

ФГБОУ ВО «Брянский государственный аграрный университет»

КОМОГОРЦЕВ В.Ф.

**МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ ПРОЦЕССОВ
В КОМПОНЕНТАХ ПРИРОДЫ**

Учебное пособие

Брянская область - 2018

УДК 517.8 (07)
ББК 22.16
К 63

Комогорцев, В. Ф. Математическое моделирование процессов в компонентах природы: учебное пособие / В. Ф. Комогорцев. - Брянск. – Изд-во Брянский ГАУ, 2018. - 131 с.

Учебное пособие соответствует федеральным государственным образовательным стандартам высшего образования и может быть использовано магистрантами очной и заочной форм обучения направления подготовки 20.04.02 «природообустройство и водопользование» в ходе изучения одноименной дисциплины «Математическое моделирование процессов в компонентах природы». В учебном пособии последовательно изложены основные понятия, определения, задачи указанной дисциплины и основные способы их решения. Учебное пособие может быть использовано как для аудиторной, так и для самостоятельной работы студентов.

РЕЦЕНЗЕНТ:

заведующая кафедрой математики, физики и информатики, ктн. **Ракул Е.А.**

Рекомендовано к изданию учебно-методической комиссией института энергетики и природопользования от 28 июня 2018г., протокол № 8.

ОГЛАВЛЕНИЕ

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

1. Физические модели	5
1.1. Критерии подобия	6
1.2. Элементы теории размерностей. π - теорема	8
2. Аналоговые модели	11
3. Математические модели.....	12
3.1. Основные этапы математического моделирования	12

ГЛАВА 1. ДЕТЕРМИНАНТНЫЕ И СТОХАСТИЧЕСКИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

§1. Примеры детерминантных математических моделей объектов различной физической природы	14
1.1. Задача о траектории брошенного тела без учета сопротивления воздуха.....	14
1.2. Задача о траектории брошенного тела при учете сопротивления воздуха.....	16
1.3. Задача об определении объема выкопанной ямы	19
1.4. Задача о баке с наименьшей площадью поверхности.....	20
1.5. Задача об определении скорости револьверной пули.....	20
1.6. Строительная задача	21
1.7. Транспортная задача	22
1.8. Задача об остывании вскипевшего чайника	23
1.9. Задача о моторной лодке	25
1.10. Задача о барометрической формуле	26
1.11. Задача об определении силы давления воды на плотину.....	28
1.12. Закон Архимеда. Плавание тел в жидкостях.....	31
1.13. Установившееся движение жидкости. Уравнение Бернулли	33
Упражнения	40
§2. Стохастическое моделирование. Регрессионный анализ	42
2.1. Корреляционная связь между числовыми величинами	42
2.2. Парная корреляция при наличии повторных опытов.....	43
2.3. Исследование уравнения парной регрессии на адекватность при наличии повторных опытов.....	58
2.4. Парная корреляция при отсутствии повторных опытов	66
2.5. Исследование уравнения парной регрессии на адекватность при отсутствии повторных опытов.....	70
2.6. Пример построения парной регрессии вручную	72

2.7. Построение парной регрессии с помощью Excel.....	78
2.8. Парные нелинейные регрессии и их построение в Excel	83
2.9. Множественная корреляция и множественная регрессия	85
2.10. Пример построения множественной регрессии в Excel	88
Упражнения	91

ГЛАВА 2. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ

§1. Основы планирования экспериментов.....	94
§2. Полные ортогональные планы типа 2^n	98
§3. Расширенная матрица планирования экспериментов	103
§4. Математическая обработка результатов экспериментов.....	104
§5. Дробный факторный эксперимент	111
§6. Крутое восхождение к оптимуму	115
Числовой пример	116
Приложение (таблицы)	124
Литература	130

ОСНОВНЫЕ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ

Моделирование – это замещение одного объекта (оригинала) другим (моделью) и изучение свойств оригинала путем изучения свойств модели.

Модели разделяют на **материальные** (физические и аналоговые) и **идеальные** (математические).

1. ФИЗИЧЕСКИЕ МОДЕЛИ

Физические модели - это натуральные модели. В частности, физическими моделями являются макеты и опытные образцы изучаемых объектов. Макет жилого района – физическая модель этого района. Уменьшенная копия самолета, испытываемая в аэродинамической трубе – физическая модель реального самолета. Физическими моделями являются и уменьшенные копии плотин, каналов, гидроузлов и других сооружений. Те параметры, которые характеризуют реальный объект, измеряют на его модели, а затем переносят (пересчитывают) их на реальный объект.

Физическое моделирование широко применяется при решении большого числа практических задач в различных областях науки и техники. Им широко пользуются в строительном деле (определение усталостных напряжений и эксплуатационных разрушений, виброзащита и сейсмостойкость различных конструкций и др.); в гидравлике и в гидротехнике (определение конструктивных и эксплуатационных характеристик различных гидротехнических сооружений, условий фильтрации в грунтах, моделирование течений рек, волн, приливов и отливов, и др.); в авиации, ракетной и космической технике (определение характеристик летательных аппаратов и их двигателей, силового и теплового воздействия среды на них, и др.); в судостроении (определение гидродинамических характеристик корпуса, ходовых качеств судов, и др.); в различных областях машиностроения; в нефте-и-газодобыче; в теплотехнике при проектировании и эксплуатации различных тепловых аппаратов; в электротехнике при исследованиях всевозможных электрических систем, и т.д.

Необходимыми условиями физического моделирования являются *геометрическое* и *физическое* подобие модели и реального объекта. Это значит, что в сходственные моменты времени и в сходственных точках пространства значения величин, характеризующих реальный объект и модель, должны быть пропорциональны. Наличие такой пропорциональности позволяет производить пересчет экспериментальных результатов, полученных для модели, путём умножения каждой из определяемых величин на постоянный для всех величин данной размерности множитель - *коэффициент подобия*.

Поскольку физические величины связаны определёнными соотношениями, вытекающими из законов и уравнений физики, то выбрав некоторые из них за основные, можно коэффициенты подобия для всех других величин выразить через коэффициенты подобия тех величин, которые приняты за основные.

Например, в механике основными величинами являются длина l , время t и масса m . Тогда

$$k_l = \frac{l_N}{l_M}; \quad k_t = \frac{t_N}{t_M}; \quad k_m = \frac{m_N}{m_M} \quad (1)$$

- коэффициенты подобия длин, времён и масс. Измерив (или вычислив) линейные размеры, времена или массы на модели (с индексом М), по формулам (1) мы найдем и сходственные им параметры для натурального объекта (с индексом N). А, например, для площадей S и объемов V их коэффициенты подобия k_S и k_V уже выразятся через коэффициент подобия k_l для длин:

$$k_S = \frac{S_N}{S_M} = k_l^2; \quad k_V = \frac{V_N}{V_M} = k_l^3 \quad (2)$$

Далее, поскольку скорость $v = l/t$, то коэффициент подобия скоростей k_v можно выразить через коэффициенты подобия длин и времен в виде

$$k_v = \frac{v_N}{v_M} = \frac{l_N / t_N}{l_M / t_M} = \frac{l_N / l_M}{t_N / t_M} = \frac{k_l}{k_t} \quad (3)$$

Аналогично, коэффициент подобия сил, на основании второго закона Ньютона $F = m \cdot a$, можно определить как $k_F = k_m \cdot k_a$, где, в свою очередь, $k_a = k_v / k_t$.

1.1 Критерии подобия

Некоторые безразмерные комбинации величин, характеризующих данный физический объект или данное физическое явление, имеют для модели и натуры одно и то же значение. Эти безразмерные комбинации физических величин называются *критериями подобия*. Равенство всех критериев подобия для модели и натуры является условием качественного физического моделирования. При этом число и вид критериев подобия для каждого моделируемого явления зависит от его природы и особенностей.

В частности, критерием подобия, содержащем действующие в оригинале и в модели силы, является *число Ньютона* (Ne).

Получим выражение для Ne. Так как для сил любой природы имеет место второй закон Ньютона $F = ma$, то

$$k_F = k_m \cdot k_a = \frac{k_m \cdot k_v}{k_t} = \frac{k_m \cdot k_v^2}{k_t} \quad (4)$$

Заменяя здесь масштабные коэффициенты их отношениями, получим:

$$k_F = \frac{F_N}{F_M} = \frac{m_N \cdot v_N^2 \cdot l_M}{m_M \cdot v_M^2 \cdot l_N},$$

откуда следует:

$$\frac{m_M \cdot v_M^2}{F_M \cdot l_M} = \frac{m_N \cdot v_N^2}{F_N \cdot l_N} = Ne \quad (5)$$

Если в исследуемой системе действуют силы разной физической природы (силы тяготения, силы трения, электромагнитные силы) и т.д., то полного подобия всех этих сил с одним и тем же коэффициентом подобия k_F достигнуть не удастся. Действительно, то соотношение, которое будет иметь место для объемных сил (например, для сил тяжести или электромагнитных сил), не будет выполняться для поверхностных сил (например, для сил трения), ибо первые из них пропорциональны объемам, а вторые пропорциональны площадям поверхностей трения. Тогда выделяют среди действующих сил основную и пренебрегают действием других сил. Такое подобие натурального объекта и его модели будет лишь частичным.

Рассмотрим критерии подобия Ne для некоторых случаев доминирования одной из сил.

Пусть преобладающей в системе силой будет сила земного тяготения. Подставляя её выражение $F = mg$ в формулу (5) критерия Ньютона, получим число (критерий) Фруда (Fr):

$$Fr = \frac{v_M^2}{gl_M} = \frac{v_N^2}{gl_N} \quad (6)$$

А если речь идет о течении вязкой жидкости, то преобладающей силой в ней может являться сила трения

$$F = \mu \rho \frac{dv}{dn} S \quad (7)$$

между слоями текущей жидкости. Здесь μ - кинематическая вязкость жидкости; ρ - плотность жидкости; $\frac{dv}{dn}$ - градиент скорости v течения жидкости, показывающий, как быстро меняется эта скорость в направлении n , перпендикулярном течению жидкости; S - площадь поверхности слоя жидкости. Так как

$$\frac{dv}{dn} \approx \frac{\Delta v}{\Delta n} \approx \frac{v}{l},$$

то

$$Ne = \frac{mv^2}{Fl} = \frac{\rho l^3 v^2 l}{\mu \rho v l^2 l} = \frac{vl}{\mu}$$

Показатели степеней в (10) определяются таким образом, чтобы комплексы π были безразмерными.

Пусть a – произвольная механическая величина. Её размерность $[a]$ всегда можно представить в виде $[a] = [l]^\alpha \cdot [m]^\beta \cdot [t]^\gamma = m^\alpha \cdot кг^\beta \cdot сек^\gamma$. А так как число π должно быть безразмерным, то выписав размерности каждой из m величин a_1, a_2, \dots, a_m , на основании равенств (10) мы сможем определить показатели и всех степеней этих величин.

Рассмотрим применение π -теоремы на примере обтекания твердого тела жидкостью. Или, что то же самое, на примере движения твердого тела в покоящейся жидкости.

Очевидно, что сила F воздействия жидкости на тело будет определяться скоростью v набегающего потока, плотностью жидкости ρ , кинематической вязкостью жидкости μ , характерной (максимальной) площадью поперечного сечения тела S . Таким образом, имеем некую зависимость

$$\varphi_1(F, v, \rho, \mu, S) = 0. \quad (11)$$

Здесь $m=5, n=3$. Согласно π -теореме равенство (11) можно привести к виду

$$\varphi_2(\pi_1, \pi_2) = 0 \quad (12)$$

Выберем в качестве первых n сомножителей (v, ρ, S) . Тогда

$$\begin{aligned} \pi_1 &= v^{x_1} \cdot \rho^{y_1} \cdot S^{z_1} \cdot F \\ \pi_2 &= v^{x_2} \cdot \rho^{y_2} \cdot S^{z_2} \cdot \mu \end{aligned} \quad (13)$$

Выпишем размерности входящих в равенства (13) величин в системе СИ:

$$[v] = [m \cdot сек^{-1}]; \quad [\rho] = [кг \cdot м^{-3}]; \quad [S] = [м^2]; \quad [F] = [кг \cdot м \cdot сек^{-2}]; \quad [\mu] = [м^2 \cdot сек^{-1}]$$

(размерность кинематической вязкости μ следует из формулы (7)). Тогда из равенств (13) вытекает:

$$\begin{aligned} \pi_1 &= [m \cdot сек^{-1}]^{x_1} \cdot [кг \cdot м^{-3}]^{y_1} \cdot [м^2]^{z_1} \cdot [кг \cdot м \cdot сек^{-2}] = [м^0 \cdot кг^0 \cdot сек^0] \\ \pi_2 &= [m \cdot сек^{-1}]^{x_2} \cdot [кг \cdot м^{-3}]^{y_2} \cdot [м^2]^{z_2} \cdot [м^2 \cdot сек^{-1}] = [м^0 \cdot кг^0 \cdot сек^0] \end{aligned} \quad (14)$$

Объединяя степени с одинаковыми основаниями, из (14) получаем:

$$\begin{aligned} [м]^{x_1-3y_1+2z_1+1} &= [м]^0; \quad [кг]^{y_1+1} = [кг]^0; \quad [сек]^{-x_1-2} = [сек]^0; \\ [м]^{x_2-3y_2+2z_2+2} &= [м]^0; \quad [кг]^{y_2} = [кг]^0; \quad [сек]^{-x_2-1} = [сек]^0 \end{aligned} \quad (15)$$

Сравнивая показатели степеней при одинаковых основаниях, получаем две системы линейных алгебраических уравнений:

$$1) \begin{cases} x_1 - 3y_1 + 2z_1 + 1 = 0 \\ y_1 + 1 = 0 \\ -x_1 - 2 = 0 \end{cases} \quad 2) \begin{cases} x_2 - 3y_2 + 2z_2 + 2 = 0 \\ y_2 = 0 \\ -x_2 - 1 = 0 \end{cases} \quad (16)$$

Решая эти системы, получаем:

$$\{x_1 = -2; y_1 = -1; z_1 = -1\}; \quad \{x_2 = -1; y_2 = 0; z_2 = -1/2\} \quad (17)$$

Подставим эти значения в выражения (13):

$$\pi_1 = v^{-2} \cdot \rho^{-1} \cdot S^{-1} \cdot F = \frac{F}{\rho S v^2} \quad (18)$$

$$\pi_2 = v^{-1} \cdot \rho^0 \cdot S^{-1/2} \cdot \mu = \frac{\mu}{v \sqrt{S}}$$

Таким образом, равенство (12) примет вид:

$$\varphi_2 \left(\frac{F}{\rho S v^2}; \frac{\mu}{v \sqrt{S}} \right) = 0 \quad (19)$$

Согласно (8), безразмерная величина

$$\pi_2 = \frac{\mu}{v \sqrt{S}} = \frac{\mu}{v l} = \frac{1}{\text{Re}}, \quad (20)$$

где Re - число (критерий) Рейнольдса (см. 8). Поэтому равенство (19) примет вид:

$$\varphi_2 \left(\frac{F}{\rho S v^2}; \frac{1}{\text{Re}} \right) = 0 \quad (21)$$

Отсюда

$$\frac{F}{\rho S v^2} = f \left(\frac{1}{\text{Re}} \right), \text{ то есть } F = C \rho S v^2, \quad (22)$$

где C - безразмерная константа, зависящая от формы тела и вязкости жидкости. Эту формулу получил еще Ньютон на основании опытных данных. Она определяет структуру зависимости силы воздействия потока жидкости на погруженное в него тело (или, что одно и то же, силу сопротивления покоящейся

жидкости движущемуся в ней телу) от геометрических и физических свойств тела и жидкости. Формула эта широко применяется при описании движения твердых тел в сплошных жидких и газообразных средах.

Следует отметить, что хотя π -теорема широко применяется для определения структуры формул, описывающих реальные объекты, она не является универсальным средством, так как неучет влияния какой-либо физической величины может привести к неверной структуре этих формул.

2. АНАЛОГОВЫЕ МОДЕЛИ

Аналоговые модели - это натуральные объекты, отличающиеся по своей физической природе от оригинала, но имеющие с оригиналом сходные процессы функционирования.

