sum_{i=1}^n \frac{1}{x_i} \right)^2} \quad (7); \quad b = \frac{\sum_{i=1}^n \frac{1}{x_i^2} \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{1}{x_i} \sum_{i=1}^n \frac{y_i}{x_i}}{n \sum_{i=1}^n \frac{1}{x_i^2} - \left(\sum_{i=1}^n \frac{1}{x_i} \right)^2} \quad (8).$$

5) Определение параметров квадратичной зависимости $f(x) = ax^2 + bx + c$.

$$\text{Составим } S = \sum_{i=1}^n \left(y_i - (ax_i^2 + bx_i + c) \right)^2 = S(a, b, c).$$

Необходимое условие минимума функции $S(a, b, c)$:

$$\begin{cases} S'_a = -2 \sum_{i=1}^n \left(y_i - (ax_i^2 + bx_i + c) \right) x_i^2 = 0 \\ S'_b = -2 \sum_{i=1}^n \left(y_i - (ax_i^2 + bx_i + c) \right) x_i = 0 \\ S'_c = -2 \sum_{i=1}^n \left(y_i - (ax_i^2 + bx_i + c) \right) = 0. \end{cases}$$

Преобразование выписанных уравнений приводит к системе

$$\begin{cases} a \sum_{i=1}^n x_i^4 + b \sum_{i=1}^n x_i^3 + c \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i^2 \\ a \sum_{i=1}^n x_i^3 + b \sum_{i=1}^n x_i^2 + c \sum_{i=1}^n x_i = \sum_{i=1}^n y_i x_i \\ a \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i + cn = \sum_{i=1}^n y_i. \end{cases} \quad (9)$$

Это система трех линейных уравнений для трех неизвестных a, b, c , её решение в общем виде выглядит громоздко, поэтому после её составления, решение системы находят одним из известных методов: Гаусса, матричным или по формулам Крамера.

Практическое занятие 1. Нахождение точечной и интервальной оценок измеряемой величины при прямых и косвенных измерениях

Задание 1. В таблице представлены результаты 10 измерений верхнего артериального давления $P_в$, нижнего артериального давления $P_н$ и частоты пульса $Ч_н$ взрослого человека.

№ изм.	1	2	3	4	5	6	7	8	9	10
$P_в$	127	131	124	135	136	139	132	134	122	121
$P_н$	80	80	75	77	80	80	79	77	77	74
$Ч_н$	78	74	73	74	73	77	73	74	70	72

Для каждой величины найти её точечную и интервальную оценки.

Задание 2. Для величины X найти точечную и интервальную оценки при следующих данных. При выполнении задания руководствоваться примером 2 из лекции 1.

1 вариант	$X = \frac{\sqrt{ab}}{\sqrt[3]{c}}$	$a = 38,50 \pm 0,01$
		$b = 2,0435 \pm 0,0004$
		$c = 962,6 \pm 0,1$
2 вариант	$X = \frac{\sqrt[3]{a \cdot b}}{27c}$	$a = 81,57 \pm 0,02$
		$b = 12,2 \pm 0,3$
		$c = 0,5 \pm 0,1$
3 вариант	$X = \frac{\sqrt{ab}}{c}$	$a = 3,845 \pm 0,004$
		$b = 16,20 \pm 0,05$
		$c = 10,8 \pm 0,1$
4 вариант	$X = \frac{a^2b}{\sqrt{c}}$	$a = 3,456 \pm 0,002$
		$b = 0,6420 \pm 0,0005$
		$c = 7,120 \pm 0,004$
5 вариант	$X = \frac{ab^3}{\sqrt{c}}$	$a = 0,6430 \pm 0,0005$
		$b = 2,170 \pm 0,002$
		$c = 5,843 \pm 0,001$
6 вариант	$X = \frac{\sqrt{a \cdot b}}{c^2}$	$a = 0,3575 \pm 0,0002$
		$b = 2,63 \pm 0,01$
		$c = 0,8540 \pm 0,0005$

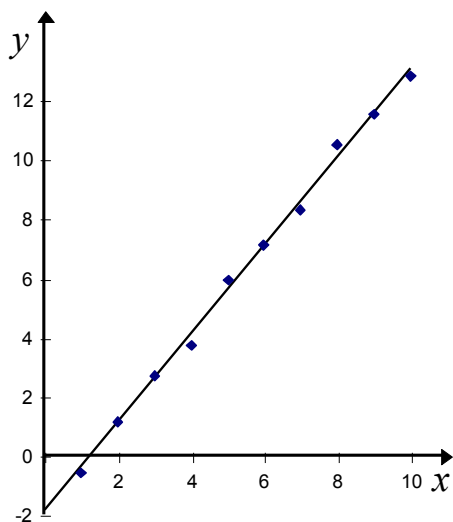
7 вариант	$X = \frac{a^2 b}{c^3}$	$a = 1,6531 \pm 0,0003$
		$b = 3,780 \pm 0,002$
		$c = 0,1580 \pm 0,0005$
8 вариант	$X = \sqrt{\frac{ab}{c}}$	$a = 0,7568 \pm 0,0002$
		$b = 21,70 \pm 0,02$
		$c = 2,65 \pm 0,1$
9 вариант	$X = \frac{2ab^2}{\sqrt{c}}$	$a = 0,051 \pm 0,001$
		$b = 2,14 \pm 0,03$
		$c = 25,66 \pm 0,06$
10 вариант	$X = \frac{ab^3}{12c}$	$a = 17,30 \pm 0,03$
		$b = 5,73 \pm 0,01$
		$c = 0,956 \pm 0,004$

Практическое занятие 2. Применение метода наименьших квадратов в случае линейной и квадратичной функций

Пример 1. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	-0,6	1,1	2,7	3,7	5,9	7,1	8,3	10,5	11,5	12,8

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.



Нанесем табличные данные на график.

Расположение точек ближе всего к графику линейной функции – прямой. Поэтому для представленной в таблице зависимости выбираем линейную функцию

$$f(x) = ax + b.$$

Вычислим суммы:

$$\sum_{i=1}^{10} x_i = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55;$$

$$\sum_{i=1}^{10} y_i = -0,6 + 1,1 + 2,7 + 3,7 + 5,9 + 7,1 + 8,3 + 10,5 + 11,5 + 12,8 = 63;$$

$$\sum_{i=1}^{10} x_i^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2 = 385;$$

$$\sum_{i=1}^{10} x_i y_i = -0,6 \cdot 1 + 1,1 \cdot 2 + 2,7 \cdot 3 + 3,7 \cdot 4 + 5,9 \cdot 5 + 7,1 \cdot 6 + 8,3 \cdot 7 +$$

$$+ 10,5 \cdot 8 + 11,5 \cdot 9 + 12,8 \cdot 10 = 470,2;$$

$$10 \sum_{i=1}^{10} x_i^2 - \left(\sum_{i=1}^{10} x_i \right)^2 = 10 \cdot 385 - 55^2 = 825.$$

Используя формулы (1) и (2) из лекции 2, найдем

$$a = \frac{10 \cdot 470,2 - 63 \cdot 55}{825} \approx 1,50; \quad b = \frac{385 \cdot 63 - 55 \cdot 470,2}{825} \approx -1,95;$$

$$f(x) = 1,5x - 1,95.$$

Нанесем прямую, описываемую данной функцией, на график по двум точкам: $x=0$; $y=-1,95$ и $x=10$; $y=13,05$.

Табличные точки располагаются вблизи полученной прямой, следовательно, вид функции выбран удачно.

Пример 2. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	4,2	2,8	5,7	10,5	13,2	20,5	33,4	46,9	60,1	71,2

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Нанесём табличные данные на координатную плоскость (результат в тексте не приведён, нанесение следует выполнить самостоятельно). Расположение точек указывает на квадратичную зависимость между исследуемыми величинами, которая описывается следующей формулой: $y = ax^2 + bx + c$.

Параметры выбранной зависимости можно найти из системы (9), представленной в лекции 2. Для составления системы вычислим суммы:

$$\sum_{i=1}^{10} y_i = 4,2 + 2,8 + \dots + 71,2 = 268,5;$$

$$\sum_{i=1}^{10} y_i x_i = 4,2 \cdot 1 + 2,8 \cdot 2 + \dots + 71,2 \cdot 10 = 2119,8;$$

$$\sum_{i=1}^{10} y_i x_i^2 = 4,2 \cdot 1^2 + 2,8 \cdot 2^2 + \dots + 71,2 \cdot 10^2 = 17929;$$

$$\sum_{i=1}^{i=10} x_i^3 = 1^3 + 2^3 + \dots + 10^3 = 3025;$$

$$\sum_{i=1}^{i=10} x_i^4 = 1^4 + 2^4 + \dots + 10^4 = 25333.$$

Подставим все необходимые суммы в систему (9):

$$\begin{cases} 25333a + 3025b + 385c = 17929; \\ 3025a + 385b + 55c = 2119,8; \\ 385a + 55b + 10c = 268,5. \end{cases}$$

Решаем систему методом Гаусса (методом исключения неизвестных). Вначале исключим параметр c из первых двух уравнений. Для этого к первому уравнению прибавим третье, умноженное на $(-38,5)$, ко второму уравнению прибавим третье, умноженное на $(-5,5)$. После выполнения перечисленных действий система примет вид:

$$\begin{cases} 10510,5a + 907,5b = 7591,75; \\ 907,5a + 82,5b = 643,05; \\ 385a + 55b + 10c = 268,5. \end{cases}$$

Разделим обе части второго уравнения на $82,5$. Система примет вид.

$$\begin{cases} 10510,5a + 907,5b = 7591,75; \\ 11a + b = 7,79(45); \\ 385a + 55b + 10c = 268,5. \end{cases}$$

К первому уравнению прибавим второе, умноженное на $(-907,5)$.

$$\begin{cases} 528a = 518,2; \\ 11a + b = 7,79(45); \\ 385a + 55b + 10c = 268,5. \end{cases}$$

Из первого уравнения находим $a = \frac{518,2}{528} = 0,981439394... \approx 0,98$. Подста-

вим найденное значение во второе уравнение, найдём

$$b = 7,79(45) - 11a = -3,001287879... \approx -3,00.$$

$$c = \frac{268,5 - 55 \cdot (-3) - 385 \cdot 0,98}{10} = 5,62.$$

Выпишем формулу функциональной зависимости

$$f(x) = 0,98x^2 - 3x + 5,62.$$

Для проверки составим таблицу значений функции и нарисуем её график вместе с табличными данными. При вычислении значения функции округлим до десятых, как и исходные данные.

x	1	2	3	4	5	6	7	8	9	10
$f(x)$	3,6	3,5	5,4	9,3	15,1	22,9	32,6	44,3	58,0	73,6

Построенный график весьма удовлетворительно соответствует табличным данным, поэтому выбор функции можно считать удачным.

Задание 1 для самостоятельной работы.

Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i		1	2	3	4	5	6	7	8	8	10
y_i	1 вар.	2,5	1,3	- 1,5	- 6,4	-9,7	-14,5	-11,2	-18,1	-17,6	-26,4
	2 вар.	10	26	35	51	60	79	85	105	110	130
	3 вар.	-2	-12	-15	-19	-35	-35	-47	-55	-60	-69
	4 вар.	-6	-6	0	6	6	13	15	21	23	30
	5 вар.	-3	-1	0	2,5	7,5	7,5	10	14	15	18,5
	6 вар.	8	1	-4	-7	-15	-15	-23	-33	-28	-38
	7 вар.	-5	3	15	19	34	34	40	41	62	70
	8 вар.	11	14	5	8	- 1	2	- 7	- 4	- 13	- 10
	9 вар.	4	7	18	21	32	35	46	49	60	63
	10 вар.	5	3	- 3	- 5	- 11	- 13	- 20	- 20	- 24	- 32

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Задание 2 для самостоятельной работы.

Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i		1	2	3	4	5	6	7	8	8	10
y_i	1 вар.	15	18	21	17	17	10	6	- 2	- 17	- 33
	2 вар.	0	- 6	- 20	- 25	- 40	- 62	- 77	- 100	-110	-150
	3 вар.	- 6	- 8	- 5	0	3	16	20	36	50	64
	4 вар.	0	- 1	6	15	35	42	65	100	120	175
	5 вар.	1	6	6	4	2	- 7	- 7	- 20	- 28	- 45
	6 вар.	- 13	- 16	3	- 6	40	65	110	151	217	290
	7 вар.	1	0	10	10	31	41	67	88	125	150
	8 вар.	3	0	- 2	- 15	- 25	- 40	- 40	- 66	- 80	-110
	9 вар.	- 6	6	25	38	65	80	110	125	170	185
	10 вар.	6	8	- 2	- 8	- 27	- 39	- 67	- 87	-125	-149

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Практическое занятие 3. Применение метода наименьших квадратов в случае показательной, логарифмической и гиперболической зависимостей.

Пример 1. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	3	6	9	14	26	38	70	105	180	290

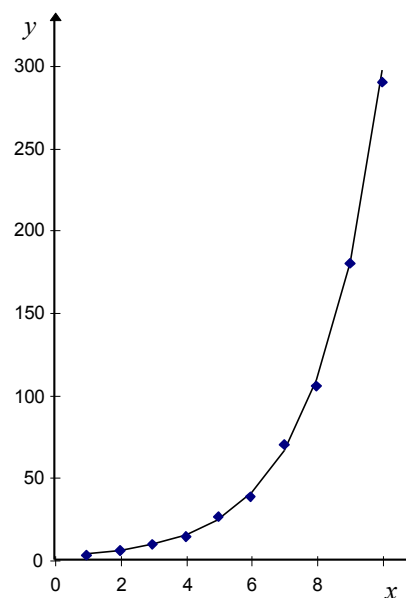
Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Нанесем табличные данные на график. Расположение точек позволяет сделать вывод: функция при малых значениях переменной ($x < 5$) меняется слабо, а при больших значениях – резко возрастает, что свидетельствует в пользу выбора экспоненциальной зависимости $f(x) = be^{ax}$.

Проведем необходимые вычисления.

$$\sum_{i=1}^{10} x_i = 55; \sum_{i=1}^{10} x_i^2 = 385$$

$$10 \sum_{i=1}^{10} x_i^2 - \left(\sum_{i=1}^{10} x_i \right)^2 = 825.$$



$$\sum_{i=1}^{10} \ln y_i = \ln y_1 y_2 \dots y_{10} = \ln(3 \cdot 6 \cdot 9 \cdot 14 \cdot 26 \cdot 38 \cdot 70 \cdot 105 \cdot 180 \cdot 290) \approx 34,388;$$

$$\sum_{i=1}^{10} x_i \ln y_i = 1 \cdot \ln 3 + 2 \ln 6 + 3 \ln 9 + 4 \ln 14 +$$

$$+ 5 \ln 26 + 6 \ln 38 + 7 \ln 70 + 8 \ln 105 + 9 \ln 180 + 10 \ln 290 \approx 230,353;$$

$$a = \frac{10 \cdot 230,35 - 34,39 \cdot 55}{825} \approx 0,4994 \approx 0,50; \quad \ln b = \frac{385 \cdot 34,39 - 55 \cdot 230,35}{825} = 0,6920;$$

$$b = e^{\ln b} = 1,997 \approx 2,00; \quad f(x) = 2e^{0,5x}.$$

Заполним таблицу значениями функции.

x	1	2	3	4	5	6	7	8	9	10
$y = 2e^{0,5x}$	3,3	5,4	9,0	14,8	24,4	40,2	66,2	109,2	180,0	296,8

Нанесем полученные точки на график и соединим плавной линией, которая весьма удовлетворительно соответствует исходным данным.

Пример 2. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	4,2	3,2	2,9	2,5	2,45	2,15	2	1,75	1,9	1,6

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Отметим табличные значения на координатной плоскости (нанесение точек выполнить самостоятельно). Плавное изменение значений функции свидетельствует в пользу выбора логарифмической зависимости $f(x) = a \ln x + b$.

Числовые параметры выбранной зависимости найдём по формулам (7) и (8), представленным в лекции 2. Чтобы воспользоваться формулами, вычислим необходимые для этого суммы.

$$\sum_{i=1}^{i=10} \ln x_i = \ln x_1 \cdot x_2 \cdot x_3 \dots x_{10} = \ln 1 \cdot 2 \cdot 3 \dots 10 \approx 15,104;$$

$$\sum_{i=1}^{i=10} (\ln x_i)^2 = (\ln 1)^2 + (\ln 2)^2 + (\ln 3)^2 + \dots + (\ln 10)^2 \approx 27,650;$$

$$\sum_{i=1}^{i=10} y_i = 4,2 + 3,2 + 2,9 + \dots + 1,9 + 1,6 = 24,65;$$

$$\sum_{i=1}^{i=10} y_i \ln x_i = 4,2 \ln 1 + 3,2 \ln 2 + 2,9 \ln 3 + \dots + 1,9 \ln 9 + 1,6 \ln 10 \approx 32,055;$$

Вычислим отдельно знаменатель в формулах (7) и (8), обозначив его Δ .

$$\Delta = 10 \cdot 27,650 - (15,104)^2 = 48,369184.$$

$$a = \frac{10 \cdot 32,055 - 24,65 \cdot 15,104}{48,369184} \approx -1,07; \quad b = \frac{27,65 \cdot 24,65 - 15,104 \cdot 32,055}{48,369184} \approx 4,08.$$

Формула искомой зависимости будет иметь вид: $f(x) = -1,07 \ln x + 4,08$.