Например, теоретически установлено, что задача о напряжениях, возникающих при кручении бруса, сводится к тому же дифференциальному уравнению, что и задача о прогибах упругой (например, резиновой) пленки, нагруженной равномерным давлением и натянутой по контуру той же формы, что и поперечное сечение бруса. Аналогом напряжения в бруске является угол, который составляет касательная к поверхности пленки к плоскости контура, а аналогом крутящего момента – объем, заключенный между плоскостью контура и поверхностью пленки. Сделав сосуд с нужной формой дна и налив в него жидкость, мы получим аналоговую модель закручиваемого бруса. Изучая (обмеряя) прогнувшуюся пленку, мы сможем установить и интересующие нас силовые факторы в закручиваемом бруске.

Еще пример аналога двух систем. Первая из них, имеющая механическую природу, состоит из оси, передающей через пружину и маховик вращение валу, жестко связанному с маховиком. При этом маховик погружен (частично или полностью) в вязкую тормозящую жидкость, Вторая система - электрическая - состоит из источника электродвижущей силы, соединённого через катушку индуктивности, конденсатор и активное сопротивление со счётчиком электрической энергии. Если подобрать значения индуктивности, ёмкости и сопротивления так, чтобы они определённым образом соответствовали упругости пружины, инерции маховика и трению жидкости, то эти системы обнаружат структурное и функциональное сходство, выражаемое в том, что они будут описываться одним и тем же дифференциальным уравнением. Это уравнение может служить «теоретической моделью» обеих систем. А любая из них — «экспериментальной моделью» этого уравнения и «аналоговой моделью» другой системы. Такого рода аналогии лежат в основе *электрического моделирования* механических систем. Электрические модели гораздо более удобны для экспериментального исследования, нежели моделируемые механические.

Современная наука и техника вообще широко использует различные аналоговые модели. В частности, в гидромеханике аналоговые модели базируются на подобии законов движения влаги в почве и законов движения электрического тока в различных проводящих системах. Метод создания этих моделей называется «электрогидродинамической аналогией» (ЭГДА). Исследования на

базе ЭГДА обычно проводят с помощью специальных приборов - электроинтегратора и сеточных интеграторов типа ЭГДА 9/60, МСМ-1 и др.

3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Математическая модель реального объекта – это нематериальная, идеальная модель этого объекта, представляющая собой систему математических соотношений и формул, связывающих наиболее существенные параметры объекта. При этом можно использовать любые математические средства – дифференциальное и интегральное исчисления, дифференциальные и интегральные уравнения, ряды, теорию вероятностей, и т.д. По существу вся математика создана для составления и исследования математических моделей реальных или воображаемых объектов или процессов. Причем в последнее время все большую роль в этом исследовании играют ЭВМ. Расчетные формулы, которые используют инженеры для анализа и синтеза всевозможных материальных систем, обычно выведены из математических моделей этих систем. Хорошая

Математическая модель объекта позволяет глубже и полнее исследовать реальный объект, чем это можно сделать при исследовании его физической модели.

Математическое моделирование и связанный с ним компьютерный эксперимент вообще незаменимы в тех случаях, когда натуральный эксперимент невозможен или затруднен по тем или иным причинам. Например, невозможно прямыми измерениями исследовать поле напряжений или деформаций в толще деформируемого материала. Нельзя поставить натуральный эксперимент в истории, чтобы проверить, «что было бы, если бы...» Невозможно проверить правильность той или иной космологической теории. В принципе возможно, но вряд ли разумно, поставить эксперимент по распространению какой-либо болезни, например чумы, или осуществить ядерный взрыв, чтобы изучить его последствия. Однако все это можно сделать на компьютере, построив предварительно математические модели изучаемых явлений.

В настоящее время наиболее широко применяется именно математическое моделирование.

Простейший пример математической модели: объект – точка числовой оси, а её координата (число) – математическая модель этого объекта. Еще простой пример: математической моделью линии на плоскости или в пространстве является её уравнение. Например, уравнение $y = x^2$ - это математическая модель хорошо всем известной параболы. Более сложные геометрические объекты имеют и более сложные математические модели (уравнения). Но если эти модели (уравнения) построены, то с помощью математического анализа этих моделей (этих уравнений) можно полно и всесторонне исследовать соответствующие им геометрические объекты.

3.1 Основные этапы математического моделирования

1) Обозначение целей моделирования. Целью математического моделирования изучаемого объекта может быть: а) простое познавательное изучение объекта; б) управление объектом. Например, при построении математических моделей атмосферы (климата), геологических подвижек земной коры и других масштабных систем, которыми человек управлять пока не в состоянии, целью математического моделирования может быть лишь просто изучение и, по мере возможности, прогнозирование процессов, происходящих в этих системах. Но во многих других случаях основной целью математического моделирования объекта является не его пассивное познавательное изучение, а управление этим объектом. Например, математическая модель технологического или экономического процесса должна позволять не только изучать процесс, но и давать рекомендации по его управлению и поиску наилучших условий для его протекания.

2) Построение математической модели. На этом этапе сначала на качественном уровне рассматривается подлежащий изучению «нематематический» объект — явление природы, конструкция, механизм, экономический план, производственный процесс и т. д. Сначала выявляются основные особенности явления и связи между ними на качественном уровне. Определяются перечень всех входных и выходных параметров объекта исследования и их предполагаемое влияние на достижение целей моделирования. Устанавливается словесная фиксация основных связей входных и выходных параметров объекта. Затем найденные качественные зависимости формулируются на языке математики, то есть строится математическая модель объекта. Это самая трудная стадия моделирования.

3) Решение математической задачи, к которой приводит модель. На этом этапе большое внимание уделяется разработке алгоритмов аналитического или численного (на ЭВМ) решения полученной математической задачи, при помощи которых результат может быть найден с необходимой точностью и за допустимое время.

4) Интерпретация полученных следствий из математической модели. Следствия, выведенные из модели на языке математики, интерпретируются на языке, принятом в данной области.

5) Проверка адекватности модели. На этом этапе выясняется, согласуются ли результаты эксперимента с теоретическими следствиями из модели в пределах определенной точности.

6) Модификация модели. На этом этапе происходит либо усложнение модели, чтобы она была более адекватной действительности, либо ее упрощение ради достижения практически приемлемого решения.

ГЛАВА 1

ДЕТЕРМИНАНТНЫЕ И СТОХАСТИЧЕСКИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Математические модели подразделяются на а) *детерминантные* и б) *стохастические*.

Детерминантные модели базируются на уравнениях, следующих из физических законов протекания процессов в объекте. А *стохастические* модели основаны на построенных с помощью теории вероятности и математической статистики уравнениях, связывающих величины, характеризующие изучаемый объект, когда исходные данные для таких уравнений получены опытным путем.

По-другому можно сказать, что стохастическая модель рассматривает изучаемый объект как «черный ящик», не обращая внимания на физическую суть протекающих внутри него процессов. А детерминантная модель объекта основывается именно на физических законах протекания процессов в нем.

На практике обычно применяются оба метода математического моделирования совместно. При этом одни характеризующие объект параметры моделируются стохастически, а другие детерминантно.

Эвристическое (экспертное) моделирование

Особым, четвертым видом моделирования, является *эвристическое моделирование*. Этот метод базируется на методах современного раздела информатики - методах «искусственного интеллекта», и заключается в использовании знаний специалистов – экспертов о поведении изучаемого объекта при различных условиях.

§1. Примеры детерминантных математических моделей объектов различной физической природы

1.1 Задача о траектории брошенного тела без учета сопротивления воздуха

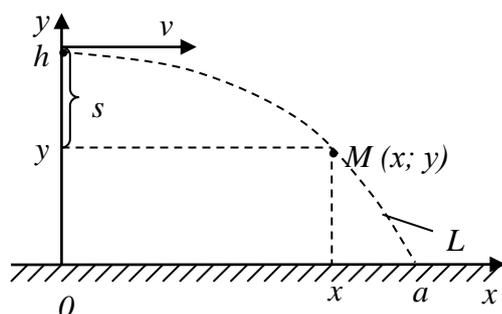


Рис. 1

Поднятое на высоту h компактное массивное тело брошено горизонтально со скоростью v (например, с балкона выбросили кирпич). Установить траекторию L брошенного тела.

Решение. Пренебрежем сопротивлением воздуха, что оправдано для массивных компактных тел, движущихся с не слишком большими скоростями. И будем

моделировать брошенное тело материальной точкой, что тоже вполне оправдано, если это тело небольшое.

Рассмотрим рис.1. Так как брошенная точка M участвует одновременно в двух независимых движениях – равномерном горизонтально со скоростью v и равноускоренном вертикально вниз с ускорением свободного падения g , то в любой момент времени $t \geq 0$ координаты $(x; y)$ этой точки найдутся по хорошо известным из школьной физики формулам:

$$\begin{cases} x = vt \\ y = h - s = h - \frac{gt^2}{2} \end{cases} \quad (t \geq 0) \quad (1.1)$$

Система уравнений (1.1) и описывает траекторию брошенной материальной точки. Она позволяет в любой заданный момент времени t найти координаты $(x; y)$ движущейся точки.

Полученную систему уравнений можно, при желании, привести и к одному уравнению вида $y = f(x)$, связывающему координаты $(x; y)$ движущейся точки M непосредственно друг с другом. Для этого из первого уравнения системы (1.1) выразим $t = \frac{x}{v}$ и подставим его во второе уравнение. В итоге получим явное уравнение траектории L :

$$y = h - \frac{gx^2}{2v^2}. \quad (1.2)$$

Положив в (1.2) $y = 0$ и $x = a$, найдем дальность полета a :

$$a = v \sqrt{\frac{2h}{g}}. \quad (1.3)$$

Уравнение (1.2) является уравнением параболы. Таким образом, построенная математическая модель траектории брошенного тела указывает на то, что эта траектория является параболой (1.2) (точнее, линией, близкой к этой параболе). А дальность полета может быть найдена по формуле (1.3). Как свидетельствует опытная проверка, для кирпича всё это действительно близко к истине. То есть построенная математическая модель достаточно хорошо (адекватно) описывает реальную траекторию брошенного кирпича. Но она уже недостаточно хорошо (недостаточно адекватно) описывает траекторию пули, ибо сопротивление воздуха её движению весьма велико, и пренебрежение им сильно искажает реальные результаты. Для пули эта модель требует существенного усложнения, связанного с необходимостью учитывать это сопротивление. И мы это усложнение проведем в следующем пункте.

Неадекватной, очевидно, эта модель будет и для бумажного самолетика, запущенного с балкона. Для самолетика построение адекватной математиче-

ской модели, очевидно, вообще сопряжено с огромными трудностями, если даже вообще возможно, ибо такая модель требует учета массы входных параметров, которыми нельзя пренебречь, причем многие из них являются случайными. В частности, любое случайное дуновение ветра может кардинально изменить траекторию движения самолетика. А вот для тяжелых самолетов, не реагирующих на случайные дуновения ветра и другие аналогичные факторы, существенные для бумажного самолетика, адекватные математические модели процессов их полета создать все же можно. И эти модели, хоть и трудным путем, с помощью и использованием серьёзных компьютерных технологий, всё же созданы. А это позволяет достаточно надежно использовать сложную авиационную технику и управлять ею.

1.2 Задача о траектории брошенного тела при учете сопротивления воздуха

А теперь снова рассмотрим задачу пункта 1, но теперь уже при учете сопротивления воздуха движущемуся телу. В частности, эта задача является актуальной для построения математической модели траектории летящей пули.

Итак, пусть стрелок, находясь на некоторой не слишком большой высоте h (на крыше или балконе дома, на невысоком холме) сделал винтовочный горизонтальный выстрел (рис. 1). Пуля массой m вылетела из винтовки со скоростью v . Как и выброшенный с балкона кирпич, она будет участвовать одновременно в двух движениях: горизонтальном с большой, но постепенно уменьшающейся со временем скоростью $v_x = v_x(t)$, и движении вертикально вниз с первоначально нулевой и затем возрастающей скоростью $v_y = v_y(t)$. Так как, по нашему предположению, высота h небольшая (несколько метров или пусть даже десятков метров), то пуля до своего падения на землю не успеет набрать вертикальную скорость, сравнимую с горизонтальной. А, значит, можно пренебречь сопротивлением воздуха вертикальному перемещению пули. И, следовательно, мы можем применить для координаты y летящей пули формулу (1.1):

$$y = y(t) = h - s = h - \frac{gt^2}{2} \quad (t \geq 0) \quad (1.4)$$

А вот для координаты x , определяющей движение пули в горизонтальном направлении, выражение должно быть отличным от того, которое давалось формулой (1.1), ибо теперь мы должны учесть сопротивление воздуха горизонтальному движению пули.

Как показывает опыт, силу сопротивления воздуха (а также жидкости) движущемуся в них телу можно приближенно считать пропорциональной скорости движения тела. А значит, согласно второму закону Ньютона,

$$mv'_x = -kv'_x, \text{ откуда } v'_x = -\frac{k}{m}v_x \quad (1.5)$$

Здесь k - коэффициент пропорциональности, зависящий от размеров (площади поперечного сечения) и формы тела – для шарообразной пули он больше, для цилиндрической с сужающейся головкой меньше, а также от свойств сопротивляющейся среды (её плотности, вязкости, и т.д.). Определяется этот коэффициент экспериментально. А знак (-) в (1.5) связан с тем, что сила сопротивления воздуха $R = -kv_x$ направлена против скорости движения тела.

Торможение воздухом летящего в нем тела в конечном итоге определяется ускорением торможения v'_x . А оно, согласно (1.5), определяется комбинацией $\frac{k}{m}v_x$. При одной и той же скорости v_x эта комбинация для тяжелой (большого калибра) пули будет меньше, чем для легкой. И еще меньше она будет для артиллерийского снаряда. Действительно, с увеличением размеров пули её площадь поперечного сечения (площадь сопротивления) увеличивается пропорционально квадрату её диаметра, а её объем, а, следовательно, и масса – пропорционально кубу диаметра. Таким образом, согласно (1.5), ускорение торможения оказывается обратно пропорциональным диаметру пули, то есть уменьшается с увеличением этого диаметра. Для артиллерийского снаряда, имеющего диаметр, в десятки раз превосходящий диаметр пули, ускорение торможения в эти же десятки раз будет меньше, чем у пули. И поэтому при построении математической модели полета артиллерийского снаряда имеется достаточно оснований сопротивлением воздуха его полету пренебречь. Тем более этим сопротивлением можно пренебречь для брошенного с небольшой скоростью кирпича или другого аналогичного компактного массивного тела, что мы и делали в пункте 1.

Но вернемся к пуле. Равенство (1.5) является дифференциальным уравнением первого порядка для неизвестной функции $v_x = v_x(t)$. Решим это уравнение:

$$\frac{dv_x}{dt} = -\frac{k}{m}v_x; \quad \frac{dv_x}{v_x} = -\frac{k}{m}dt; \quad \int \frac{dv_x}{v_x} = -\frac{k}{m} \int dt; \quad \ln v_x = -\frac{k}{m}t + C;$$

$$v_x = e^{-\frac{k}{m}t+C} = e^{-\frac{k}{m}t} \cdot e^C = Ce^{-\frac{k}{m}t}; \quad v_x = Ce^{-\frac{k}{m}t}$$

Здесь учтено, что если C – произвольная константа (константа интегрирования), то и e^C - произвольная константа, которую опять можно обозначить буквой C .

Учитывая начальное условие: $v_x(0) = v$, получим для горизонтальной скорости $v_x = v_x(t)$ окончательно:

$$v_x = ve^{-\frac{k}{m}t} \tag{1.6}$$

А теперь можем найти и горизонтальное перемещение пули $x = x(t)$. Учитывая, что $x' = v_x$, получим:

$$x = \int v_x dt = -v \frac{m}{k} e^{-\frac{k}{m}t} + C$$

А так как $x(0) = 0$, то $C = v \frac{m}{k}$. И с учетом этого получаем окончательно:

$$x = x(t) = v \frac{m}{k} \left(1 - e^{-\frac{k}{m}t} \right) \quad (t \geq 0) \quad (1.7)$$

Равенства (1.4) и (1.7) и определяют траекторию летящей пули при учете сопротивления воздуха её полету. А уравнения (1.1) определяют эту траекторию, если сопротивление воздуха не учитывать.