Для проверки результата составим таблицу значений (округлённых так же, как и обрабатываемые данные, до сотых) найденной аналитической зависимости и нарисуем её график вместе с исходными данными.

x	1	2	3	4	5	6	7	8	9	10
$f(x)$	4,08	3,34	2,90	2,60	2,36	2,16	2,00	1,85	1,73	1,62

Построенный график очень хорошо соответствует экспериментальным данным, значит, выбор аналитической зависимости можно считать верным.

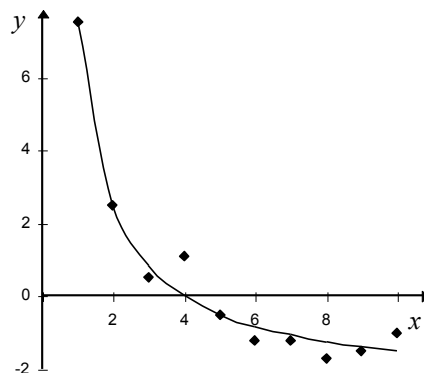
Пример 3. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	7,5	2,5	0,5	1,1	- 0,5	- 1,2	- 1,2	- 1,7	- 1,5	- 1

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Нанесем табличные данные на график. Сильное изменение функции при небольших $x(x < 5)$ и очень незначительное при больших $x(5 < x \leq 10)$ указывает на возможность выбрать для данного случая гиперболическую зависимость

$$f(x) = \frac{a}{x} + b.$$



Выполним необходимые вычисления:

$$\sum_{i=1}^{10} \frac{1}{x_i} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} \approx 2,929;$$

$$\sum_{i=1}^{10} \frac{1}{x_i^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2} + \frac{1}{8^2} + \frac{1}{9^2} + \frac{1}{10^2} \approx 1,550;$$

$$10 \sum_{i=1}^{10} \frac{1}{x_i^2} - \left(\sum_{i=1}^{10} \frac{1}{x_i} \right)^2 \approx 6,920959;$$

$$\sum_{i=1}^{10} y_i = 7,5 + 2,5 + 0,5 + 1,1 - 0,5 - 1,2 - 1,2 - 1,7 - 1,5 - 1 = 4,5;$$

$$\sum_{i=1}^{10} \frac{y_i}{x_i} = \frac{7,5}{1} + \frac{2,5}{2} + \frac{0,5}{3} + \frac{1,1}{4} - \frac{0,5}{5} - \frac{1,2}{6} - \frac{1,2}{7} - \frac{1,7}{8} - \frac{1,5}{9} - \frac{1}{10} \approx 8,241;$$

$$a = \frac{10 \cdot 8,241 - 4,5 \cdot 2,929}{6,920959} \approx 10,00;$$

$$b = \frac{1,550 \cdot 4,5 - 2,929 \cdot 8,241}{6,920959} \approx -2,48;$$

$$f(x) = \frac{10}{x} - 2,48.$$

Составим таблицу значений функции, округлив их до десятых, как и в исходной таблице.

x	1	2	3	4	5	6	7	8	9	10
$f(x) = \frac{10}{x} - 2,48$	7,5	2,5	0,9	0,0	-0,5	-0,8	-1,1	-1,2	-1,4	-1,5

Нанесем рассчитанные значения на график и соединим плавной линией. Точки, соответствующие исходным данным, хорошо группируются вблизи проведенной линии, значит, зависимость выбрана удовлетворительно.

Задание 1 для самостоятельной работы.

Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i		1	2	3	4	5	6	7	8	8	10
y_i	1 вар.	4,4	5,1	5,4	6,7	6,2	7,5	7,7	9,2	9,9	11,5
	2 вар.	15	13	25	30	45	55	82	120	140	200
	3 вар.	7	8	15	20	32	42	60	100	120	220
	4 вар.	0,7	0,7	1,2	2	4,2	3,4	7,1	6,9	15	21,9
	5 вар.	0,6	0,1	3	7,5	7,5	21	38,3	80,1	171	340
	6 вар.	0,3	0,51	0,7	1,1	1,3	1,7	2,6	3,1	4,5	6,3
	7 вар.	10	11	25	32	65	91	170	253	460	742
	8 вар.	0,4	0,2	1,5	1,5	9,2	4,6	21,6	56,1	101,3	238,5
	9 вар.	1,5	0,9	4,5	2,6	10,3	10,3	27,5	35,7	73,6	120
	10 вар.	12,2	6,1	3	49,7	24,7	110	191	423	801	1600

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Задание 2 для самостоятельной работы.

Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i		1	2	3	4	5	6	7	8	8	10
y_i	1 вар.	4,9	6,5	7,1	7,9	8,1	8,9	8,6	9,1	9,5	9,7
	2 вар.	-10,5	- 8,1	- 4,4	- 2,1	- 2,1	- 1,1	0,5	0,3	1,5	1,8
	3 вар.	6,8	5,8	5	4,3	3,6	3,6	3,1	2,9	2,5	2,3
	4 вар.	- 7,5	- 5,2	- 3,1	- 2,9	- 1,5	- 1	0	0,5	0,5	1,3
	5 вар.	1,5	- 0,1	- 1,1	- 2,1	- 2,2	- 2,6	- 3,2	- 2,9	- 3,5	- 3,5
	6 вар.	2	7,1	13,5	18,1	15,9	21,5	19,9	24	22,8	25
	7 вар.	2	- 5,3	-13,6	-13,6	-17,9	- 21	-22,4	-25,4	- 24	-26,6
	8 вар.	2,5	9	16,1	16,1	22,5	20,3	24,4	27,4	26	28,6
	9 вар.	13,1	5,1	- 2,1	- 5,5	-10,1	-10,1	-16,2	-14,2	-19,5	- 18
	10 вар.	- 5,2	3,4	12,7	15,6	21,4	22,6	26,1	30,4	28,4	32,1

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Задание 3 для самостоятельной работы.

Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i		1	2	3	4	5	6	7	8	8	10
y_i	1 вар.	- 6,5	- 1,2	0,1	- 0,1	0,3	0,9	1,1	1,1	1,3	1,2
	2 вар.	12	11	10,7	10,5	10,4	10,3	10,3	10,3	10,1	10
	3 вар.	- 3	- 0,5	0,3	0,7	1	1,3	1,2	1,4	1,4	1,6
	4 вар.	8,5	3,9	2,1	1,9	1	0,5	0,5	0,1	0,3	0
	5 вар.	26	18	14	11	11	10,5	9,5	9,9	9	9,2
	6 вар.	-14,5	-12,5	-11,3	-11,7	-10,8	- 11	-10,7	-10,5	-10,5	-10,2
	7 вар.	6,5	1,5	- 0,5	0,1	- 1,5	- 2,2	- 2,2	- 2,7	- 2,5	- 2,8
	8 вар.	18,1	10,6	6,2	3,7	3,7	2	2,5	1,3	1,6	1,1
	9 вар.	- 15	- 7,5	- 2,1	- 1,3	0,1	0,6	0,9	1,6	1,4	1,9
	10 вар.	20	8	0,1	- 3	- 3,3	- 4,5	- 5,2	- 6,2	- 6,8	- 6,4

Методом наименьших квадратов найти формулу функциональной зависимости $y = f(x)$.

Практическое занятие 4. Установление тесноты связи между величинами в однофакторном эксперименте

Как уже говорилось, при экспериментальном исследовании зависимостей они искажаются действием случайных факторов. Поэтому интересно выяснить не только характер зависимости между функцией (откликом) и независимой переменной (фактором), но и определить, насколько значения отклика определяются действием фактора, т.е. установить степень тесноты связи между величинами.

Тесноту связи можно оценить вычислением коэффициента корреляции в случае линейной зависимости между величинами или индекса корреляции в случае нелинейной зависимости.

Коэффициент корреляции обозначается r_{xy} и вычисляется по формуле:

$$r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - (\bar{x})^2} \cdot \sqrt{\overline{y^2} - (\bar{y})^2}}.$$

В этой формуле используются различные средние значения:

$$\bar{x} = \frac{\sum_{i=1}^{i=n} x_i}{n}; \quad \bar{y} = \frac{\sum_{i=1}^{i=n} y_i}{n}; \quad \overline{x^2} = \frac{\sum_{i=1}^{i=n} x_i^2}{n}; \quad \overline{y^2} = \frac{\sum_{i=1}^{i=n} y_i^2}{n}; \quad \overline{x \cdot y} = \frac{\sum_{i=1}^{i=n} x_i y_i}{n};$$

n - количество зарегистрированных в эксперименте пар значений фактора и отклика.