Сравним результаты, которые дают обе эти модели. Пусть $v = 800 \text{ м/сек}$ – начальная скорость пули (примерно такая она при выстреле из 7,62 мм винтовки, масса пули 11,75 г). Известно, что при настильной (близкой к горизонтальной) траектории из-за сопротивления воздуха скорость такой пули через секунду её полета уменьшается примерно вдвое, то есть до 400 м/сек. Найдем дальность полета пули, если она пущена в горизонтальный полет с высоты $h = 20$ метров, а также найдем вертикальную и горизонтальную скорости, с которыми она упадет на землю.

Согласно представленным выше данным о скорости пули, $v = 800$; $v_x(1) = 400$, откуда, на основании (1.6), имеем:

$$400 = 800 e^{-\frac{k}{m}} \Rightarrow e^{-\frac{k}{m}} = \frac{1}{2} \Rightarrow -\frac{k}{m} = \ln\left(\frac{1}{2}\right) = -\ln 2 \Rightarrow \frac{k}{m} = \ln 2 \approx 0,693$$

Подставляя $v = 800$ и найденное значение $\frac{k}{m}$ в выражения (1.6) и (1.7), получим:

$$v_x(t) = 800 \cdot \left(\frac{1}{2}\right)^t; \quad x(t) = 1154 \cdot \left(1 - \left(\frac{1}{2}\right)^t\right) \quad (1.8)$$

Теперь найдем время T полета пули. В момент времени $t = T$ пуля упадет на землю, и станет $y = 0$ (см. рис. 1). Полагая в (1.4) $y = 0$, $t = T$, и учитывая, что $h = 20$ метров, получим: $T = 2$ сек. Подставляя $t = T = 2$ в выражения (1.8), получим конечную горизонтальную скорость пули $v_x = 200 \text{ м/сек}$ и дальность полета пули $a \approx 865 \text{ м}$.

Если же воспользоваться формулами (1.1)-(1.3), при выведении которых сопротивление воздуха не учитывалось, то для этой же пули получим: $T = 2$ сек., конечная горизонтальная скорость пули $v_x = 800 \text{ м/сек}$ (она равна начальной скорости), и дальность полета пули $a = 1600 \text{ м}$. Вертикальная скорость па-

дения в обоих случаях примерно одна и та же и равна $v_y = y'(T) = gT = 20 \text{ м/сек}$.

Как видим, учет сопротивления воздуха существенно влияет на параметры траектории движения пули. И пренебрегать этим сопротивлением нельзя.

1.3 Задача об определении объема выкопанной ямы

Решение. Допустим, что осмотр ямы и предварительный её обмер позволяет считать её прямоугольным параллелепипедом (отклонения её формы от прямоугольного параллелепипеда для нас несущественны). Тогда в качестве математической модели нашей ямы принимаем прямоугольный параллелепипед. Остается измерить длину ямы x , ширину y и высоту z и затем по формуле $V = xyz$ вычислить её объем V .

Теперь оценим точность вычисления объема ямы. Допустим, что все измерения делались мерной лентой, которая дает результат измерения с точностью до сантиметра, то есть до 0,01 метра (а размеры земляной ямы и трудно получить точнее). Следовательно, при измерении размеров ямы x , y и z могли быть допущены их максимальные абсолютные погрешности $\Delta x = 0,01$, $\Delta y = 0,01$, $\Delta z = 0,01$ метра. А, следовательно, максимальные относительные погрешности размеров ямы будут таковы: $\frac{\Delta x}{x} = \frac{0,01}{x}$, $\frac{\Delta y}{y} = \frac{0,01}{y}$, $\frac{\Delta z}{z} = \frac{0,01}{z}$ (размеры ямы x , y и z здесь должны быть указаны в метрах). А так как при произведении величин их относительные погрешности, как известно, складываются, то максимальная относительная погрешность $\frac{\Delta V}{V}$ объема ямы $V = xyz$ будет такой:

$$\frac{\Delta V}{V} = \frac{\Delta x}{x} + \frac{\Delta y}{y} + \frac{\Delta z}{z} = 0,01 \left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z} \right) \quad (1.9)$$

Например, если оказалось, что $x = 0,5 \text{ м}$, $y = 2 \text{ м}$, $z = 2 \text{ м}$, то объем ямы $V = 2 \text{ м}^3$. Причем максимальная относительная погрешность этого результата $\frac{\Delta V}{V} = 0,03 = 3\%$. Если такая относительная погрешность вычисления объема ямы (в 3%) является допустимой, то поставленная задача решена. А принятая математическая модель ямы (прямоугольный параллелепипед) признается адекватной самому исследуемому объекту, то есть реальной яме. Если же нужна более высокая точность определения объема ямы (с меньшей допустимой погрешностью), то эту модель нужно усложнять: учитывать и отклонения формы ямы от прямоугольного параллелепипеда, и все размеры ямы измерять точнее.

1.4 Задача о баке с наименьшей площадью поверхности.

Требуется найти высоту h_0 и диаметр d_0 жестяного бака объемом $V = 30 \text{ м}^3$, имеющего форму закрытого кругового цилиндра, при которых площадь полной его поверхности S будет минимальной (в этом случае на его изготовление пойдет наименьшее количество жести).

Решение. Запишем известные формулы для объема V и площади полной поверхности S цилиндра высоты h и радиуса r :

$$V = \pi r^2 h; \quad S = \pi r^2 + \pi r^2 + 2\pi r h = 2\pi r^2 + 2\pi r h \quad (1.10)$$

Выражая h через r и V из первой формулы и подставляя полученное выражение во вторую, получим:

$$S = S(r) = 2\pi r^2 + \frac{2V}{r} \quad (1.11)$$

Таким образом, с математической точки зрения, задача сводится к определению такого значения r , при котором достигает своего минимума функция $S(r)$. Для этого найдем те значения r , при которых производная S' функции S обратится в нуль:

$$S' = 4\pi r - \frac{2V}{r^2}; \quad S' = 0 \Rightarrow 4\pi r - \frac{2V}{r^2} = 0; \Rightarrow r = \sqrt[3]{\frac{V}{2\pi}} \approx \sqrt[3]{\frac{30}{6,28}} \approx 1,684 \text{ м}. \quad (1.12)$$

Других значений r , при которых $S' = 0$, нет. Значит, найденное значение $r \approx 1,684 \text{ м}$ и есть то значение r_0 радиуса бака, при котором площадь его полной поверхности будет минимальной. Соответствующее ему значение высоты бака

$h_0 = \frac{V}{\pi r_0^2} = 2r_0 \approx 3,368 \text{ м}$. Округляя полученные значения до сантиметров, получим ответ: $d_0 = 2r_0 \approx 3 \text{ м } 37 \text{ см}$; $h_0 \approx 3 \text{ м } 37 \text{ см}$. Таким образом, наиболее экономным баком по расходу металла на его изготовление является бак, у которого диаметр совпадает с высотой. И это верно для бака любого объема.

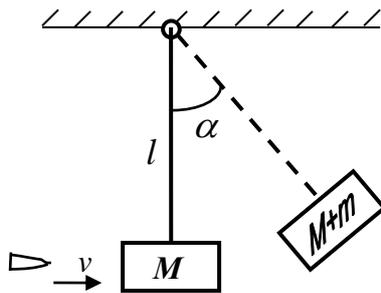


Рис.2

5. Задача об определении скорости револьверной пули

Решение. Эту задачу можно решить следующим простым способом. На легком жестком и свободно вращающемся стержне подвешивается массивный груз, в котором застревает пуля. Она передаст грузу свой импульс и свою кинетическую энергию. В результате стержень с грузом отклонится от вертикали на некоторый угол α

(рис. 2). Измерив этот угол, можно вычислить и скорость пули. Для этого к системе «пуля-стержень с грузом» нужно применить закон сохранения импульса и закон сохранения энергии.

Пренебрегая массой стержня по сравнению с массой груза, согласно первому из этих законов имеем:

$$mv = (m + M)V \quad (1.13)$$

А согласно второму

$$(m + M)\frac{V^2}{2} = (m + M)gl(1 - \cos \alpha) \quad (1.14)$$

В равенстве (1.14) слева стоит кинетическая энергия системы «пуля-груз», а справа - потенциальная энергия этой системы в положении её максимального отклонения от положения равновесия. Здесь m - масса пули, M - масса груза, V - скорость системы «пуля-груз» после вхождения пули в толщу груза, g - ускорение свободного падения, l - длина стержня, α - угол наибольшего отклонения стержня.

Из уравнений (1.13) - (1.14) легко находится скорость v пули:

$$v = \frac{m + M}{m} \sqrt{2gl(1 - \cos \alpha)} \quad (1.15)$$

Отметим, что точность результата (1.15), то есть адекватность этой математической модели будет тем выше, чем легче будет стержень по сравнению с грузом и чем точнее будут измерены все входящие в формулу (1.15) величины.

1.6. Строительная задача

Городской бюджет имеет возможность потратить на социальное жилье не более 600 млн. рублей, располагая при этом проектами и участками земли под 8 пятиэтажных домов на 90 квартир каждый и под 5 девятиэтажных домов на 120 квартир каждый. Средняя сметная стоимость одной квартиры в пятиэтажном доме составляет 600 тысяч рублей, а в девятиэтажном 800 тысяч рублей. Сколько пятиэтажных и сколько девятиэтажных домов должен построить город, чтобы получить максимальное число квартир?

Решение. Постройка одного пятиэтажного дома обойдется городскому бюджету в $90 \cdot 600000 = 54$ млн. руб., а постройка одного девятиэтажного в $120 \cdot 800000 = 96$ млн. руб. Так как квартиры в пятиэтажных домах стоят дешевле, то для городского бюджета выгоднее строить пятиэтажные дома – построенных квартир будет больше.

Вариант 1. Если город построит 8 пятиэтажных домов (максимально возможное их число), то он получит 720 квартир, затратив на это $54 \cdot 8 = 432$ млн. руб. На оставшиеся в бюджете 168 млн. руб. он может построить лишь один

девятиэтажный дом за 96 млн. руб. – на второй денег не хватит. Этот один построенный девятиэтажный дом добавит городу еще 120 квартир, итого их будет $720+120=840$.

Вариант 2. Если город построит 7 пятиэтажных домов (630 квартир), то это обойдется ему в $54 \cdot 7 = 378$ млн. руб. На оставшиеся 222 млн. руб. город может построить еще два девятиэтажных дома по 96 млн. руб. (на три не хватит), добавив еще 240 квартир. Итого их станет $630+240 = 870$. То есть больше, чем в первом варианте.

Вариант 3. Если город построит 6 пятиэтажных домов (540 квартир), то это обойдется ему в $54 \cdot 6 = 324$ млн. руб. Оставшихся 276 млн. руб. не хватит на три девятиэтажных дома по 96 млн. руб. – только на два, которые добавят еще 240 квартир. Итого их будет построено 780 – меньше, чем во втором варианте.

Завершая рассмотрение оставшихся вариантов, приходим к выводу: вариант 2 – лучший.

1.7. Транспортная задача

В городе имеются два склада муки и два хлебозавода. Ежедневно на хлебозаводы с первого склада вывозят 50 т муки, а со второго — 70 т, причем на первый хлебозавод завозят 40 т, а на второй 80 т. Обозначим через a_{ij} стоимость перевозки 1 т муки с i -го склада на j -й завод ($i, j = 1, 2$). Пусть $a_{11} = 12$ р., $a_{12} = 16$ р., $a_{21} = 8$ р., $a_{22} = 10$ р. Как нужно спланировать перевозки, чтобы их общая стоимость была минимальной?

Решение. Придадим задаче математическую формулировку. Обозначим через x_1 и x_2 количество муки, которое надо перевезти с первого склада на первый и второй заводы, а через y_1 и y_2 - со второго склада на первый и второй заводы соответственно. Тогда:

$$x_1 + x_2 = 50; \quad y_1 + y_2 = 70; \quad x_1 + y_1 = 40; \quad x_2 + y_2 = 80 \quad (1.16)$$

Общая стоимость F всех перевозок определяется формулой

$$F = 12x_1 + 16x_2 + 8y_1 + 10y_2 \quad (1.17)$$

С математической точки зрения задача заключается в том, чтобы найти четыре неотрицательных числа x_1, x_2, y_1 и y_2 , удовлетворяющих условиям (1.16) и дающих минимум функции F .

Систему уравнений (1.16) легко решить методом исключения неизвестных. При этом получим (убедитесь в этом самостоятельно):

$$x_2 = 50 - x_1; \quad y_1 = 40 - x_1; \quad y_2 = 30 + x_1, \quad (1.18)$$

а x_1 не может быть определено однозначно из этой системы: x_1 остается произвольным. Но неотрицательность всех величин, фигурирующих в равенствах

(1.18), приводит к условию: $0 \leq x_1 \leq 40$. Подставляя выражения для x_2, y_1, y_2 из (1.18) в формулу для F , получим

$$F = 1420 - 2x_1.$$

Очевидно, что минимум этой функции достигается при максимально возможном значении x_1 , то есть при $x_1 = 40$. Соответствующие значения других неизвестных определяются по формулам (1.18): $x_2 = 10, y_1 = 0, y_2 = 70$.

1.8. Задача об остывании вскипевшего чайника

Вскипевший чайник остывает в комнате с температурой воздуха 20°C . Известно, что через 20 минут после снятия чайника с плиты его температура снижается с исходных 100°C до 60°C . Через какое время с начала остывания чайника его температура опустится до 30°C ? И через какое время чайник остынет полностью (то есть его температура сравняется с температурой воздуха)?

Решение. Пусть t (мин) – время, прошедшее с момента времени $t = 0$, когда начался процесс остывания чайника. И пусть $T = T(t)$ – функция, описывающая изменение температуры T чайника со временем. Ответы на поставленные вопросы мы получим, если найдем эту функцию.

Согласно условиям задачи, $T(0) = 100$; $T(20) = 60$; $T_0 = 20$ – температура окружающего чайник воздуха (указание на градусы Цельсия опускаем). Между чайником и воздухом идет теплообмен. Согласно закону теплообмена Ньютона, скорость остывания тела в окружающей его среде пропорциональна разности температур тела и окружающей среды. Так как скорость остывания нашего чайника – это скорость изменения его температуры, то есть скорость изменения функции $T = T(t)$, а скорость изменения любой функции – это её производная, то скорость остывания чайника равна $T' = T'(t)$. А $T - T_0 = T - 20$ – это разность температур чайника и воздуха. Следовательно, по закону Ньютона получаем: $T' = k(T - 20)$. Здесь k – коэффициент пропорциональности, который пока неизвестен. Собирая все выписанные выше равенства в систему, получим:

$$\begin{cases} T' = k(T - 20) \\ T(0) = 100 \\ T(20) = 60 \end{cases} \quad (1.19)$$

Первое из этих равенств – это дифференциальное уравнение первого порядка для неизвестной функции $T = T(t)$. Второе равенство представляет собой начальное условие для этого уравнения. А третье равенство – дополнительное условие для неизвестной функции $T = T(t)$. То есть система уравнений (1.19)

представляет собой задачу Коши для дифференциального уравнения первого порядка с дополнительным условием для неизвестной функции $T = T(t)$.

Решим эту задачу.

1. Сначала решим (проинтегрируем) дифференциальное уравнение:

$$T' = k(T - 20); \quad \frac{dT}{dt} = k(T - 20); \quad \frac{dT}{T - 20} = k dt; \quad \int \frac{dT}{T - 20} = k \int dt;$$

$$\ln(T - 20) = kt + C; \quad T - 20 = e^{kt+C} = e^C \cdot e^{kt} = Ce^{kt}; \quad T = 20 + Ce^{kt} \quad (1.20)$$

Полученное выражение $T = 20 + Ce^{kt}$, где C - произвольная константа интегрирования, представляет собой общее решение дифференциального уравнения $T' = k(T - 20)$. В этом общем решении содержится всё бесчисленное количество его частных решений.