Коэффициент корреляции должен удовлетворять двойному неравенству:
 $-1 \leq r_{xy} \leq 1$.

Существует шкала Чеддока, в которой устанавливается соответствие между значениями коэффициента корреляции и степенью тесноты связи между величинами.

r_{xy}	$ 0,1 - 0,3 $	$ 0,3 - 0,5 $	$ 0,5 - 0,7 $	$ 0,7 - 0,9 $	$ 0,9 - 0,99 $
Сила связи	слабая	умеренная	заметная	высокая	очень высокая

При положительных значениях коэффициента корреляции связь является положительной (наблюдается прямая зависимость между фактором и откликом); при отрицательных значениях коэффициента корреляции связь является отрицательной (наблюдается обратная зависимость между величинами).

Индекс корреляции, обозначаемый ρ_{xy} , можно вычислить по формуле:

$$\rho_{xy} = \sqrt{1 - \frac{\sum_{i=1}^{i=n} (y_i - y(x_i))^2}{\sum_{i=1}^{i=n} (y_i - \bar{y})^2}}.$$

В этой формуле y_i - зарегистрированные в эксперименте значения функции (отклика);

$y(x_i)$ - значения отклика, рассчитанные по формуле функциональной зависимости, найденной методом наименьших квадратов;

$\bar{y}$ - среднее арифметическое экспериментальных значений отклика;

n - количество зарегистрированных в эксперименте пар значений фактора и отклика.

Индекс корреляции может принимать значения от 0 до 1.

Существует и другой вариант вычисления индекса корреляции, однако он очень чувствителен к точности нахождения $y(x_i)$ и при недостаточной точности может дать неправдоподобные результаты.

Пример 1. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	-0,6	1,1	2,7	3,7	5,9	7,1	8,3	10,5	11,5	12,8

Методом наименьших квадратов найдена формула линейной зависимости между величинами $y = 1,5x - 1,95$.

Оценить тесноту связи между величинами, вычислив коэффициент корреляции.

Найдём необходимые средние величины, используя вычисления, выполненные в примере 1 практического занятия 2.

$$\bar{x} = \frac{55}{10} = 5,5; \quad \bar{y} = \frac{63}{10} = 6,3; \quad \overline{x^2} = \frac{385}{10} = 38,5; \quad \overline{y^2} = \frac{583}{10} = 58,3; \quad \overline{x \cdot y} = \frac{470,2}{10} = 47,02.$$

Подставим найденные средние величины в формулу для вычисления коэффициента корреляции:

$$r_{xy} = \frac{47,02 - 5,5 \cdot 6,3}{\sqrt{38,5 - (5,5)^2} \cdot \sqrt{58,3 - (6,3)^2}} \approx 0,998.$$

Значение коэффициента корреляции указывает на очень высокую положительную силу связи между величинами.

Пример 2. Результаты экспериментального исследования величины y от величины x представлены в таблице.

x_i	1	2	3	4	5	6	7	8	9	10
y_i	7,5	2,5	0,5	1,1	- 0,5	- 1,2	- 1,2	- 1,7	- 1,5	- 1

С помощью метода наименьших квадратов установлено, что зависимость гиперболическая и найдена её формула $y = \frac{10}{x} - 2,48$.

Оценить тесноту связи между величинами, вычислив индекс корреляции.

Вычисленные по формуле значения функции возьмём из примера 3 практического занятия 3.

x	1	2	3	4	5	6	7	8	9	10
$y(x_i)$	7,5	2,5	0,9	0,0	-0,5	-0,8	-1,1	-1,2	-1,4	-1,5

Среднее экспериментальное значение функции (отклика) также найдём, используя этот пример: $\bar{y} = \frac{4,5}{10} = 0,45$.

Составим необходимые суммы.

$$\sum_{i=1}^{i=10} (y_i - y(x_i))^2 = (7,5 - 7,5)^2 + (2,5 - 2,5)^2 + (0,5 - 0,9)^2 + \dots + (-1,5 - (-1,4))^2 + (-1 - (-1,5))^2 = 2,05.$$

$$\sum_{i=1}^{i=10} (y_i - \bar{y})^2 = (7,5 - 0,45)^2 + (2,5 - 0,45)^2 + (0,5 - 0,45)^2 + \dots + (-1,5 - 0,45)^2 + (-1 - 0,45)^2 = 71,205.$$

Подставим найденные суммы в формулу для вычисления индекса корреляции: $\rho_{xy} = \sqrt{1 - \frac{2,05}{71,205}} \approx 0,985$. Значение индекса указывает на очень высокую тесноту связи между величинами.

Задание 1 для самостоятельной работы: используя задание 1 для самостоятельной работы практического занятия 2, оценить тесноту связи между величинами с помощью вычисления коэффициента корреляции.

Задание 2 для самостоятельной работы: используя задание 3 для самостоятельной работы практического занятия 3, оценить тесноту связи между величинами с помощью вычисления индекса корреляции.

Лабораторная работа №1. Нахождение точечной и интервальной оценок измеряемой величины при косвенных измерениях

Цель работы: провести математическую обработку, необходимую при выполнении эксперимента простейшего вида, состоящего в косвенном измерении значения какой-либо величины. В результате обработки для измеряемой величины должны быть указаны приближённое значение и абсолютная погрешность, т.е. даны её точечная и интервальная оценки.

Содержание работы: вывод формулы для нахождения величины, исходя из описания эксперимента; вывод формулы абсолютной погрешности согласно определению, данному в лекции 1, на основе формулы полного дифференциала функции нескольких переменных; вычисление абсолютной погрешности с округлением результата до одной значащей цифры; вычисление приближённого значения с округлением его до того же десятичного знака, что и погрешность; указание значения измеряемой величины.

Описание работы.

В качестве примера соответствующего эксперимента выбрано определение плотности жидкости. Плотность жидкости находят, сравнивая её с плотностью воды при температуре измерения.

Эксперимент проводят следующим образом. Берут стеклянный сосуд, для которого очень точно можно зафиксировать объём наливаемой в него жидкости. Такой сосуд называется пикнометром. Вначале находят массу пикнометра. Затем наливают в него воду и находят массу пикнометра вместе с водой. Затем воду выливают, пикнометр сушат и заливают в него жидкость, плотность которой нужно найти. Находят массу пикнометра вместе с жидкостью. При такой процедуре плотность жидкости ρ можно найти по формуле:

$$\rho = \frac{M_{\text{ж}} - m}{M_{\text{в}} - m} \rho_{\text{в}}.$$

В данной формуле $M_{\text{ж}}$ - масса жидкости вместе с пикнометром; $M_{\text{в}}$ - масса воды вместе с пикнометром; m - масса сухого пикнометра; $\rho_{\text{в}}$ - плотность воды при температуре измерения.

Чтобы найти приближённое значение (точечную оценку) плотности, в данную формулу нужно подставить взятое из справочника значение плотности воды при температуре проведения измерения и приближённые значения всех масс. Это значения, которые зафиксированы при взвешивании на рычажных весах.

Составим формулу абсолютной погрешности $\Delta\rho$. Плотность ρ является функцией трёх величин: $M_{\text{ж}}$, $M_{\text{с}}$, m . Плотность воды $\rho_{\text{с}}$ считается константой. Тогда полный дифференциал функции ρ будет иметь вид:

$$\begin{aligned} d\rho &= \rho'_{M_{\text{ж}}} dM_{\text{ж}} + \rho'_{M_{\text{с}}} dM_{\text{с}} + \rho'_m dm = \\ &= \frac{\rho_{\text{с}}}{M_{\text{с}} - m} dM_{\text{ж}} - \frac{\rho_{\text{с}}(M_{\text{ж}} - m)}{(M_{\text{с}} - m)^2} dM_{\text{с}} + \frac{-(M_{\text{с}} - m) + M_{\text{ж}} - m}{(M_{\text{с}} - m)^2} \rho_{\text{с}} dm = \\ &= \frac{\rho_{\text{с}}}{M_{\text{с}} - m} dM_{\text{ж}} - \frac{\rho_{\text{с}}(M_{\text{ж}} - m)}{(M_{\text{с}} - m)^2} dM_{\text{с}} + \frac{M_{\text{ж}} - M_{\text{с}}}{(M_{\text{с}} - m)^2} \rho_{\text{с}} dm = \\ &= \rho_{\text{с}} \left(\frac{dM_{\text{ж}}}{M_{\text{с}} - m} + \frac{m - M_{\text{ж}}}{(M_{\text{с}} - m)^2} dM_{\text{с}} + \frac{M_{\text{ж}} - M_{\text{с}}}{(M_{\text{с}} - m)^2} dm \right). \end{aligned}$$

Чтобы из формулы полного дифференциала получить формулу абсолютной погрешности, знаки дифференциалов заменим на знаки погрешностей, а все частные производные возьмём по модулю.

$$\Delta\rho = \rho_{\text{с}} \left(\frac{\Delta M_{\text{ж}}}{M_{\text{с}} - m} + \frac{|m - M_{\text{ж}}|}{(M_{\text{с}} - m)^2} \Delta M_{\text{с}} + \frac{|M_{\text{ж}} - M_{\text{с}}|}{(M_{\text{с}} - m)^2} \Delta m \right).$$

Задание. Дать точечную и интервальную оценку плотности жидкости (в г/см³) при следующих данных:

$M_{\text{ж}} = 235,12$ г; $M_{\text{с}} = 151,24$ г; $m = 51,10$ г; $\rho_{\text{с}} = 0,998$ г/см³; цена деления рычажных весов, на которых проводилось взвешивание – 0,02 г.