2. Реализуем теперь начальное условие $T(0) = 100$ и с его помощью найдем константу C , содержащуюся в общем решении:

$$T(0) = 100; \quad 100 = 20 + Ce^{k \cdot 0} = 20 + C; \quad C = 80; \quad T = 20 + 80e^{kt} \quad (1.21)$$

Осталось найти константу k . Её найдем из последнего, дополнительного условия (1.14):

$$T(20) = 60; \quad 60 = 20 + 80e^{20k}; \quad e^{20k} = \frac{1}{2} \quad (1.22)$$

Используя этот результат в последнем равенстве (1.21), получим окончательно:

$$T = 20 + 80(e^{20k})^{\frac{t}{20}} = 20 + 80 \cdot \left(\frac{1}{2}\right)^{\frac{t}{20}}; \quad T = T(t) = 20 + 80 \cdot \left(\frac{1}{2}\right)^{\frac{t}{20}} \quad (1.23)$$

Найденная функция $T = T(t)$ и описывает изменение температуры чайника со временем. Её график показан на рис.3.

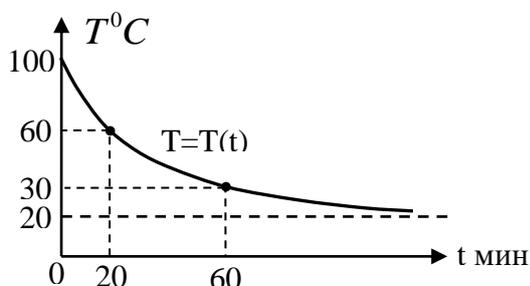


Рис. 3

А теперь с помощью формулы (1.23) можем ответить и на поставленные в задаче вопросы.

1) $T = 30; \quad t = ?$

Решение.

$$30 = 20 + 80 \cdot \left(\frac{1}{2}\right)^{\frac{t}{20}};$$

$$80 \cdot \left(\frac{1}{2}\right)^{\frac{t}{20}} = 10; \quad \left(\frac{1}{2}\right)^{\frac{t}{20}} = \left(\frac{1}{2}\right)^3; \quad \frac{t}{20} = 3;$$

$t = 60$ (мин). То есть до 30 градусов чайник остынет через 1 час.

2) $T = 20$; $t = ?$

Решение.

$$20 = 20 + 80 \cdot \left(\frac{1}{2}\right)^{\frac{t}{20}}; \quad \left(\frac{1}{2}\right)^{\frac{t}{20}} = 0; \quad t - \text{ не существует.}$$

То есть до 20 градусов (до температуры воздуха) чайник никогда не остынет. С течением времени его температура, снижаясь, будет лишь приближаться к комнатной температуре, но никогда её не достигнет. Это наглядно демонстрирует график изменения температуры чайника со временем (рис. 3). И это не противоречит нашему ощущению, что через достаточно большое время (через 2-3 часа) чайник кажется уже полностью остывшим. Просто спустя это время температура чайника настолько приблизится к температуре воздуха, что разница в их температурах нами уже не ощущается.

1.9 Задача о моторной лодке

Моторная лодка движется в спокойной воде со скоростью 20 км/час. На полном ходу её мотор выключается, и через 40 секунд после этого скорость лодки уменьшается до 8 км/час. Считая, что сопротивление воды пропорционально скорости движения лодки, определить скорость лодки через 2 минуты после остановки мотора.

Решение. После выключения мотора на лодку действует единственная сила – сила сопротивления воды F , ибо две другие силы, вес лодки и выталкивающая сила Архимеда, уравновешивают друг друга, так что их результирующая равна нулю. Сила F , по условию, пропорциональна скорости v движения лодки. То есть $F = kv$. Имея в виду, что $a = v' = \frac{dv}{dt}$ – ускорение лодки, по второму закону Ньютона $ma = F$ получаем:

$$mv' = kv \tag{1.24}$$

Это – дифференциальное уравнение первого порядка относительно функции $v = v(t)$ – скорости лодки. Решая его, получаем:

$$m \frac{dv}{dt} = kv; \quad \frac{dv}{v} = \frac{k}{m} dt; \quad \int \frac{dv}{v} = \frac{k}{m} \int dt; \quad \ln v = \frac{k}{m} t + C; \quad v = e^{\frac{k}{m}t + C} = e^{\frac{k}{m}t} \cdot e^C = Ce^{\frac{k}{m}t}$$

Здесь учтено, что если C – произвольная константа (константа интегрирования), то и e^C – произвольная константа, которую опять можно обозначить буквой C .

Полученное выражение

$$v = Ce^{\frac{k}{m}t} \quad (1.25)$$

представляет собой общее решение дифференциального уравнения (1.24). Оно содержит в себе все частные решения $v = v(t)$ этого уравнения (их бесчисленное количество). Среди них находится и та функция $v = v(t)$, которая описывает скорость нашей лодки. Но чтобы найти эту функцию, нужно в выражении (1.25) найти константу C и комбинацию констант $\frac{k}{m}$. Для этого воспользуемся

остальными условиями задачи.

- 1) По условию, в момент времени $t = 0$ скорость лодки была $v = 20$ км/час. Подставляя эти данные в формулу (1.25), получим: $C = 20$. Таким образом,

$$v = 20e^{\frac{k}{m}t} \quad (1.26)$$

- 2) В момент времени $t = 40$ сек. $= \frac{2}{3}$ мин. скорость лодки была 8 км/час. Подставляя эти данные в формулу (1.26), получим:

$$8 = 20e^{\frac{k \cdot 2}{m \cdot 3}}, \text{ откуда } e^{\frac{2k}{3m}} = \frac{8}{20} = \frac{2}{5} \quad (1.27)$$

С учетом (1.27) выражение (1.26) примет следующий окончательный вид:

$$v = 20 \left(e^{\frac{2k}{3m}} \right)^{\frac{3}{2}t} = 20 \cdot \left(\frac{2}{5} \right)^{\frac{3}{2}t}; \quad v = v(t) = 20 \cdot \left(\frac{2}{5} \right)^{\frac{3}{2}t} \quad (1.28)$$

При $t = 2$ мин. получаем:

$$v = 20 \cdot \left(\frac{2}{5} \right)^3 = \frac{32}{25} \approx 1,3 \text{ км/час}$$

1.10 Задача о барометрической формуле

Барометрическая формула – это формула, по которой определяется атмосферное давление воздуха в зависимости от высоты над уровнем моря. Выведем эту формулу.

Решение. Пусть $p = p(h)$ - давление воздуха на высоте h над уровнем моря. Функцию $p(h)$ и требуется установить.

Рассмотрим на высоте h горизонтальную площадку единичной площади $S_0 = 1$. Так как, по своему физическому смыслу, давление – это сила, приходящаяся на единицу площади, то атмосферное давление $p = p(h)$ - это сила дав-

ления воздуха на эту площадку. А эта сила, очевидно, просто равна весу столба воздуха над этой площадкой.

Наряд с указанной выше площадкой, рассмотрим еще одну площадку единичной площади $S_0 = 1$, расположенную строго над первой площадкой на расстоянии dh от неё (dh бесконечно мало). Давление на эту площадку будет уже не $p(h)$, а $p(h + dh) = p(h) + dp$, где $dp < 0$ - изменение давления при переходе с нижней площадки на верхнюю ($dp < 0$, так как атмосферное давление с высотой убывает). Величина dp , взятая со знаком минус, равна, очевидно, весу воздуха между указанными площадками. Если $\rho = \rho(h)$ - плотность воздуха на высоте h , то учитывая, что объем воздуха между площадками равен $dV = S_0 \cdot dh = 1 \cdot dh = dh$, получим вес $\rho(h)dV \cdot g = g\rho(h)dh$ этого воздуха, то есть $-dp$:

$$dp = -g\rho(h)dh \quad (1.29)$$

Здесь g - ускорение свободного падения, которое будем считать постоянным (его незначительным уменьшением с увеличением высоты h пренебрежем).

Далее, атмосферный воздух будем считать идеальным газом (как известно, воздух, особенно разреженный, с достаточной для практики точностью можно считать идеальным газом). Уравнение состояния идеального газа имеет вид:

$$pV = \frac{m}{\mu} RT \quad (1.30)$$

Здесь m масса газа; V - объем, занимаемый газом; p - давление в газе; T - абсолютная температура газа (по Кельвину); μ - молярная масса газа (для воздуха $\mu = 0,029 \frac{\text{кг}}{\text{моль}}$); $R = 8,31 \frac{\text{дж}}{\text{моль} \cdot \text{К}}$ - универсальная газовая постоянная.

Из уравнения (1.30) можно выразить плотность газа ρ :

$$\rho = \frac{m}{V} = \frac{\mu p}{RT} \quad (1.31)$$

Подставив (1.31) в (1.29), получим

$$dp = -\frac{\mu g}{RT} p dh \Rightarrow \frac{dp}{p} = -\frac{\mu g}{RT} dh \Rightarrow \int \frac{dp}{p} = -\int \frac{\mu g}{RT} dh \quad (1.32)$$

Пренебрежем изменением с высотой абсолютной температуры воздуха. То есть будем считать, что воздух на разных высотах имеет некую одну и ту же усредненную абсолютную температуру T . Для температуры по Цельсию это является большой натяжкой – как известно, на большой высоте заметно холоднее, чем у поверхности Земли, особенно летом. У поверхности Земли может быть

+30⁰С, а на десятикилометровой высоте, на которой летают самолеты, может быть –50⁰С. Но для абсолютной температуры T по Кельвину эта разница уже не столь существенна: 303⁰К и 223⁰К соответственно. Тем более эта разница будет еще менее существенной, если мы ограничимся не слишком большими высотами над поверхностью земли. Для таких высот абсолютную температуру T атмосферы действительно приближенно можно считать постоянной и равной её среднему значению.

Тогда в правой части интеграла (1.32) всё, кроме dh - константы, которые можно вынести за знак интеграла. В результате из равенства (1.32) получим:

$$\ln p = -\frac{\mu g}{RT}h + C \Rightarrow p = Ce^{-\frac{\mu g}{RT}h} \quad (1.33)$$

Если p_0 - давление воздуха на высоте моря (при $h = 0$), то $C = p_0$, и формула (1.33) примет следующий окончательный вид:

$$p = p(h) = p_0 e^{-\frac{\mu g}{RT}h} \quad (1.34)$$

Это и есть искомая барометрическая формула. По ней можно вычислять атмосферное давление $p = p(h)$ на разных высотах h . И для не слишком больших высот она дает результаты, близкие к истинным.

1.11 Задача об определении силы давления воды на плотину

Плотина имеет форму равнобокой трапеции с верхним основанием $2a$ (м), нижним основанием $2b$ (м), высотой h (м) (рис.4). Требуется получить формулу, выражающую силу давления воды на эту плотину.

Решение. Пусть F - искомая сила давления. Если бы давление воды на плотину в любой её точке было одним и тем же и равным p , то силу F мы нашли бы по простой и очевидной формуле: $F = pS$, где S - площадь плотины. Но давление воды с глубиной меняется. Из курса физики известно, что на глубине x давление воды $p = p(x)$ определяется формулой:

$$p = p(x) = \rho g x \quad (1.35)$$

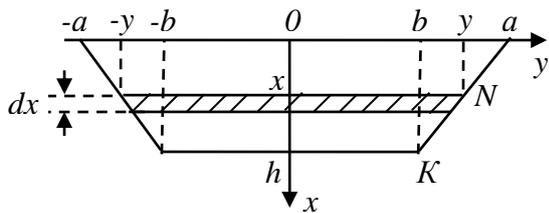


Рис.4

Здесь ρ - плотность воды ($\rho = 10^3 \text{ кг/м}^3$), а $g = 9,8 \text{ м/сек}^2$ - ускорение свободного падения. В связи с неравномерностью давления воды на плотину есть смысл при подсчете силы давления F воды на плотину мысленно разбить поверхность плоти-

ны на узкие (бесконечно узкие) горизонтальные полоски – так, чтобы все точки каждой полоски находились на одной глубине, то есть при одинаковом давлении воды. После этого найти силу давления воды на каждую полоску, а затем просуммировать эти силы по всем полоскам. То есть использовать идею вычисления силы F через определенный интеграл.

Реализуем эту идею. Введем на поверхности плотины систему координат xoy так, как показано на рис. 4, и выделим на поверхности одну из таких полосок (она заштрихована). Верхний край полоски находится на глубине x , а её ширина dx бесконечно мала (полоска бесконечно узкая). Площадь полоски $dS = 2y \cdot dx$. Давление воды на глубине x равно $p = \rho gx$. И в силу того, что наша полоска бесконечно узкая, таким будет давление в любой точке полоски. То есть вся полоска находится под равномерным давлением воды $p = \rho gx$. А тогда сила давления dF найдется умножением этого давления на площадь полоски:

$$dF = p(x) \cdot dS = \rho gx \cdot 2y dx = 2\rho gxy dx \quad (1.36)$$

Найденная сила dF является бесконечно малой частью (дифференциалом) искомой силы F . Заметим, что в формулу (1.36), кроме x , входит переменная y , которую можно (и нужно!) выразить через переменную x - ту переменную, чей дифференциал dx входит в выражение для dF . Для этого используем подобие прямоугольных треугольников, вершины которых отмечены знаками $\{K; b; a\}$ и $\{N; y; a\}$. Пропорциональность их катетов дает:

$$\frac{h}{a-b} = \frac{x}{a-y}, \text{ откуда } y = a - \frac{(a-b)x}{h} \quad (1.37)$$

Подставляя y из (1.37) в (1.36), получим для dF окончательно:

$$dF = 2\rho gx \left(a - \frac{(a-b)x}{h} \right) dx \quad (1.38)$$

А теперь просуммируем силы dF по всем полоскам, на которые мы мысленно разбивали поверхность плотины – от верхней полоски, где $x = 0$, до нижней, где $x = h$, и получим искомую силу F давления воды на всю плотину:

$$F = \sum dF = \sum 2\rho gx \left(a - \frac{(a-b)x}{h} \right) dx = \int_0^h 2\rho gx \left(a - \frac{(a-b)x}{h} \right) dx \quad (1.39)$$

Вычислим определенный интеграл (1.39) с использованием формулы Ньютона-Лейбница:

$$\begin{aligned}
F &= 2\rho g \left(a \int_0^h x dx - \frac{(a-b)}{h} \int_0^h x^2 dx \right) = 2\rho g \left(a \frac{x^2}{2} \Big|_0^h - \frac{(a-b)}{h} \frac{x^3}{3} \Big|_0^h \right) = \\
&= 2\rho g \left(a \frac{h^2}{2} - \frac{(a-b)h^2}{3} \right) = \rho g h^2 \left(\frac{a+2b}{3} \right); \quad F = \rho g h^2 \left(\frac{a+2b}{3} \right)
\end{aligned} \tag{1.40}$$

Формула для вычисления суммарной силы F давления воды на трапециевидную плотину получена.

А теперь, в дополнение к решенной задаче, найдем центр давления найденной силы. То есть найдем на поверхности плотины точку приложения такой сосредоточенной силы, которая будет являться равнодействующей рассредоточенных по поверхности плотины сил давления воды на плотину.

По величине эта равнодействующая равна, естественно, F . А точка её приложения должна быть такой, чтобы момент вращения, создаваемый этой сосредоточенной силой относительно любой оси, проходящей по плоскости плотины, совпадал с моментом вращения относительно этой же оси рассредоточенных по поверхности плотины сил давления воды на плотину. В частности, момент вращения этих рассредоточенных сил относительно любой оси, проходящей через центр давления, должен равняться нулю, ибо он равен нулю и для равнодействующей этой силы (плечо у этой равнодействующей отсутствует). То есть рассредоточенные по поверхности плотины силы давления воды должны быть самоуравновешены относительно любой оси, проходящей через центр давления.

Ясно, что в силу симметрии плотины центр давления должен находиться в некоторой точке x_0 на оси ox (рис. 4), ибо относительно этой вертикальной оси силы давления воды на плотину очевидным образом самоуравновешены. Но они должны быть самоуравновешены и относительно горизонтальной оси, проходящей через точку x_0 . То есть суммарный момент M всех сил dF давления воды на рассмотренные выше горизонтальные полоски относительно горизонтальной линии, проходящей через точку x_0 , должен равняться нулю.

В частности, относительно указанной горизонтальной линии момент dM силы dF , действующей на полоску, изображенную на рис.4, равен $dM = dF \cdot (x - x_0)$ (равен произведению силы на плечо этой силы). А полный момент всех сил dF относительно этой линии равен

$$M = \sum dM = \sum (x - x_0) dF = \int_0^h (x - x_0) \cdot 2\rho g x \left(a - \frac{(a-b)x}{h} \right) dx \tag{1.41}$$

Приравнивая этот момент нулю, после несложных выкладок, которые опускаем (проделайте их самостоятельно), получим:

$$x_0 = \frac{(a + 3b)h}{2(a + 2b)} \quad (1.42)$$

В частности, если $a = b$ (плотина прямоугольная), то $x_0 = \frac{2}{3}h$. На такой глубине находится центр давления воды на прямоугольную плотину.

1.12 Закон Архимеда. Плавание тел в жидкостях

Закон Архимеда о телах, погруженных в жидкость, принадлежит к числу самых первых из установленных наукой физических законов. Он установлен Архимедом еще до нашей эры, и звучит так:

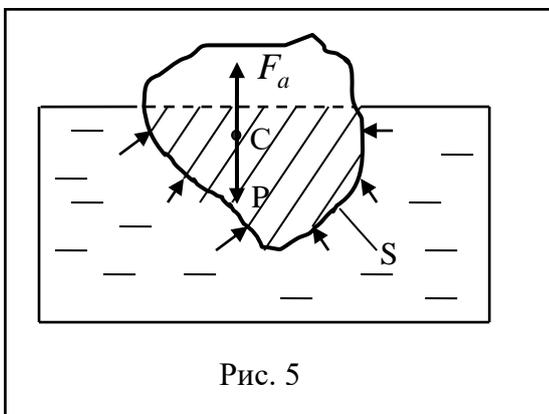


Рис. 5

на тело, погруженное в жидкость, действует выталкивающая сила, равная весу жидкости, вытесненной телом. Эта сила направлена вертикально вверх и приложена в центре тяжести вытесняемой телом жидкости.

Для вывода этого закона рассмотрим рис.5. Здесь изображено плавающее в жидкости тело с поверхностью S , погруженной в жидкость. Жидкость, окружающая тело, давит на эту поверхность. Суммарная сила

этого давления и есть сила Архимеда F_a .

Давление жидкости на поверхность S погруженного в него тела уравновешивается точно таким же давлением поверхности тела на окружающую её жидкость с той же суммарной силой F_a , только противоположно направленной. Это давление, очевидно, не изменится, если заменить погруженную в жидкость часть тела такой же жидкостью. Жидкость, заключенная внутри поверхности, давит на жидкость вне поверхности своим весом. Он приложен в центре тяжести объема жидкости внутри поверхности. А так как вся жидкость находится в равновесии, то точно такой же будет и сила, действующая со стороны внешней жидкости на внутреннюю. А это и есть сила Архимеда. То есть она действительно равна весу жидкости, вытесняемой погруженным в неё телом, приложена к центру тяжести области, ограниченной поверхностью S , и направлена строго вверх – против веса вытесняемой жидкости.

Если V - объем погруженной в жидкость части тела, а плотность жидкости ρ , то сила Архимеда F_a равна

$$F_a = \rho g V \quad (1.43)$$

Если P - вес всего тела, то для плавающего в жидкости или на его поверхности тела $F_a = P$. Максимальной будет сила Архимеда при полном погружении тела в жидкость: $F_{\max} = \rho g V_0$, где V_0 - объем всего тела. И если вес тела $P \leq F_{\max}$, то тело будет плавать. А если $P > F_{\max}$ - тело потонет.

Пример 1 (историческая задача Архимеда).

Слиток сплава двух металлов (золота и серебра) с плотностями ρ_1 и ρ_2 в воздухе имеет вес P , а в воде P_* . Найти вес каждого из металлов в слитке.

Решение. На слиток, погруженный в воду, действует сила Архимеда $F_a = P - P_*$. С другой стороны, если V - объем слитка, то $F_a = \rho g V$, где ρ - плотность воды. Следовательно,

$$\rho g V = P - P_* \Rightarrow V = \frac{P - P_*}{\rho g}$$

Далее, пусть V_1 и V_2 - объемы, занимаемые каждым из металлов в слитке. И пусть P_1 и P_2 - веса металлов (в воздухе). Тогда

$$V_1 = \frac{P_1}{\rho_1 g}; \quad V_2 = \frac{P_2}{\rho_2 g}$$

А так как $V_1 + V_2 = V$, то получаем:

$$\frac{P_1}{\rho_1 g} + \frac{P_2}{\rho_2 g} = \frac{P - P_*}{\rho g}$$

С другой стороны, $P_1 + P_2 = P$. Решая систему из последних двух уравнений, получаем:

$$P_1 = \frac{(P - P_*)\rho_1\rho_2 - P\rho\rho_1}{\rho(\rho_2 - \rho_1)}; \quad P_2 = \frac{P\rho\rho_2 - (P - P_*)\rho_1\rho_2}{\rho(\rho_2 - \rho_1)}.$$

Пример 2. Резиновый мяч с массой m и радиусом R погружают в воду на глубину h и отпускают. Найти скорость, с которой мяч вылетит из воды, и высоту, на которую он поднимется в воздух.

Решение. На погруженный в воду мяч действуют две силы: его вес $P = mg$ (направлен вниз), и архимедова сила $F_a = \rho g V$, где ρ - плотность воды, а $V = \frac{4}{3}\pi R^3$ - объем мяча. То есть $F_a = \frac{4}{3}\pi\rho g R^3$. Результирующая сила F , действующая на мяч, равна

$$F = F_a - P = \frac{4}{3}\pi\rho g R^3 - mg$$

Эта сила будет поднимать мяч к поверхности. Так как мяч, поднимаясь, пройдет до поверхности воды расстояние h , то сила F на этом пути совершит работу

$$A = Fh = \left(\frac{4}{3}\pi\rho gR^3 - mg\right)h.$$

Эта работа на выходе мяча на поверхность будет равна приобретенной им кинетической энергии $E_k = \frac{mv^2}{2}$, где v - вертикальная скорость, которую на выходе приобретет мяч. Таким образом

$$E_k = A; \quad \frac{mv^2}{2} = \left(\frac{4}{3}\pi\rho gR^3 - mg\right)h; \quad v = \sqrt{\frac{\left(\frac{4}{3}\pi\rho gR^3 - mg\right) \cdot 2h}{m}}$$

Подлетев на максимальную высоту y , мяч приобретет потенциальную энергию $E_p = mgy$, равную E_k . Которая, в свою очередь, равна работе A . Итак,

$$E_p = A; \quad mgy = \left(\frac{4}{3}\pi\rho gR^3 - mg\right)h; \quad y = \frac{\left(\frac{4}{3}\pi\rho R^3 - m\right)h}{m}.$$

1.13 Установившееся движение жидкости. Уравнение Бернулли

Различают два вида движения жидкостей: а) установившееся; б) неустановившееся.

Установившимся называют такое движение, при котором давление в жидкости и скорость её течения в любой точке не меняются со временем. Например, установившимся является течение воды по трубе и истечение её из крана при постоянном напоре воды; движение воды в канале или реке при постоянном уровне воды, и т.д. А истечение воды из сосуда при его опорожнении будет движением неустановившимся.

При установившемся движении частицы жидкости описывают не меняющиеся со временем криволинейные траектории, называемые линиями тока. Они между собой не пересекаются – в противном случае частица, попавшая в точку пересечения двух линий тока, должна была бы в дальнейшем двигаться одновременно в двух разных направлениях. А при неустановившемся движении линии тока частиц могут пересекаться. Вектор скорости частицы, движущейся по своей линии тока, в любой её точке направлен по касательной к линии тока.

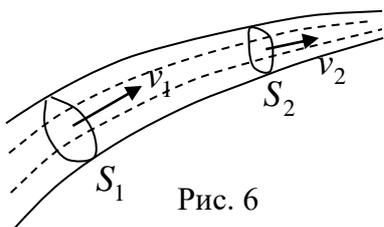


Рис. 6

Если в движущейся жидкости в некоторый момент времени выделить некоторый

компактный объем жидкости, то в дальнейшем он образует трубку тока (рис. б): Так как жидкость считается несжимаемой, не разрывается и не переходит через стенки трубки тока, то за одно и то же время dt через каждое сечение трубки проходит один и тот же объем жидкости:

$$V_1 = V_2 \Rightarrow S_1 v_1 dt = S_2 v_2 dt \Rightarrow S_1 v_1 = S_2 v_2 \quad (1.44)$$

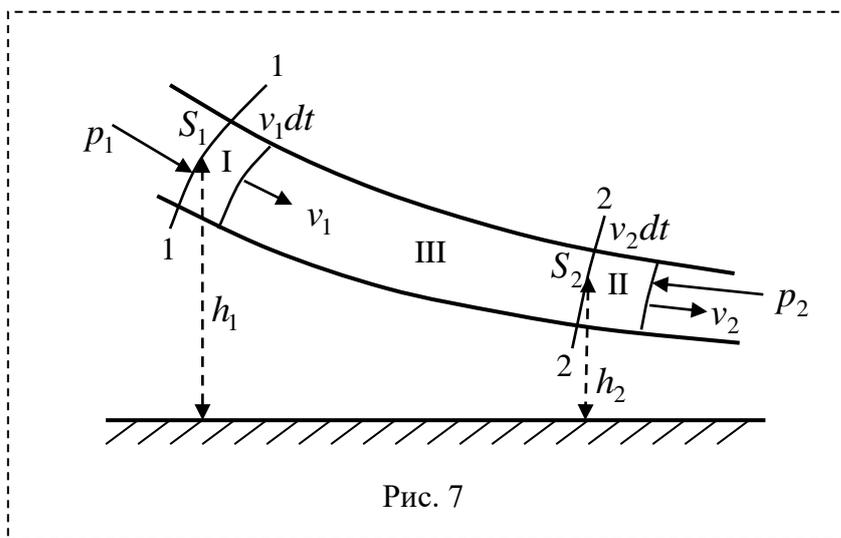
Последнее равенство называется уравнением неразрывности движения жидкости. Оно справедливо не только для трубки тока, но и для всякой реальной трубы, по которой течет жидкость, для русла реки, и т.д. Причем в этих случаях оно справедливо и при неустановившемся движении жидкости. Согласно уравнению (1.44), в более широких сечениях трубки тока скорости частиц жидкости меньше, а в более узких – больше. И это вполне естественно.

Кстати, произведение Sv представляет собой объем жидкости, протекающей через сечение S трубки тока за единицу времени, и называется он *расходом* текущей жидкости. Согласно (1.44), расход жидкости один и тот же в любом сечении трубки тока.

Уравнение Бернулли для идеальной жидкости

Уравнение Бернулли является одним из основных уравнений механики движения жидкостей, имеющем большое прикладное значение. Выведем его для установившегося движения идеальной жидкости – то есть несжимаемой и не имеющей вязкости (не имеющей внутреннего трения между своими слоями). Оно очень важно для гидравлики. В гидравлике в качестве жидкости в основном рассматривается вода. А её свойства близки свойствам идеальной жидкости.

Пусть по наклонной неподвижной трубке тока (или реальной трубе) переменного сечения движется жидкость (рис. 7).



Мысленно выделим область трубки, ограниченную сечениями 1-1 и 2-2 с площадями сечений S_1 и S_2 соответственно, в которых скорости течения жидкости v_1 и v_2 , а высоты центров тяжести сечений относительно некоторой горизонтальной поверхности отсчета равны h_1 и h_2 .

Через бесконечно малое время dt сечение S_1 переместится вдоль трубки на расстояние $v_1 dt$, а сечение S_2 на расстояние $v_2 dt$. И если в исходный момент времени t выделенный объем жидкости занимал области I и III, то через время dt он будет занимать области III и II.

Как известно из физики, изменение кинетической энергии любого движущегося тела равно работе внешних сил, действующих на это тело.

Подсчитаем сначала изменение кинетической энергии выделенной нами жидкости. При переходе из начального положения (жидкость занимала объемы отсеков I и III) в последующий (отсеки II и III) при установившемся движении кинетическая энергия общего отсека III останется неизменной, а потому приращение кинетической энергии произойдет лишь за счет изменения кинетической энергии отсеков I и II:

$$\Delta E_{кин} = E_{II+III} - E_{I+III} = E_{II} - E_I = \frac{m_2 v_2^2}{2} - \frac{m_1 v_1^2}{2} \quad (1.45)$$

Здесь m_1 и m_2 - массы жидкости в отсеках I и II соответственно:

$$m_1 = \rho V_1 = \rho S_1 v_1 dt; \quad m_2 = \rho V_2 = \rho S_2 v_2 dt; \quad (1.46)$$

Учитывая уравнение неразрывности (1.44) и вводя обозначение

$$S_1 v_1 = S_2 v_2 = q, \quad (1.47)$$

получаем из (1.46):

$$m_1 = m_2 = m = \rho q dt \quad (1.48)$$

И тогда выражение (1.45) примет вид:

$$\Delta E_{кин} = \left(\frac{v_2^2}{2} - \frac{v_1^2}{2} \right) \rho q dt \quad (1.49)$$

А теперь подсчитаем работу всех внешних сил, действующих на выделенный объем жидкости. Внешними силами для него являются:

1. Поверхностные силы нормального давления со стороны жидкости, окружающей выделенный объем;
2. Силы трения, действующие по боковой поверхности трубки тока;

3. Силы тяжести.

Из сил давления работу могут совершать только силы, действующие на торцевые сечения S_1 и S_2 . Это силы

$$P_1 = p_1 S_1 \quad \text{и} \quad P_2 = p_2 S_2 \quad (1.50)$$

Здесь p_1 и p_2 - давление в жидкости в сечениях S_1 и S_2 соответственно. Силы давления, действующие на боковую поверхность трубки, работы не совершают, так как направлены перпендикулярно течению жидкости. Поэтому работа сил давления, согласно рисунку 7, равна:

$$A_{\text{давл}} = P_1 v_1 dt - P_2 v_2 dt = p_1 S_1 v_1 dt - p_2 S_2 v_2 dt = (p_1 - p_2) q dt \quad (1.51)$$

Что касается работы сил трения по боковой поверхности трубки тока, то она отсутствует, так как силы трения в идеальной жидкости отсутствуют.

Осталось рассмотреть работу силы тяжести. Она состоит в том, что отсек I переместится в отсек II, так как отсек III остается на месте. Массы и, следовательно, веса отсеков I и II одинаковы. Поэтому работа сил тяжести равна убыли потенциальной энергии массы $m = \rho q dt$ отсеков при изменении высоты h , на которой находится эта масса над горизонтальной поверхностью, со значения h_1 до значения h_2 . То есть

$$A_{\text{тяж}} = mgh_1 - mgh_2 = (h_1 - h_2)mg = (h_1 - h_2)\rho g q dt \quad (1.52)$$

А теперь приравняем изменение кинетической энергии (1.49) сумме работ всех внешних сил, приложенных к трубке тока:

$$\Delta E_{\text{кин}} = A_{\text{давл}} + A_{\text{тяж}} \Rightarrow \left(\frac{v_2^2}{2} - \frac{v_1^2}{2} \right) \rho q dt = (p_1 - p_2) q dt + (h_1 - h_2) \rho g q dt \quad (1.53)$$

Сокращая обе части этого уравнения на $q dt$, после группировки слагаемых получим:

$$\frac{\rho v_1^2}{2} + \rho g h_1 + p_1 = \frac{\rho v_2^2}{2} + \rho g h_2 + p_2 \quad (1.54)$$

Поскольку сечения S_1 и S_2 в трубке тока выбирались произвольно, то получаем следующий вывод: в любом сечении выделенной трубки тока идеальной жидкости при установившемся её течении выполняется равенство:

$$\frac{\rho v^2}{2} + \rho g h + p = \text{Const} \quad (1.55)$$

Это и есть уравнение Даниила Бернулли, выведенное им в 18 веке.

Все три слагаемых в уравнении Бернулли имеют размерность давления.

При этом $\frac{\rho v^2}{2}$ называют динамическим давлением; ρgh - гидравлическим давлением; p - статическим давлением. А в сумме они составляют полное (гидродинамическое) давление.