В отчёте указать результаты вычислений абсолютной погрешности $\Delta\rho$, приближённого значения плотности $\rho_{\text{нр}}$ и ответ в виде: $\rho = \rho_{\text{нр}} \pm \Delta\rho$.

Лабораторная работа № 2. Обработка результатов однофакторного пассивного эксперимента в случае линейной зависимости

Цель работы: по результатам экспериментальных наблюдений за фактором и откликом составить уравнение зависимости фактора от отклика (уравнение регрессии) и оценить тесноту связи между исследуемыми величинами.

Содержание работы: предварительная обработка данных, графическое изображение данных для получения вывода о характере зависимости, вывод уравнения зависимости, оценка тесноты связи между величинами.

Описание работы. В таблице приведены результаты обследования группы предприятий относительно величины X – стоимости основных фондов предприятия в млн. рублей, и Y – выработки продукции на одного рабочего в тыс. рублей.

$X \backslash Y$	9,9	10,0	10,1	10,2	10,3	10,4	10,5
0,8	1	3					
0,9	1	3	4	1			
1,0		1	4	2	1		
1,1				1	1	3	1
1,2		1			2	1	2
1,3						1	1

В клетках таблицы указаны частоты N_{ij} соответствующих пар значений x_i и y_j . Представленная таблица называется корреляционной. Исследуем зависимость выработки продукции на одного рабочего от стоимости основных фондов предприятия.

Для прояснения характера зависимости найдём средние значения величины Y для каждого значения величины X :

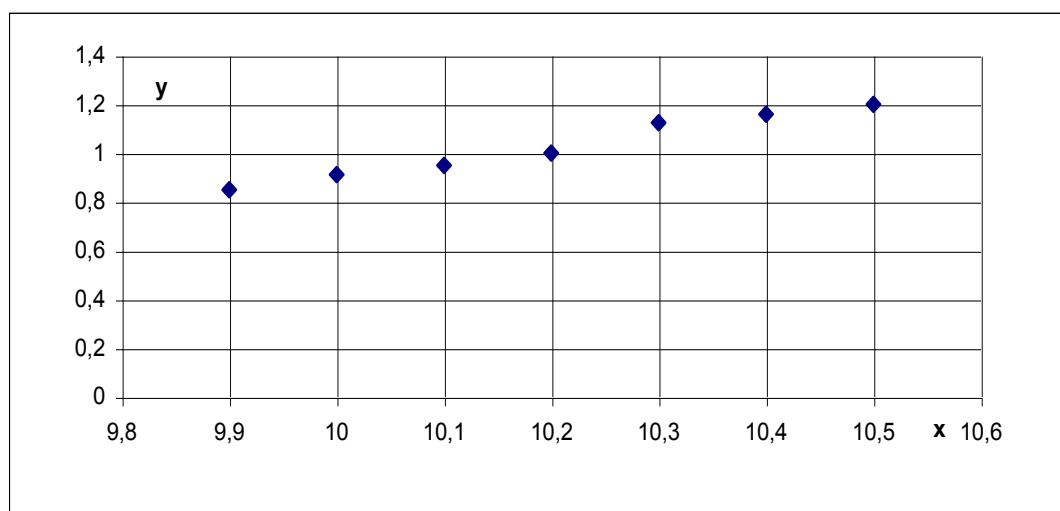
$$\bar{y}(X = 9,9) = \frac{0,8 + 0,9}{2} = 0,85; \quad \bar{y}(X = 10,0) = \frac{0,8 \cdot 3 + 0,9 \cdot 3 + 1,0 + 1,2}{8} = 0,9125;$$

$$\bar{y}(X = 10,1) = \frac{0,9 \cdot 4 + 1,0 \cdot 4}{8} = 0,95; \quad \bar{y}(X = 10,2) = \frac{0,9 + 1,0 \cdot 2 + 1,1}{2} = 1;$$

$$\bar{y}(X = 10,3) = \frac{1,0 + 1,1 + 1,2 \cdot 2}{4} = 1,125;$$

$$\bar{y}(X = 10,4) = \frac{1,1 \cdot 3 + 1,2 + 1,3}{5} = 1,16; \quad \bar{y}(X = 10,5) = \frac{1,1 + 1,2 \cdot 2 + 1,3}{4} = 1,2.$$

Изобразим точками на координатной плоскости соответствующие пары значений X и соответствующих им средних значений Y . Назовём такую картинку полем рассеяния.



Из положения точек на плоскости можно сделать вывод, что в исследуемом диапазоне изменения величин зависимость между ними линейная.

Уравнение линейной зависимости значений y величины Y от значений x величины X можно найти из следующей формулы:

$$y - \bar{y} = \frac{N \sum_{i,j=1}^{i=n, j=m} x_i y_j N_{ij} - \sum_{i=1}^{i=n} x_i l_i \cdot \sum_{j=1}^{j=m} y_j k_j}{N \sum_{i=1}^{i=n} x_i^2 l_i - \left(\sum_{i=1}^{i=n} x_i l_i \right)^2} (x - \bar{x}) \quad (1).$$

В формуле использованы следующие обозначения:

$N = \sum N_{ij}$; n - количество отдельных значений величины X ; m - количество отдельных значений величины Y ; l_i - частоты отдельных значений величины X ; k_j - частоты отдельных значений величины Y ; $\bar{x}$ - среднее значение

величины X , вычисляется по формуле: $\bar{x} = \frac{\sum_{i=1}^{i=n} x_i l_i}{\sum_{i=1}^{i=n} l_i}$; $\bar{y}$ - среднее значение

величины Y , вычисляется по формуле: $\bar{y} = \frac{\sum_{j=1}^{j=m} y_j k_j}{\sum_{j=1}^{j=m} k_j}$. Из корреляционной

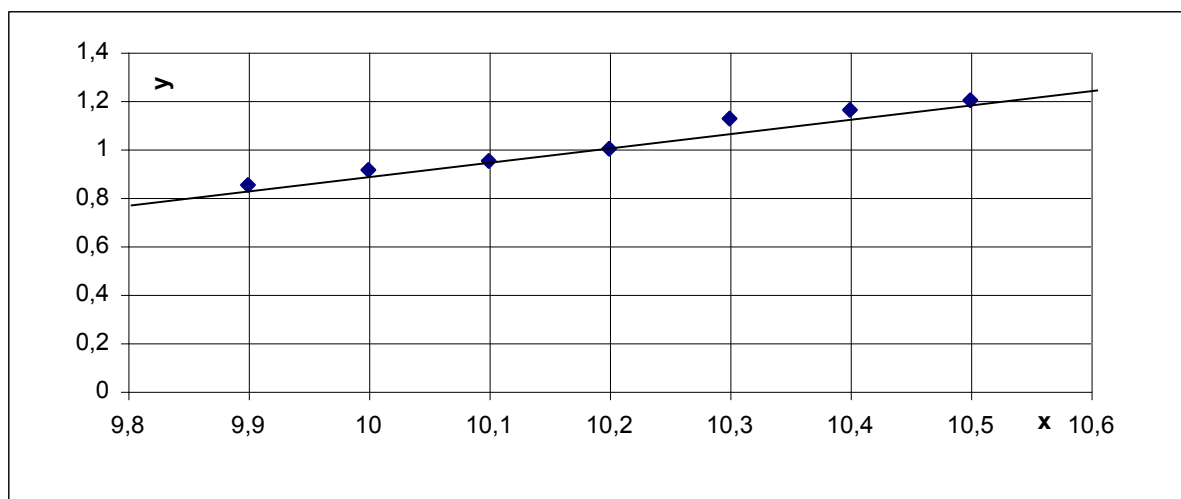
таблицы найдем частоты отдельных значений X и Y , результаты выпишем в отдельную таблицу:

x_i	9,9	10,0	10,1	10,2	10,3	10,4	10,5	y_j	0,8	0,9	1,0	1,1	1,2	1,3
l_i	2	8	8	4	4	5	4	k_j	4	9	8	6	6	2

Проведя необходимые вычисления, придём к уравнению зависимости:

$$y = 0,62x - 5,29.$$

Для проверки нарисуете график полученной прямой вместе с табличными данными. Для построения графика используем точки: (9,8; 0,786) и (10,6; 1,282).