Для горизонтальной трубки тока (или реальной горизонтальной трубы) уравнение Бернулли принимает вид:

$$\frac{\rho v^2}{2} + p = p_0 = Const, \quad (1.56)$$

ибо в этом случае $\rho gh = Const$. Из этого уравнения, в частности, следует, что при горизонтальном течении жидкости по трубе с переменным сечением в узких местах, где скорость больше, статическое давление, передающееся на стенки трубы, меньше. А всё выражение p_0 будет полным давлением в трубе. Например, человек, стоящий в ручье или в реке лицом против течения, испытывает на свою переднюю часть полное давление воды p_0 . А на свои бока – лишь статическое.

И полное, и статическое давление в движущейся жидкости можно измерить с помощью так называемой скоростной трубки – обычной открытой стеклянной или пластмассовой трубки с изогнутым под прямым углом концом. Трубку погружают в поток жидкости в двух положениях: 1) когда изогнутое колено ориентировано вдоль течения жидкости так, что жидкость вливается в это колено, и 2) в положении, когда колено повернуто перпендикулярно течению жидкости. В первом случае уровень поднявшейся в трубку жидкости будет характеризовать полное давление p_0 , а во втором – более низкое статическое давление p . По разности высот Δh этих уровней определится динамическое давление $\frac{\rho v^2}{2}$:

$$\frac{\rho v^2}{2} = p_0 - p = \rho g \Delta h \quad \Rightarrow \quad v = \sqrt{2g\Delta h} \quad (1.57)$$

То есть по разности уровней Δh жидкости в скоростной трубке можно вычислить скорость течения жидкости. Если же скорость v потока жидкости измерена прямым путем (по скорости мелких легких предметов, плывущих по поверхности потока), то по ней можно найти динамическое давление $\frac{\rho v^2}{2}$. А тогда из равенства (1.57) можно найти и превышение $p_0 - p$ полного давления в жидкости над статическим.

С помощью уравнения Бернулли можно вывести формулу для скорости истечения идеальной жидкости из малого отверстия в стенке или дне сосуда.

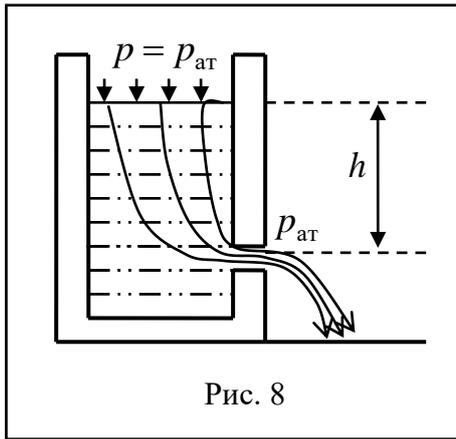


Рис. 8

Выведем эту формулу. Для этого рассмотрим рис.8. Если рассматривать процесс вытекания жидкости в течение малого (бесконечно малого) промежутка времени dt , то уровень h жидкости над отверстием понизится лишь на бесконечно малую величину dh , то есть фактически не изменится. Значит, в течение времени dt процесс истечения жидкости из сосуда можно считать стационарным, а значит, к этому сосуду можно применять уравнение Бернулли (1.55). На поверхности сосуда $v = 0$; $p = p_{ат}$; $h = h$. У отвер-

стия $v \neq 0$; $p = p_{ат}$; $h = 0$ (высоту h отсчитываем от отверстия до уровня жидкости в сосуде). Тогда из уравнения Бернулли (1.55) следует:

$$\rho gh + p_{ат} = \frac{\rho v^2}{2} + p_{ат} \Rightarrow v = \sqrt{2gh} \quad (1.58)$$

Формула (1.58) носит название формулы Торричелли (итальянский физик 17-го века). Согласно формуле Торричелли, скорость истечения идеальной жидкости из малого отверстия в сосуде совпадает со скоростью v свободно падающего с высоты h тела, если пренебречь сопротивлением воздуха.

Впрочем, для реальных жидкостей формула (1.58) даёт несколько завышенные значения скорости истечения жидкостей из отверстий и требует поправки:

$$v = \lambda \sqrt{2gh}, \text{ где } 0 < \lambda < 1 \quad (1.59)$$

Поправочный множитель λ связан с тем, что реальные жидкости не идеальны – они имеют вязкость, что тормозит частицы вытекающей жидкости и уменьшает скорость их истечения из отверстия. Играет свою роль и форма отверстия, и толщина стенок сосуда, а также такой эффект, как сжатие поперечного сечения струи сразу после выхода её из отверстия. Имеются и полуэмпирические формулы для подсчета величины λ . Но эти вопросы мы здесь обсуждать не будем, они исследуются в специальной дисциплине - гидравлике.

Пример 1. Из опрыскивателя плодовых деревьев выбрасывается струя жидкости. Давление в баке опрыскивателя составляет два атмосферных давления. С какой скоростью выбрасывается струя жидкости из опрыскивателя?

Решение. Используем уравнение Бернулли в форме (1.54)

$$\frac{\rho v_1^2}{2} + \rho gh_1 + p_1 = \frac{\rho v_2^2}{2} + \rho gh_2 + p_2 \quad (1.60)$$

и применим его к шлангу опрыскивателя. Здесь $\{v_1; h_1; p_1\}$ - параметры жидкости (воды) на его входе, а $\{v_2; h_2; p_2\}$ - на выходе. Перепишем уравнение (1.60) в виде

$$\frac{\rho v_2^2}{2} = \frac{\rho v_1^2}{2} + \rho g(h_1 - h_2) + p_1 - p_2 \quad (1.61)$$

Оценим все слагаемые правой части этого равенства. Учтем, что $p_2 = p_{\text{ат}} = 10^5$ паскалей, так как струя выбрасывается в атмосферу; $p_1 = 2p_{\text{ат}} + \rho gh$, где h - высота жидкости в бачке опрыскивателя. Так как эта высота составляет 0,2-0,3, максимум 0,4 метра, то $\rho gh < 1000 \cdot 10 \cdot 0,4 = 4 \cdot 10^3$ па = 0,04 ат, и этим слагаемым в выражении для p_1 можно пренебречь. То есть $p_1 \approx 2p_{\text{ат}} = 2 \cdot 10^5$ па. И тогда $p_1 - p_2 \approx 1 \text{ атм} = 10^5$ па. Слагаемым $\rho g(h_1 - h_2)$ в (1.61) тоже, очевидно, можно пренебречь, так как $h_1 \approx h_2$. Наконец, пренебрежем и величиной $\frac{\rho v_1^2}{2}$ по сравнению с $\frac{\rho v_2^2}{2}$, так как скорость v_1 движения воды в шланге много меньше скорости v_2 истечения воды из конца шланга. В итоге равенство (1.61) примет вид

$$\frac{\rho v_2^2}{2} \approx p_1 - p_2 = 10^5 \text{ па, откуда } v_2 \approx \sqrt{\frac{2 \cdot 10^5}{\rho}} = \sqrt{\frac{2 \cdot 10^5}{10^3}} = \sqrt{200} \approx 14 \text{ м/сек.}$$

Пример 2. Цилиндрический бак с площадью основания $S = 1 \text{ м}^2$ и высотой $H = 3 \text{ м}$ заполнен водой. Пренебрегая вязкостью воды, определить время t , необходимое для опорожнения бака, если на дне бака открыть отверстие площадью $S_1 = 10 \text{ см}^2 = 0,001 \text{ м}^2$.

Решение. Согласно формуле Торричелли, скорость v истечения воды из отверстия равна $v = \sqrt{2gh}$, где h - высота уровня воды в баке над отверстием. Эта высота меняется от своего первоначального значения H до конечного значения $h = 0$, а значит меняется (уменьшается) и скорость истечения воды из бака.

Пусть t - время, прошедшее с начала опорожнения бака, а h - уровень жидкости в баке в этот момент. Рассмотрим, что произойдет за бесконечно малое время dt , следующее за моментом времени t . За это время из бака вытечет объем жидкости $dV = S_1 v dt = S_1 \sqrt{2gh} dt$. С другой стороны, за это время высота h воды в баке изменится на величину $dh < 0$ ($dh < 0$, так как высота h с течением времени уменьшается), поэтому $dV = S \cdot (-dh)$. Приравнявая друг другу эти два выражения для dV , получим

$$S_1 \sqrt{2gh} dt = -S dh, \text{ откуда } dt = -\frac{S}{S_1} \cdot \frac{dh}{\sqrt{2gh}}$$

Интегрируя обе части этого равенства, получим: $t = -\frac{S}{S_1 g} (\sqrt{2gh} + C)$. Константу интегрирования C найдем из условия, что $h = H$ при $t = 0$: $C = -\sqrt{2gH}$. В итоге получаем следующую зависимость между временем t и уровнем h воды в баке:

$$t = \frac{S}{S_1 g} (\sqrt{2gH} - \sqrt{2gh})$$

Бак опустеет, когда станет $h = 0$. Это произойдет через время

$$t_* = \frac{S}{S_1 g} \sqrt{2gH} = \frac{1}{0,001 \cdot 9,8} \sqrt{2 \cdot 9,8 \cdot 3} \approx 775 \text{ сек} \approx 13 \text{ мин.}$$

Упражнения

1. С поверхности земли под углом α к горизонту с начальной скоростью v брошено материальное тело. При каком значении угла α дальность полета тела будет максимальной? Сопротивлением воздуха пренебречь.

Ответ: $\alpha = 45^\circ$.

2. С поверхности земли произведен винтовочный выстрел с начальной скоростью пули 800 м/сек. Через 1 секунду пуля упала на землю, имея скорость 400 м/сек. Найти дальность полета пули.

Ответ: ≈ 577 м.

3. При измерении размеров кругового цилиндра (его радиуса основания R и его высоты h) получены значения $R = 40$ см и $h = 80$ см, причем эти значения получены с максимально возможной ошибкой 5 мм (в ту или в другую сторону). Найти максимально возможные абсолютную ΔV и относительную $\frac{\Delta V}{V}$ погрешности при вычислении объема V цилиндра.

Ответ: $V \approx 0,402 \text{ м}^3$; $\Delta V \approx 0,012 \text{ м}^3$; $\frac{\Delta V}{V} \approx 0,03 = 3\%$.

4. Из квадратного листа картона со стороной a вырезаются по углам одинаковые квадраты, а из оставшейся части склеивается прямоугольная коробка. Какова должна быть сторона вырезаемого квадрата, чтобы объем коробки был максимальным?

Ответ: $\frac{1}{6} a$.

5. Вблизи завода A проходит железная дорога к городу B . Под каким углом α к железной дороге нужно провести шоссе, соединяющее завод A с железной дорогой, чтобы доставка продукции до города B была максимально дешевой, если стоимость 1 тонны/километра при перевозке по шоссе в n раз дороже, чем по железной дороге?

Ответ: $\cos \alpha = \frac{1}{n}$.

6. Груз, лежащий на горизонтальной поверхности, нужно сдвинуть с места. Под каким углом α к горизонту нужно направить сдвигающую силу, чтобы её величина была максимальной? Коэффициент трения покоя между грузом и плоскостью равен k .

Ответ: $\operatorname{tg} \alpha = k$.

7. Дождевая капля, начальная масса которой равна m_0 , а начальная скорость равна нулю, падает под действием силы тяжести (сопротивлением воздуха пренебрегаем) так, что убыль её массы пропорциональна времени (коэффициент пропорциональности равен k). Через какое время после начала падения кинетическая энергия капли будет максимальной, и какова она?

Ответ: $t = \frac{2m_0}{3k}$; $E_{\max} = \frac{2g^2 m_0^3}{27k^2}$.

8. Через небольшую пробоину в трюме судна бьет струя воды. Один человек не в состоянии преодолеть силу струи, чтобы закрыть отверстие доской. Однако когда с помощью товарища доска была наложена, человек оказался в состоянии удержать доску. Почему?

Ответ: При наложенной на отверстие доске сила давления воды на неё уменьшилась вдвое.

9. Труба, лежащая горизонтально, поперечным значением которой является круг радиуса R , наполовину наполнена водой. Найти: а) силу F , с которой вода давит на вертикальную заслонку, закрывающую трубу; б) центр давления этой силы.

Ответ: а) $F = 2\rho g \int_0^R \sqrt{R^2 - x^2} dx = \frac{2\rho g R^3}{3}$; б) центр давления лежит на

вертикальной оси симметрии заслонки в точке, отстоящей от поверхности воды на величину $\frac{3\pi}{16} R$.

10. Из трубы с площадью сечения S_1 бьет вертикально вверх струя воды. Расход воды (объем воды, вытекающей из трубы за единицу времени) равен Q . Найти: а) скорость v_1 истечения воды из трубы; б) максимальную высоту h_{\max}

подъема воды; в) площадь сечения S_2 струи на высоте h над отверстием трубы ($h < h_{\max}$).

$$\text{Ответ: а) } v_1 = \frac{Q}{S_1}; \text{ б) } h_{\max} = \frac{Q^2}{2gS_1^2}; \text{ в) } S_2 = \frac{QS_1}{\sqrt{Q^2 - 2ghS_1^2}}.$$

§2. Стохастическое моделирование. Регрессионный анализ

Как уже указывалось выше, стохастическое моделирование – это вид математического моделирования, основывающийся на экспериментальных данных наблюдений за объектом и на методах теории вероятности и математической статистики. То есть это моделирование объекта, рассматривающее его как «черный ящик», часть параметров которого является входными (их задают на входе в ящик), часть выходными (их измеряют на выходе), а оставшиеся (неконтролируемые) параметры считаются помехами. Одним из основных видов стохастического моделирования является *регрессионный анализ*.

В регрессионном анализе изучается *корреляционная зависимость* (зависимость «в среднем») между числовыми величинами, характеризующими изучаемый объект или процесс.

2.1. Корреляционная связь между величинами

Рассмотрим суть корреляционной зависимости (связи) между числовыми величинами на примере *парной корреляции*, то есть на примере корреляционной связи между некоторыми двумя числовыми величинами x и y . Относительно этой связи имеется, в принципе, три возможности.

1) Первая возможность: величины x и y независимы друг от друга. Это значит, что каждая из этих величин принимает свои значения независимо от значений, принимаемых другой величиной.

2) Вторая возможность - обратная первой: величины x и y связаны жесткой (функциональной) зависимостью, т.е. зависимостью вида $y = f(x)$. В этом случае каждому возможному значению величины x соответствует вполне определенное значение $y = f(x)$ величины y . То есть возможные значения величины y жестко привязаны к возможным значениям величины x . Например, объем куба $y = x^3$ жестко (функционально) зависит от его стороны x .

3) Третья возможность - промежуточная между первыми двумя: x и y в принципе связаны между собой (независимыми они не являются), но эта связь не жесткая (размытая). Это значит, что каждому возможному значению величины x могут соответствовать различные значения ($y_1; y_2; \dots$) величины y . Причём набор этих значений и (или) вероятности появления этих значений меняются с изменением значения x . Такого рода связь между величинами называются *статистической* (или *вероятностной*) связью. Статистическая связь

между величинами x и y означает, что изменение значения одной из них ведет к изменению внешних условий для реализации другой величины. Например, меняющаяся среднесуточная температура статистически влияет на плотность сельскохозяйственных вредителей на засеянном поле; объем денежной массы у покупателей статистически влияет на объем закупаемых ими товаров, и т.д.

Если при статистической связи между величинами x и y при изменении значения величины x еще и меняется *среднее значение* \bar{y}_x величины y , то говорят, что y *корреляционно* (в среднем) зависит от x . Аналогично понимается корреляционная зависимость x от y . В частности, очевидно, что между температурой x воздуха и количеством y вредителей имеет место не просто статистическая, а корреляционная зависимость, ибо с изменением температуры изменяется и среднее количество сельскохозяйственных вредителей. Аналогично между количеством x денег у покупателей и их тратами y на покупку товаров тоже имеется, очевидно, корреляционная зависимость, ибо чем больше денег у покупателей, тем больше в среднем они покупают. Корреляционно (в среднем) связаны также урожайность различных культур с количеством внесенных под них удобрений; продуктивность животных с количеством и качеством их кормов; производительность труда рабочих с их квалификацией, и т.д.