Как видим, полученная формула соответствует имеющимся данным.

Оценим теперь тесноту связи между исследуемыми величинами. В случае линейной зависимости это можно сделать, вычислив коэффициент корреляции r_{xy} . Обычно его вычисляют по формуле: $r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sqrt{x^2 - (\bar{x})^2} \cdot \sqrt{y^2 - (\bar{y})^2}}$, где чертой обозначены соответствующие средние значения.

При вычислении средних значений необходимо выполнить деление, что приводит к получению бесконечных десятичных дробей и необходимости их округления. Вследствие использования округлённых значений могут произойти катастрофические неточности при вычислении коэффициента корреляции, поэтому приведённую формулу для вычисления коэффициента корреляции преобразуем к следующему виду:

$$r_{xy} = \frac{N \sum_{i,j=1}^{i=n, j=m} x_i y_j N_{ij} - \sum_{i=1}^{i=n} x_i l_i \cdot \sum_{j=1}^{j=m} y_j k_j}{\sqrt{N \sum_{i=1}^{i=n} x_i^2 l_i - \left(\sum_{i=1}^{i=n} x_i l_i \right)^2} \cdot \sqrt{N \sum_{j=1}^{j=m} y_j^2 k_j - \left(\sum_{j=1}^{j=m} y_j k_j \right)^2}} \quad (2)$$

Используя уже найденные значения сумм и вычислив недостающие, получим $r_{xy} \approx 0,79$.

Задание.

В таблице представлены результаты обследования вагонных колёс: X - время пробега колеса в месяцах, Y - износ обода колеса в мм.

$Y \backslash X$	0	2	7	12	17	22	27	32	37	42
0	3	6								
1	25	108	44	8	2					
2	3	50	60	21	5	5				
3	1	11	33	32	13	2	3	1		
4		5	5	13	13	7	2			
5			1	2	12	6	3	2	1	
6		1		1			2	1		1
7			1	1				1		

Составить формулу зависимости Y от X , оценить тесноту связи. В отчёте представить поле рассеяния, значения всех сумм, необходимых для вычисления формулы зависимости и коэффициента корреляции, формулу полученной зависимости и её график (на поле рассеяния) и значение коэффициента корреляции.

Лабораторная работа № 3. Обработка результатов однофакторного пассивного эксперимента в случае нелинейной зависимости

Цель работы: по результатам экспериментальных наблюдений за фактором и откликом составить уравнение зависимости отклика от фактора (уравнение регрессии) и оценить тесноту связи между исследуемыми величинами.

Содержание работы: предварительная обработка данных, графическое изображение данных для получения вывода о характере зависимости, вывод уравнения зависимости, оценка тесноты связи между величинами.

Описание работы. В таблице представлены данные о потреблении электроэнергии владельцами 10 индивидуальных домов.

Число совместно проживающих членов семьи, чел.	2	3	3	4	4	5	5	6	7	7
Годовое потребление электроэнергии, тыс. кВт-час.	15	14	16	19	20	22	23	25	24	22

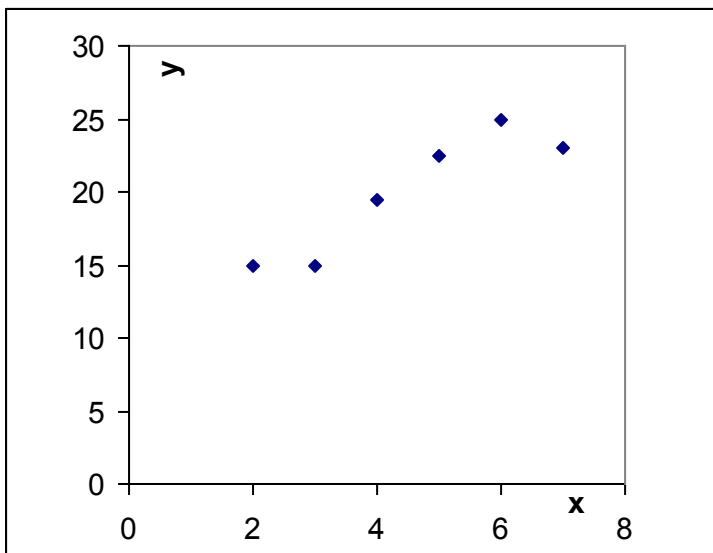
Составим корреляционную таблицу для величин X – число совместно проживающих членов семьи, и Y – годовое потребление электроэнергии.

$X \backslash Y$	14	15	16	19	20	22	23	24	25	l_i
2		1								1
3	1		1							2
4				1	1					2
5						1	1			2
6									1	1
7						1		1		2
k_j	1	1	1	1	1	2	1	1	1	10

В корреляционной таблице указаны отдельные значения x_i и y_j величин X и Y , частоты этих значений l_i и k_j , в клетках таблицы указаны частоты N_{ij} соответствующих пар значений x_i и y_j .

Найдём средние значения величины Y для каждого значения величины X , внесём их в отдельную таблицу.

x_i	2	3	4	5	6	7
$\bar{y}_i$	15	15	19,5	22,5	25	23



Построим поле рассеяния, отложив данные этой таблицы на координатной плоскости.

Расположение точек на графике указывает на нелинейную зависимость между величинами, для её описания выбираем квадратичную зависимость:

$$y = ax^2 + bx + c,$$

графиком которой является парабола.

Параметры этой зависимости a , b и c найдём из сле-

дующей системы, которая составлена на основе системы (9), приведённой в лекции 2, с учётом неединичных частот отдельных значений X и Y :

$$\begin{cases} a \sum_{i=1}^n x_i^4 l_i + b \sum_{i=1}^n x_i^3 l_i + c \sum_{i=1}^n x_i^2 l_i = \sum_{i=1, j=1}^{i=n, j=m} y_i x_i^2 N_{ij} \\ a \sum_{i=1}^n x_i^3 l_i + b \sum_{i=1}^n x_i^2 l_i + c \sum_{i=1}^n x_i l_i = \sum_{i=1, j=1}^{i=n, j=m} y_i x_i N_{ij} \\ a \sum_{i=1}^n x_i^2 l_i + b \sum_{i=1}^n x_i l_i + c N = \sum_{j=1}^{j=m} y_j k_j \end{cases}$$

В данной системе n – количество отдельных значений величины X , m – количество отдельных значений величины Y , $N = \sum N_{ij}$. В рассматриваемом примере $n = 6$, $m = 9$. Вычислим необходимые суммы.

$$\sum_{i=1}^6 x_i^4 l_i = 2^4 \cdot 1 + 3^4 \cdot 2 + 4^4 \cdot 2 + 5^4 \cdot 2 + 6^4 \cdot 1 + 7^4 \cdot 2 = 8038,$$

$$\sum_{i=1}^6 x_i^3 l_i = 2^3 \cdot 1 + 3^3 \cdot 2 + 4^3 \cdot 2 + 5^3 \cdot 2 + 6^3 \cdot 1 + 7^3 \cdot 2 = 1342,$$

$$\sum_{i=1}^6 x_i^2 l_i = 2^2 \cdot 1 + 3^2 \cdot 2 + 4^2 \cdot 2 + 5^2 \cdot 2 + 6^2 \cdot 1 + 7^2 \cdot 2 = 238,$$

$$\sum_{i=1}^6 x_i l_i = 2 \cdot 1 + 3 \cdot 2 + 4 \cdot 2 + 5 \cdot 2 + 6 \cdot 1 + 7 \cdot 2 = 46,$$

$$\sum_{i=1, j=1}^{i=6, j=9} y_j x_i^2 N_{ij} = 15 \cdot 2^2 \cdot 1 + 14 \cdot 3^2 \cdot 1 + 16 \cdot 3^2 \cdot 1 + 19 \cdot 4^2 \cdot 1 + 20 \cdot 4^2 \cdot 1 +$$

$$+ 22 \cdot 5^2 \cdot 1 + 23 \cdot 5^2 \cdot 1 + 25 \cdot 6^2 \cdot 1 + 22 \cdot 7^2 \cdot 1 + 24 \cdot 7^2 \cdot 1 = 5233,$$

$$\sum_{i=1, j=1}^{i=6, j=9} y_j x_i N_{ij} = 15 \cdot 2 \cdot 1 + 14 \cdot 3 \cdot 1 + 16 \cdot 3 \cdot 1 + 19 \cdot 4 \cdot 1 + 20 \cdot 4 \cdot 1 +$$

$$+ 22 \cdot 5 \cdot 1 + 23 \cdot 5 \cdot 1 + 25 \cdot 6 \cdot 1 + 22 \cdot 7 \cdot 1 + 24 \cdot 7 \cdot 1 = 973,$$

$$\sum_{j=1}^9 y_j k_j = 14 \cdot 1 + 15 \cdot 1 + 16 \cdot 1 + 19 \cdot 1 + 20 \cdot 1 + 22 \cdot 2 + 23 \cdot 1 + 24 \cdot 1 + 25 \cdot 1 = 200.$$

После подстановки получим систему:

$$\begin{cases} 8038a + 1342b + 238c = 5233, \\ 1342a + 238b + 46c = 973, \\ 238a + 46b + 10c = 200. \end{cases}$$

Решая эту систему так, как описано в практическом занятии №3, найдём уравнение зависимости Y от X :

$$y = -0,395x^2 + 5,707x + 3,151.$$

Для сравнения с опытными данными составим таблицу значений y в зависимости от x , округлив эти значения до десятых.

x	2	3	4	5	6	7
y	13,0	16,7	19,7	21,8	23,2	23,7

Нарисуем график полученной зависимости на поле рассеяния, используя полученные значения, а также то, что вершина параболы находится в точке (7,2; 23,8).

Как видим, полученная зависимость весьма удовлетворительно описывает экспериментальные данные в исследованном диапазоне изменения величин.

Оценим тесноту связи между величинами. В случае нелинейной зависимости это можно сделать, вычислив индекс корреляции ρ_{xy} по формуле:

$$\rho_{xy} = \sqrt{1 - \frac{\sum_{i=1}^{i=n} (y_i - y(x_i))^2}{\sum_{i=1}^{i=n} (y_i - \bar{y})^2}}.$$

В этой формуле y_i - зарегистрированные в эксперименте значения функции (отклика) для $X = x_i$;

$y(x_i)$ - значения отклика, рассчитанные по формуле функциональной зависимости, найденной методом наименьших квадратов;

$\bar{y}$ - среднее арифметическое экспериментальных значений отклика;

n - количество зарегистрированных значений фактора (величины X).

Вычислим необходимые суммы:

$$\begin{aligned} \sum_{i=1}^{i=6} (y_i - y(x_i))^2 &= (15 - 13)^2 + (14 - 16,7)^2 + (16 - 16,7)^2 + (19 - 19,7)^2 + (20 - 19,7)^2 + \\ &+ (22 - 21,8)^2 + (23 - 21,8)^2 + (25 - 23,2)^2 + (22 - 23,7)^2 + (24 - 23,7)^2 = 20,06; \\ \sum_{i=1}^{i=6} (y_i - \bar{y})^2 &= (15 - 20)^2 + (14 - 20)^2 + (16 - 20)^2 + (19 - 20)^2 + (20 - 20)^2 + \\ &+ (22 - 20)^2 + (23 - 20)^2 + (25 - 20)^2 + (22 - 20)^2 + (24 - 20)^2 = 136. \end{aligned}$$

Подставив найденные суммы в формулу для индекса корреляции, найдём:

$$\rho_{xy} = \sqrt{1 - \frac{20,06}{136}} \approx 0,92.$$

Значение индекса корреляции указывает на высокую тесноту связи между количеством совместно проживающих в доме членов семьи и годовым потреблением электроэнергии.

Задание

В таблице представлены результаты анализа спроса на различные модели автомобилей одной марки в зависимости от их цены.

Модель	1	2	3	4	5	6	7	8	9	10	11	12
Цена, тыс. долл.	25	28	29	27	29	28	29	24	25	23	24	27
Кол-во автомобилей, проданных за месяц	55	54	40	51	30	45	46	45	50	35	30	53

Составить корреляционную таблицу, изобразить поле рассеяния, выбрать вид зависимости между величиной X - ценой автомобиля и Y - величиной спроса на него, определить её параметры, оценить тесноту связи между величинами.

В отчёте представить корреляционную таблицу, поле рассеяния, значения всех сумм, необходимых для вычисления формулы зависимости и индекса корреляции, формулу полученной зависимости и её график (на поле рассеяния) и значение индекса корреляции.

Лабораторная работа № 4. Обработка результатов двухфакторного пассивного эксперимента

Цель работы: по результатам наблюдений составить уравнение зависимости отклика от факторов (уравнение множественной регрессии) и оценить тесноту связи между исследуемыми величинами.

Содержание работы: предварительная обработка данных, графическое изображение данных для получения вывода о характере зависимости, вывод уравнения зависимости, оценка тесноты связи между величинами.

Описание работы. В таблице представлены результаты изучения влияния изменения объёма промышленного производства и среднедушевого дохода на товарооборот. Для этого по 20 регионам РФ были получены следующие данные.

№	Розничный товарооборот в % к предыдущему году	Объём промышленного производства в % к предыдущему году	Среднедушевой денежный доход в % к предыдущему году
1	89	85	88
2	75	70	85
3	82	86	81
4	84	80	87

5	91	97	87
6	92	79	110
7	89	92	102
8	107	99	108
9	89	83	94
10	87	77	92
11	96	88	82
12	75	89	85
13	74	72	84
14	86	80	105
15	73	81	94
16	100	97	98
17	87	73	92
18	87	75	95
19	98	84	101
20	81	75	88

Нужно найти уравнение зависимости отклика Z (розничного товарооборота) от факторов X (объёма промышленного производства) и Y (среднедушевого денежного дохода), оценить тесноту связи между величинами.

В общем случае, зависимость значений z отклика Z от значений x фактора X и значений y фактора Y описывается формулой:

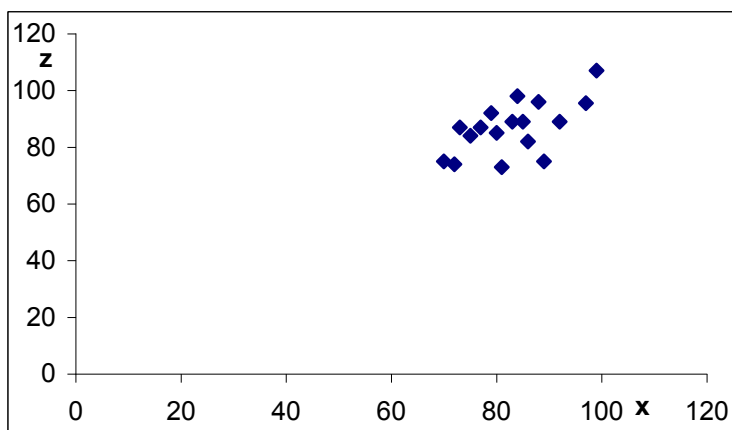
$$z = a_0 + a_1x + b_1y + a_2x^2 + c_2xy + b_2y^2 + \dots$$

Чтобы выбрать вид зависимости (другими словами, указать, какие слагаемые в общей формуле следует оставить), проанализируем полученные данные. Составим корреляционную таблицу для величин X и Z :

$Z \backslash X$	70	72	73	75	77	79	80	81	83	84	85	86	88	89	92	97	99
73								1									
74		1															
75	1													1			
81				1													
82												1					
84							1										
86							1										
87			1	1	1												
89									1		1				1		
91																1	
92						1											
96													1				
98										1							
100																1	
107																	1

Пользуясь таблицей, для каждого значения X найдём средние значения Z :

x_i	70	72	73	75	77	79	80	81	83	84	85	86	88	89	92	97	99
$\bar{z}_i$	75	74	87	84	87	92	85	73	89	98	89	82	96	75	89	95,5	107



Нанесём полученные данные на координатную плоскость.

Расположение точек указывает на линейную зависимость между X и Z .

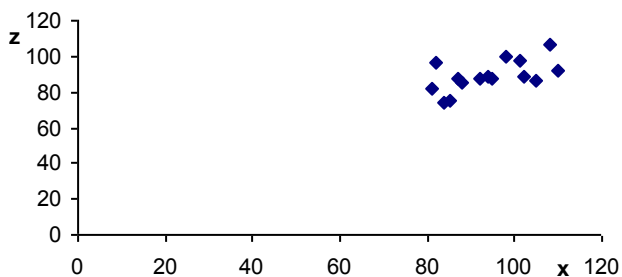
Составим теперь корреляционную таблицу для величин Y и Z .