Величина y может корреляционно (в среднем) зависеть сразу от нескольких величин (x_1, x_2, \dots, x_n) . Тогда говорят не о парной, а о *множественной* корреляции.

2.2 Парная корреляция при наличии повторных опытов

Рассмотрим сначала парную корреляцию. Пусть экспериментальным путем исследовалась зависимость одного выходного (искомого) числового параметра y исследуемого объекта от одного входного числового параметра x . И пусть результаты этого исследования представлены в таблице 2.1:

Таблица 2.1

$x_i \backslash y_j$	x_1	x_2	...	x_n	$m_j = \sum_{i=1}^n n_{ij}$
y_1	n_{11}	n_{21}	...	n_{n1}	m_1
y_2	n_{12}	n_{22}	...	n_{n2}	m_2
...
y_m	n_{1m}	n_{2m}	...	n_{nm}	m_m
$n_i = \sum_{j=1}^m n_{ij}$	n_1	n_2	...	n_n	$\sum_{i=1}^n n_i = \sum_{j=1}^m m_j = N$

(2.1)

Эта таблица называется *корреляционной*. Из нее видно, что объект исследовался при значениях величины $x=x_i = (x_1 ; x_2 ; \dots ; x_n)$, а величина y принимала при

этом значения $y=y_j = (y_1 ; y_2 ; \dots ; y_m)$. При каждом $x=x_i$ проводились повторные измерения с целью определения степени влияния на величину y других, неконтролируемых, величин (помех).

Согласно таблице 2.1, пара значений $(x_1; y_1)$ встретилась в измерениях n_{11} раз, пара $(x_2; y_1) - n_{21}$ раз, и т.д. Числа $n_i = (n_{i1}; n_{i2}; \dots ; n_{im})$ определяют общее количество измерений со значениями x , равными $(x_1; x_2; \dots ; x_n)$ соответственно, а числа $(m_1; m_2; \dots ; m_m)$ – общее количество измерений со значениями $(y_1; y_2; \dots ; y_m)$ величины y соответственно. При этом ясно, что

$$n_i = \sum_{j=1}^m n_{ij}; \quad m_j = \sum_{i=1}^n n_{ij}; \quad \sum_{i=1}^n n_i = \sum_{j=1}^m m_j = \sum_{i=1}^n \sum_{j=1}^m n_{ij} = N, \quad (2.2)$$

где N - общее число опытов.

Данные корреляционной таблицы 2.1 для наглядности удобно изобразить и в виде так называемого *корреляционного поля*. Корреляционное поле – это нанесенное на плоскость $хоу$ множество всех N экспериментальных точек с координатами $(x_i; y_j)$ с учетом их кратности n_{ij} . Это значит, что при построении корреляционного поля нужно показать, что в точке $(x_i; y_j)$ плоскости $хоу$ содержится не одна, а n_{ij} точек. Чтобы это было видно на корреляционном поле, нужно эти точки немного отделить друг от друга. Они тогда образуют видимую компактную кучку из n_{ij} точек, окружающих точку $(x_i; y_j)$ (рис.2.1).

В регрессионном анализе ставятся и решаются **две основные задачи**:

1) построение уравнений, описывающих усредненные зависимости выходных параметров объекта от входных (эти уравнения называются *уравнениями регрессии*);

2) Определение степени адекватности этих уравнений опытными данными. В рассматриваемом сейчас случае парной корреляции первой из этих двух задач является построение уравнения регрессии $\bar{y}_x = f(x)$, описывающего зависимость среднего значения \bar{y}_x величины y от величины x . А второй задачей является выяснение вопроса о степени адекватности этого уравнения опытными данными, представленным в таблице 2.1.

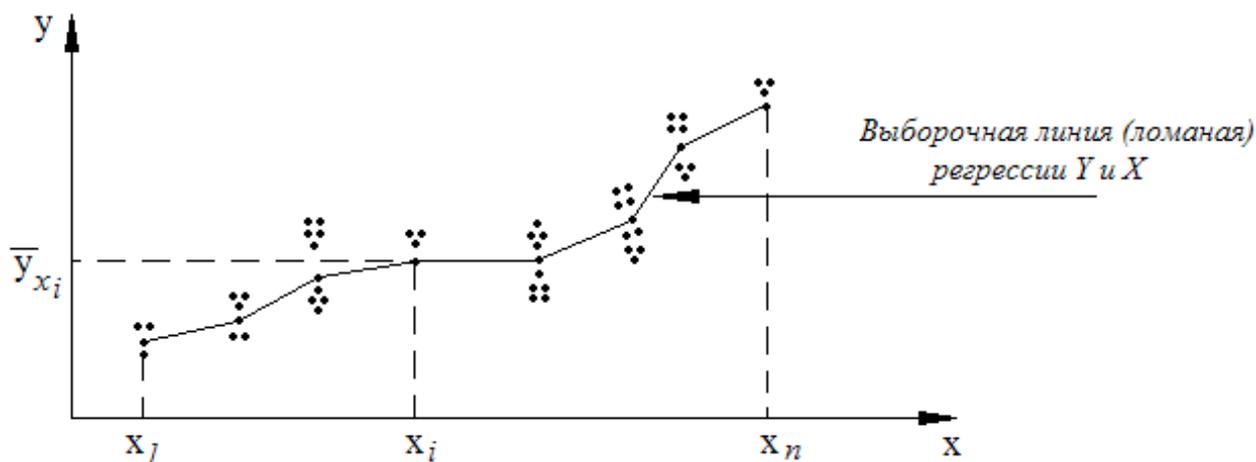


Рис. 2.1

График уравнения регрессии $\bar{y}_x = f(x)$ называется *линией регрессии*. Примерный (оценочный) вид этой линии мы получим, если по данным корреляционной таблицы (2.1) для каждого экспериментального значения x_i величины x найдем среднее значение \bar{y}_{x_i} величины y

$$\bar{y}_{x_i} = \frac{\sum_{j=1}^m y_j n_{ij}}{n_i} \quad (i = 1, 2, \dots, n), \quad (2.3)$$

и затем нанесем на корреляционное поле ломаную, соединяющую точки с координатами $(x_i; \bar{y}_{x_i})$ ($i = 1, 2, \dots, n$) – рис. 2.1. Эта ломаная называется *экспериментальной линией регрессии*.

Корреляционное поле с нанесенной на него экспериментальной линией регрессии показывает, как *в проведенном эксперименте* в среднем (корреляционно) меняются значения величины y при изменении величины x .

Таким образом, построив корреляционное поле с нанесенной на него экспериментальной линией регрессии, мы можем визуально оценить и *силу*, и *тесноту* корреляционной зависимости y от x . Если экспериментальная линия регрессии окажется горизонтальной прямой или линией, близкой к такой прямой, то это будет свидетельствовать об отсутствии или о слабости корреляционной зависимости y от x (\bar{y}_x слабо зависит от x). А чем более крутой будет эта линия, тем сильнее будет указанная зависимость. А степень разбросанности точек корреляционного поля вокруг экспериментальной линии регрессии указывает на *степень тесноты* этой корреляционной зависимости: чем меньше разброс точек корреляционного поля вокруг экспериментальной линии регрессии, тем эта зависимость теснее.

Корреляционную зависимость y от x описывает, как мы уже говорили, уравнение регрессии $\bar{y}_x = f(x)$. А в качестве такого уравнения принимается уравнение

$$\bar{y}_x^* = f^*(x) \quad (2.4)$$

линии, наилучшим образом приближающей (сглаживающей) экспериментальную линию регрессии. Это уравнение стараются построить так, чтобы оно удовлетворяло *двум требованиям*:

- а) чтобы оно было как можно проще;
- б) чтобы линия, описываемая этим уравнением, располагалась возможно ближе к экспериментальной линии регрессии.

Отметим, однако, что два указанных выше требования, вообще говоря, противоречивы, ибо повысить указанную близость сглаживающей линии регрессии к экспериментальной можно лишь за счет усложнения уравнения $\bar{y}_x^* = f^*(x)$ сглаживающей линии. Поэтому на практике стараются добиться некой золотой середины: и чтобы сглаживающее уравнение регрессии (2.4) было не слишком сложным, и чтобы соответствующая ему сглаживающая ли-

ния регрессии, тем не менее, в целом была достаточно близкой к экспериментальной ломаной регрессии. Как из многих возможных вариантов выбрать лучший (найти эту золотую середину) – об этом будет сказано ниже.

В качестве наиболее простых форм сглаживающего уравнения чаще всего принимаются следующие его формы:

1) Уравнение прямой

$$\bar{y}_x^* = kx + b \quad (2.5)$$

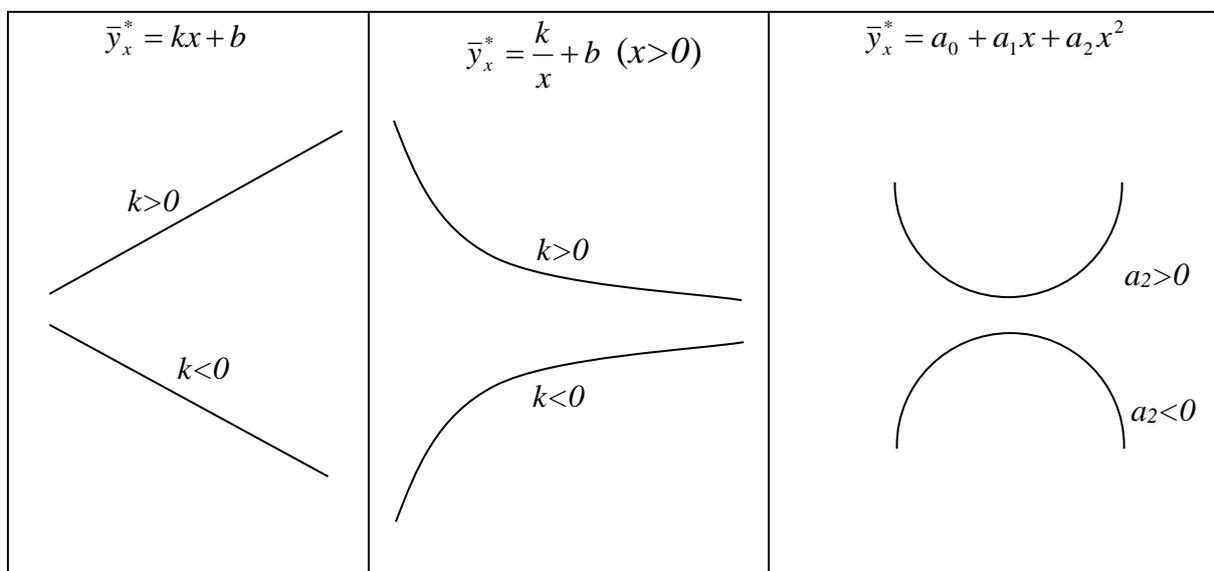
2) Уравнение гиперболы

$$\bar{y}_x^* = \frac{k}{x} + b \quad (2.6)$$

3) Уравнение параболы

$$\bar{y}_x^* = a_0 + a_1x + a_2x^2 \quad (2.7)$$

Напомним, что эти линии в принципе имеют следующий вид:



Линейная зависимость (2.5) наиболее проста по форме, ее параметры k и b легко интерпретируются. В частности, коэффициент k указывает, на сколько в среднем увеличится (при $k > 0$) или уменьшится (при $k < 0$) величина y , если значение x увеличится на единицу. А параметр b указывает среднее значение величины y при $x = 0$. Благодаря этим преимуществам, а также благодаря простоте вычисления параметров k и b уравнение (2.5) используется в качестве приближенного (сглаживающего) уравнения регрессии даже в тех случаях, когда более логичным представляется использование уравнение кривой.

Гиперболы (2.6) – это либо монотонно возрастающая, либо монотонно убывающая кривая. Однако, в отличие от прямой, рост или убывание гиперболы имеет тенденцию к затуханию, практически сходя на нет при больших значениях x . Параметр b при этом представляет собой предельное значение \bar{y}_x^* при $x \rightarrow \infty$. Убывающей гиперболой, например, хорошо выражается зависимость себестоимости продукции растениеводства от урожайности, а возрастающей – зависимость продуктивности животных от расхода кормов.

Парабола (2.7) имеет вершину или впадину и применяется для приближенного описания зависимостей, в которых с изменением x убывание \bar{y}_x^* может меняться на возрастание, и наоборот. Примером параболической зависимости является, например, зависимость средней урожайности от дозы удобрений, когда начальные дозы удобрений приводят к значительному увеличению урожая, последующие – к постепенно уменьшающимся прибавкам, а чрезмерные – к снижению урожая и даже к его гибели. Иногда график параболы используется только частично (только восходящая или только нисходящая ветвь). Параметры параболы, за исключением a_0 , интерпретировать сложно. Ну, а параметр a_0 является, очевидно, оценкой среднего значения величины y при $x = 0$.

Уравнения регрессии (2.5)–(2.7) линейны по их числовым параметрам. Построение и анализ линейных по параметрам уравнений регрессии называется *линейным регрессионным анализом*. Он наиболее прост по своей реализации.

В некоторых случаях и нелинейные по параметрам уравнения регрессии с помощью специальных линеаризирующих преобразований могут быть преобразованы в линейные. Рассмотрим несколько примеров.

1. Нелинейная по параметру b функция $\bar{y}_x^* = f^*(x) = ax^b$ при $x > 0$ с помощью логарифмирования преобразуется к виду: $\ln \bar{y}_x^* = \ln a + b \ln x$. Производя замену $\ln x = t$; $\ln a = a_*$; $\ln \bar{y}_x^* = \bar{z}_t^*$, получим линейное по параметрам уравнение регрессии $\bar{z}_t^* = a_* + bt$.

2. Функция $\bar{y}_x^* = \frac{ax}{b+x}$ линеаризуется так:

$$b+x = a \frac{x}{\bar{y}_x^*}, \text{ или } \frac{x}{\bar{y}_x^*} = \frac{b}{a} + \frac{1}{a}x.$$

После замены

$$\frac{x}{\bar{y}_x^*} = \bar{z}_x^*; \quad \frac{b}{a} = a_*; \quad \frac{1}{a} = a_*$$

получим линейную по параметрам функцию $\bar{z}_x^* = a_* + b_*x$.

3. Так называемая *логистическая функция* $\bar{y}_x^* = \frac{e^{a+bx}}{1+e^{a+bx}}$ при помощи преобразования $\bar{z}_x^* = \ln \frac{\bar{y}_x^*}{1-\bar{y}_x^*}$ примет линейный по параметрам вид: $\bar{z}_x^* = a + bx$.

Подходящую форму сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$ выбирают, исходя из общих теоретических соображений или, что чаще, по виду корреляционного поля (рис. 2.1). При этом, как уже говорилось выше, наиболее часто уравнение регрессии выбирается в одной из форм (2.5) – (2.7). А для

нахождения параметров выбранного уравнения используется универсальный стандартный метод, называемый *методом наименьших квадратов*.

Суть этого метода в следующем. Из корреляционной таблицы (2.1) для каждого x_i по формуле (2.3) находим экспериментальное среднее значение \bar{y}_{x_i} величины y . Далее, для выбранной формы уравнения регрессии записываем сглаживающие средние $\bar{y}_{x_i}^* = f^*(x_i)$. В итоге получаем следующую таблицу соответствий экспериментальных \bar{y}_{x_i} и сглаживающих $\bar{y}_{x_i}^*$ средних значений величины y , соответствующих величине x , принимающей значения x_i с частотами n_i ($i = 1, 2, \dots, n$):

Таблица 2.2

x_i	x_1	x_2	...	x_n
n_i	n_1	n_2	...	n_n
\bar{y}_{x_i}	\bar{y}_{x_1}	\bar{y}_{x_2}	...	\bar{y}_{x_n}
$\bar{y}_{x_i}^* = f^*(x_i)$	$f^*(x_1)$	$f^*(x_2)$...	$f^*(x_n)$

(2.8)

Параметры выбранного сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$ считаются наилучшими, если они обеспечивают минимально возможные отклонения экспериментальных средних \bar{y}_{x_i} от подсчитанных по уравнению регрессии сглаживающих средних $\bar{y}_{x_i}^* = f^*(x_i)$. В методе наименьших квадратов за меру отклонений \bar{y}_{x_i} от $\bar{y}_{x_i}^* = f^*(x_i)$ принимается сумма квадратов их разностей. При этом должно быть учтено, что в образовании каждого \bar{y}_{x_i} участвуют n_i точек корреляционного поля. То есть в средней \bar{y}_{x_i} как бы сливаются n_i значений величины y . С учетом сказанного указанная сумма принимает вид:

$$Q = \sum_{i=1}^n (\bar{y}_{x_i} - \bar{y}_{x_i}^*)^2 n_i = \sum_{i=1}^n [\bar{y}_{x_i} - f^*(x_i)]^2 n_i \rightarrow \min \quad (2.9)$$

В эту сумму входят параметры выбранной функции $f^*(x)$. И эти параметры подбираются таким образом, чтобы сумма Q была минимально возможной. А это – стандартная задача математического анализа об исследовании функции нескольких переменных на экстремум (минимум или максимум), где Q – функция, а ее переменные – параметры функции $f^*(x)$. Решая эту задачу, находим наилучшие параметры функции $f^*(x)$, а вместе с ними получаем и искомое наилучшее (для выбранной формы) сглаживающее уравнение регрессии (2.4), являющееся оценкой искомого уравнения регрессии $\bar{y}_x = f(x)$.