$Z \backslash Y$	81	82	84	85	87	88	92	94	95	98	101	102	105	108	110
73								1							
74			1												
75				2											
81						1									
82	1														
84					1										
86													1		
87							2		1						
89						1		1				1			
91					1										
92															1
96		1													
98											1				
100										1					
107														1	

Найдём соответствующие средние значения величины Z для каждого значения величины Y :

y_i	81	82	84	85	87	88	92	94	95	98	101	102	105	108	110
$\bar{z}_i$	82	96	74	75	87,5	85	87	89	87	100	98	89	86	107	92

Отложим полученные данные на координатной плоскости.



По расположению точек выбираем линейную зависимость между Z и Y .

Таким образом, общая формула зависимости величины Z от величин X и Y будет иметь вид:

$$z = a_0 + a_1x + b_1y.$$

Числовые параметры этой зависимости найдём из системы, составленной применением метода наименьших квадратов:

$$\begin{cases} a_0 n + a_1 \sum_{i=1}^n x_i + b_1 \sum_{i=1}^n y_i = \sum_{i=1}^n z_i, \\ a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 + b_1 \sum_{i=1}^n x_i y_i = \sum_{i=1}^n x_i z_i, \\ a_0 \sum_{i=1}^n y_i + a_1 \sum_{i=1}^n x_i y_i + b_1 \sum_{i=1}^n y_i^2 = \sum_{i=1}^n y_i z_i. \end{cases}$$

В данной системе n - количество зарегистрированных значений каждой величины, в данном случае $n = 20$.

Найдя необходимые суммы, получим систему:

$$\begin{cases} 20a_0 + 1662a_1 + 1858b_1 = 1742, \\ 1662a_0 + 139508a_1 + 154721b_1 = 145676, \\ 1858a_0 + 154721a_1 + 174040b_1 = 162648. \end{cases}$$

Решая её, как описывалось ранее, найдём уравнение зависимости:

$$z = -0,371 + 0,554x + 0,446y.$$

Для проверки полученной формулы для каждой пары значений X и Y , представленных в таблице, вычислим по ней значения Z (округлив их до целых) и сравним с имеющимися в таблице.

№	x_i	y_i	$Z_{i\text{выч}}$	$Z_{i\text{табл}}$	$\Delta Z = Z_{i\text{выч}} - Z_{i\text{табл}}$
1	85	88	86	89	-3
2	70	85	76	75	1
3	86	81	83	82	1
4	80	87	84	84	-1
5	97	87	92	91	1
6	79	110	92	92	0
7	92	102	96	89	7
8	99	108	103	107	-4
9	83	94	88	89	-1
10	77	92	83	87	-4
11	88	82	85	96	-11
12	89	85	87	75	12
13	72	84	77	74	3
14	80	105	91	86	5
15	81	94	86	73	13
16	97	98	97	100	-3
17	73	92	81	87	-6
18	75	95	84	87	-3
19	84	101	91	98	-7
20	75	88	80	81	-1

Результаты, представленные в таблице, показывают, что отклонения вычисленных значений от табличных небольшие: намного меньше самих значений, а их сумма практически равна нулю. Таким образом, можно признать, что полученная формула адекватно описывает статистические данные.

Оценим тесноту связи между величинами. В случае линейной зависимости это можно сделать, вычислив парные коэффициенты корреляции:

$$\text{коэффициент корреляции между } X \text{ и } Y \quad r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{\overline{x^2} - (\bar{x})^2} \cdot \sqrt{\overline{y^2} - (\bar{y})^2}};$$

$$\text{коэффициент корреляции между } X \text{ и } Z \quad r_{xz} = \frac{\overline{xz} - \bar{x} \cdot \bar{z}}{\sqrt{\overline{x^2} - (\bar{x})^2} \cdot \sqrt{\overline{z^2} - (\bar{z})^2}};$$

$$\text{коэффициент корреляции между } Y \text{ и } Z \quad r_{yz} = \frac{\overline{yz} - \bar{y} \cdot \bar{z}}{\sqrt{\overline{y^2} - (\bar{y})^2} \cdot \sqrt{\overline{z^2} - (\bar{z})^2}}.$$

В приведённых формулах использованы следующие обозначения:

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^n x_i}{n}; & \bar{y} &= \frac{\sum_{i=1}^n y_i}{n}; & \bar{z} &= \frac{\sum_{i=1}^n z_i}{n}; & \overline{x^2} &= \frac{\sum_{i=1}^n x_i^2}{n}; & \overline{y^2} &= \frac{\sum_{i=1}^n y_i^2}{n}; & \overline{z^2} &= \frac{\sum_{i=1}^n z_i^2}{n}; \\ \overline{xy} &= \frac{\sum_{i=1}^n x_i \cdot y_i}{n}; & \overline{xz} &= \frac{\sum_{i=1}^n x_i \cdot z_i}{n}; & \overline{yz} &= \frac{\sum_{i=1}^n y_i \cdot z_i}{n}. \end{aligned}$$

Используя уже найденные суммы и вычислив $\sum_{i=1}^n z_i^2 = 153276$, найдём:

$$r_{xy} \approx 0,227, \quad r_{xz} \approx 0,623; \quad r_{yz} \approx 0,548.$$

Также можно вычислить частные коэффициенты корреляции, которые характеризуют тесноту связи между откликом и каким-либо фактором при фиксированном воздействии другого фактора:

$$r_{zx}(y = \text{const}) = \frac{r_{zx} - r_{zy} \cdot r_{xy}}{\sqrt{(1 - r_{zy}^2)} \cdot \sqrt{(1 - r_{xy}^2)}}; \quad r_{zy}(x = \text{const}) = \frac{r_{zy} - r_{zx} \cdot r_{xy}}{\sqrt{(1 - r_{zx}^2)} \cdot \sqrt{(1 - r_{xy}^2)}}.$$

Для рассматриваемого примера эти коэффициенты будут равны:

$$r_{zx}(y = \text{const}) \approx 0,612, \quad r_{zy}(x = \text{const}) \approx 0,533.$$

Задание

В таблице представлены результаты обследования 20 предприятий некоторой отрасли промышленности.

№	Производительность труда (в млн. руб. на 1 рабочего)	Энерговооружённость (кВт-час на 1 рабочего)	Доля рабочих ручного труда в %
1	9,8	4,8	40
2	6,7	2,8	59
3	12,4	7,0	38
4	6,9	3,8	97
5	11,8	5,5	31
6	7,3	3,0	56
7	8,4	3,4	45
8	10,7	5,2	35
9	11,1	5,4	32
10	7,3	2,9	54
11	8,8	3,9	43
12	10,0	5,0	36
13	12,1	6,2	29
14	11,0	5,1	32
15	6,8	2,7	55
16	7,0	2,9	50
17	12,0	6,1	30
18	7,5	2,3	48
19	8,0	3,4	45
20	9,2	3,8	49

Провести обработку данных для установления типа зависимости величины Z - производительности труда от величин X - энерговооружённости на 1 рабочего и Y - доли рабочих ручного труда, составить уравнение регрессии, сделать проверку найденного уравнения, оценить тесноту связи между величинами.

В отчёте представить корреляционные таблицы зависимости Z от X и Z от Y , их графическое изображение, уравнение регрессии с необходимыми расчётами, таблицу проверки уравнения регрессии, парные и частные коэффициенты корреляции.

Оглавление

Введение.....	3
Лекция 1. Виды экспериментов. Обработка результатов простейшего эксперимента	3
1.1. Эксперимент и его виды	3
1.2. Обработка результатов простейшего эксперимента	4
Лекция 2. Нахождение формулы функциональной зависимости методом наименьших квадратов	9
2.1. Основные положения метода наименьших квадратов	9
2.2. Определение параметров функциональных зависимостей.....	10
Практическое занятие 1. Нахождение точечной и интервальной оценок измеряемой величины при прямых и косвенных измерениях.....	13
Практическое занятие 2. Применение метода наименьших квадратов в случае линейной и квадратичной функций	14
Практическое занятие 3. Применение метода наименьших квадратов в случае показательной, логарифмической и гиперболической зависимостей	18
Практическое занятие 4. Установление тесноты связи между величинами в однофакторном эксперименте	22
Лабораторная работа № 1. Нахождение точечной и интервальной оценок измеряемой величины при косвенных измерениях	25
Лабораторная работа № 2. Обработка результатов однофакторного пассивного эксперимента в случае линейной зависимости	26
Лабораторная работа № 3. Обработка результатов однофакторного пассивного эксперимента в случае нелинейной зависимости.....	30
Лабораторная работа № 4. Обработка результатов двухфакторного пассивного эксперимента	33

Подписано в печать 05.09.2013.	Усл. печ. л. 2,44	Тираж 16 экз.
Печать офсетная.	Бумага писчая.	Заказ № 425.

Отпечатано: РИО ВоГТУ, г. Вологда, ул. Ленина, 15