Заметим что и экспериментальные средние \bar{y}_{x_i} , и сглаживающие их средние $\bar{y}_{x_i}^*$ имеют одно и то же среднее значение – общую среднюю \bar{y} величины y . То есть

$$\frac{\sum_{i=1}^n \bar{y}_{x_i} \cdot n_i}{N} = \bar{y}; \quad \frac{\sum_{i=1}^n \bar{y}_{x_i}^* \cdot n_i}{N} = \bar{y} \quad (2.10)$$

Первое из этих равенств следует из выражения (2.3):

$$\frac{\sum_{i=1}^n \bar{y}_{x_i} \cdot n_i}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m y_j n_{ij}}{N} = \frac{\sum_{j=1}^m y_j \sum_{i=1}^n n_{ij}}{N} = \frac{\sum_{j=1}^m y_j m_j}{N} = \bar{y}$$

А второе равенство получим, рассудив следующим образом. Так как сглаживающая кривая $\bar{y}_x^* = f^*(x)$ получена методом наименьших квадратов в процессе минимизации суммы (2.9), то она наилучшим способом вписывается в выборочную линию регрессии, то есть имеет от нее в целом минимальное отклонение. Поэтому если эту сглаживающую линию поднять или опустить, то есть если заменить её на линию $\bar{y}_x^* = f^*(x) + C$, где C – некоторая константа, то взамен функции (2.9) получим функцию

$$Q = Q(C) = \sum_{i=1}^n (\bar{y}_{x_i} - f^*(x_i) - C)^2 n_i,$$

значения которой при всех C больше величины (2.9). А свое наименьшее значение (экстремум) функция $Q(C)$ должна иметь при $C = 0$. Но это значит, что

$$Q' = Q'(C) = -2 \sum_{i=1}^n (\bar{y}_{x_i} - f^*(x_i) - C) n_i = 0 \quad \text{при} \quad C = 0$$

Отсюда следует:

$$\sum_{i=1}^n \bar{y}_{x_i} n_i - \sum_{i=1}^n f^*(x_i) n_i = 0; \quad N\bar{y} - \sum_{i=1}^n f^*(x_i) n_i = 0; \quad \frac{\sum_{i=1}^n f^*(x_i) n_i}{N} = \bar{y}.$$

То есть и второе равенство (2.10) доказано.

Равенства (2.10) можно использовать для контроля правильности подсчета и реальных экспериментальных средних \bar{y}_{x_i} , и сглаживающих средних $\bar{y}_{x_i}^*$.

Рассмотрим, в частности, приложение метода наименьших квадратов к случаю, когда теоретические соображения или конфигурация корреляционного поля позволяют в качестве сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$ использовать уравнение прямой (2.5). Сформируем для этого случая сумму Q :

$$Q = Q(k; b) = \sum_{i=1}^n (\bar{y}_{x_i} - kx_i - b)^2 n_i \rightarrow \min$$

Вспоминая, что необходимым условием минимума (или максимума) функции нескольких переменных является равенство нулю всех ее частных производных первого порядка, получим следующую систему уравнений (так называемую *нормальную систему*) для нахождения параметров k и b функции (2.5):

$$\begin{cases} Q'_k = -2 \sum_{i=1}^n (\bar{y}_{x_i} - kx_i - b)x_i n_i = 0 \\ Q'_b = -2 \sum_{i=1}^n (\bar{y}_{x_i} - kx_i - b)n_i = 0 \end{cases} \quad (2.11)$$

Сократив на (-2), разбив суммы (2.11) на отдельные суммы и разделив затем обе части каждого уравнения на N , получим:

$$\begin{cases} k \cdot \frac{\sum_{i=1}^n x_i^2 n_i}{N} + b \cdot \frac{\sum_{i=1}^n x_i n_i}{N} = \frac{\sum_{i=1}^n \bar{y}_{x_i} x_i n_i}{N} \\ k \cdot \frac{\sum_{i=1}^n x_i n_i}{N} + b \cdot \frac{\sum_{i=1}^n n_i}{N} = \frac{\sum_{i=1}^n \bar{y}_{x_i} n_i}{N} \end{cases} \quad (2.12)$$

Учтем, что

$$\begin{aligned} \frac{\sum_{i=1}^n x_i^2 n_i}{N} &= \overline{x^2}; \quad \frac{\sum_{i=1}^n x_i n_i}{N} = \bar{x}; \quad \frac{\sum_{i=1}^n n_i}{N} = \frac{N}{N} = 1; \\ \frac{\sum_{i=1}^n \bar{y}_{x_i} \cdot x_i n_i}{N} &= \left| \text{см. (2.3)} \right| = \frac{\sum_{i=1}^n \sum_{j=1}^m x_i y_j n_{ij}}{N} = \overline{xy}; \\ \frac{\sum_{i=1}^n \bar{y}_{x_i} \cdot n_i}{N} &= \frac{\sum_{i=1}^n \sum_{j=1}^m y_j \cdot n_{ij}}{N} = \frac{\sum_{j=1}^m y_j \sum_{i=1}^n n_{ij}}{N} = \frac{\sum_{j=1}^m y_j m_j}{N} = \bar{y} \end{aligned} \quad (2.13)$$

где черточкой сверху обозначены средние значения соответствующих величин. Тогда система (2.12) примет вид:

$$\begin{cases} k\overline{x^2} + b\bar{x} = \overline{xy} \\ k\bar{x} + b = \bar{y} \end{cases} \quad (2.14)$$

Решая ее, находим k и b :

$$k = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2}; \quad b = \bar{y} - k\bar{x} \quad (2.15)$$

Подставляя найденные значения k и b в уравнение (2.5), получим искомое сглаживающее линейное уравнение регрессии:

$$\bar{y}_x^* - \bar{y} = k(x - \bar{x}) \quad (2.16)$$

Коэффициент k имеет, как уже отмечалось выше, простой смысл: он показывает, на сколько единиц величины y изменится предсказываемое среднее значение \bar{y}_x^* этой величины, если величина x изменится на свою единицу измерения.

Коэффициент k , таким образом, зависит от единиц измерения этих величин. Чтобы исключить эту зависимость, есть смысл перейти к универсальным единицам измерения величин x и y . В качестве таких единиц в математической статистике используются средние квадратические отклонения величин от их средних значений. То есть используются квадратные корни из дисперсий величин. Для наших переменных x и y это будут

$$\begin{aligned} \sigma_x &= \sqrt{D_x}, \text{ где дисперсия } D_x = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{N}; \\ \sigma_y &= \sqrt{D_y}, \text{ где дисперсия } D_y = \frac{\sum_{j=1}^m (y_j - \bar{y})^2 \cdot m_j}{N} \end{aligned} \quad (2.17)$$

Средние квадратические отклонения σ_x и σ_y имеют, очевидно, смысл средних отклонений величин x и y от их средних значений \bar{x} и \bar{y} соответственно без учета знака этих отклонений. Кстати, формулы (2.17) легко приводятся к виду

$$D_x = \sigma_x^2 = \overline{x^2} - (\bar{x})^2; \quad D_y = \sigma_y^2 = \overline{y^2} - (\bar{y})^2 \quad (2.18)$$

(проделайте необходимые для этого преобразования самостоятельно).

Представим теперь уравнение (2.16) в равносильном виде:

$$\frac{\bar{y}_x^* - \bar{y}}{\sigma_y} = \rho \cdot \frac{x - \bar{x}}{\sigma_x}, \quad (2.19)$$

где

$$\rho = k \cdot \frac{\sigma_x}{\sigma_y} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y} \quad (2.20)$$

Безразмерный коэффициент ρ называется *коэффициентом линейной корреляции* между величинами x и y . Он, согласно (2.19), показывает, на сколько величин σ_y изменится в среднем величина y , когда x изменится на одно σ_x .

Этот коэффициент играет ключевую роль при оценке тесноты линейной корреляционной связи между величинами x и y . Для выяснения его роли вы-

ведем уравнение (2.16) другим способом.

Пусть $\bar{y}_x^* = kx + b$ - искомое наилучшее усредненное линейное приближение зависимости y от x . Представим величину y в виде

$$y = kx + b + z, \quad (2.21)$$

где z - разность между величиной y и её наилучшим усредненным линейным приближением $kx + b$. То есть введем отклонение z величины y от её наилучшего усредненного линейного приближения:

$$y_j = kx_i + b + z_{ij} \quad (i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m) \quad (2.22)$$

Согласно (2.22), значения z_{ij} величины z - это отклонения по вертикали точек корреляционного поля от наилучшей прямой регрессии $\bar{y}_x^* = kx + b$. Для точек, лежащих выше этой прямой, $z_{ij} > 0$, а для точек, лежащих ниже, $z_{ij} < 0$.

Естественно считать ту прямую регрессии наилучшей, для которой величины z_{ij} были бы минимальными. А они будут минимальными, если z_{ij} будут удовлетворять следующим двум условиям:

$$1) \bar{z} = \frac{\sum_{i=1}^n \sum_{j=1}^m z_{ij} n_{ij}}{N} = 0; \quad 2) D_z = \sigma_z^2 = \frac{\sum_{i=1}^n \sum_{j=1}^m (z_{ij} - \bar{z})^2 n_{ij}}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m z_{ij}^2 n_{ij}}{N} \rightarrow \min \quad (2.23)$$

Реализуем сначала условие 1 из (2.23):

$$\bar{z} = \frac{\sum_{i=1}^n \sum_{j=1}^m z_{ij}}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - kx_i - b) n_{ij}}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m y_j n_{ij}}{N} - k \frac{\sum_{i=1}^n \sum_{j=1}^m x_i n_{ij}}{N} - b \frac{\sum_{i=1}^n \sum_{j=1}^m n_{ij}}{N} = 0 \quad (2.24)$$

Но

$$\begin{aligned} \frac{\sum_{i=1}^n \sum_{j=1}^m y_j n_{ij}}{N} &= \frac{\sum_{j=1}^m y_j \sum_{i=1}^n n_{ij}}{N} = \frac{\sum_{j=1}^m y_j m_j}{N} = \bar{y}; \\ \frac{\sum_{i=1}^n \sum_{j=1}^m x_i n_{ij}}{N} &= \frac{\sum_{i=1}^n x_i \sum_{j=1}^m n_{ij}}{N} = \frac{\sum_{i=1}^n x_i n_i}{N} = \bar{x}; \quad \frac{\sum_{i=1}^n \sum_{j=1}^m n_{ij}}{N} = \frac{N}{N} = 1 \end{aligned} \quad (2.25)$$

С учетом (2.25) равенство (2.24) примет вид:

$$\bar{y} - k\bar{x} - b = 0, \quad \text{откуда} \quad b = \bar{y} - k\bar{x} \quad (2.26)$$

Подставляя это значение b в уравнение $\bar{y}_x^* = kx + b$ искомой сглаживающей прямой линии регрессии, получим следующий его вид:

$$\bar{y}_x^* - \bar{y} = k(x - \bar{x}) \quad (2.27)$$

А подставляя это же значение b в (2.21), получим

$$y - \bar{y} = k(x - \bar{x}) + z, \quad (2.28)$$

откуда

$$z = (y - \bar{y}) - k(x - \bar{x}); \quad z_{ij} = (y_j - \bar{y}) - k(x_i - \bar{x}) \quad (2.29)$$

Осталось найти коэффициент k уравнения (2.27). Для этого с помощью выражений (2.29) реализуем условие 2 из (2.23):

$$\begin{aligned} D_z = \sigma_z^2 &= \frac{\sum_{i=1}^n \sum_{j=1}^m z_{ij}^2 n_{ij}}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m [(y_j - \bar{y}) - k(x_i - \bar{x})]^2 n_{ij}}{N} = \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})^2 n_{ij}}{N} - 2k \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})(x_i - \bar{x}) n_{ij}}{N} + k^2 \frac{\sum_{i=1}^n \sum_{j=1}^m (x_i - \bar{x})^2 n_{ij}}{N} \rightarrow \min \quad (2.30) \end{aligned}$$

Но

$$\begin{aligned} \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})^2 n_{ij}}{N} &= \frac{\sum_{j=1}^m (y_j - \bar{y})^2 \sum_{i=1}^n n_{ij}}{N} = \frac{\sum_{j=1}^m (y_j - \bar{y})^2 m_j}{N} = D_y = \sigma_y^2; \\ \frac{\sum_{i=1}^n \sum_{j=1}^m (x_i - \bar{x})^2 n_{ij}}{N} &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^m n_{ij}}{N} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 n_i}{N} = D_x = \sigma_x^2; \\ \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})(x_i - \bar{x}) n_{ij}}{N} &= \frac{\sum_{i=1}^n \sum_{j=1}^m x_i y_j n_{ij}}{N} - \bar{x} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^m y_j n_{ij}}{N} - \bar{y} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^m x_i n_{ij}}{N} + \bar{x} \cdot \bar{y} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^m n_{ij}}{N} = \\ &= \overline{xy} - \bar{x} \cdot \bar{y} - \bar{y} \cdot \bar{x} + \bar{x} \cdot \bar{y} = \overline{xy} - \bar{x} \cdot \bar{y} = |\text{учтем (2.20)}| = \rho \sigma_x \sigma_y \quad (2.31) \end{aligned}$$

Значит, выражение (2.30) примет вид:

$$D_z = \sigma_z^2 = \sigma_y^2 - 2k\rho\sigma_x\sigma_y + k^2\sigma_x^2 \rightarrow \min \quad (2.32)$$

Выделим из выражения (2.32) полный квадрат:

$$D_z = \sigma_z^2 = \sigma_y^2(1 - \rho^2) + (\rho\sigma_y - k\sigma_x)^2 \rightarrow \min \quad (2.33)$$

Первое из полученных слагаемых неотрицательно и не зависит от параметра k . Таким образом, дисперсия D_z будет минимальной при том значении k , которое обеспечит обращение в нуль второго слагаемого. То есть при

$$k = \rho \frac{\sigma_y}{\sigma_x} \quad (2.34)$$

При этом дисперсия D_z (её минимальное значение) примет вид:

$$D_z = \sigma_z^2 = \sigma_y^2(1 - \rho^2) \quad (2.35)$$

Итак, вывод: наилучшее приближение величины y линейной функцией будет иметь вид (2.27) при k , определяемым формулой (2.34). То есть в конечном итоге оно будет иметь вид (2.19). Эти выводы совпадают с полученными ранее другим путем.

По уравнению (2.27) (или (2.19)) для каждого значения x можно найти линейризованное усредненное значение \bar{y}_x^* величины y . Но теперь мы имеем ещё и формулу (2.35) для оценки тесноты этой сглаживающей линейризованной зависимости y от x . Теснота эта определяется степенью разброса точек корреляционного поля вокруг сглаживающей прямой. А этот разброс, в свою очередь, определяется значениями z_{ij} отклонений z точек корреляционного поля от указанной прямой. То есть их дисперсией (2.35).

Так как $D_z \geq 0$, то, согласно (2.35), всегда $-1 \leq \rho \leq 1$. Если $\rho = 1$ или $\rho = -1$, то дисперсия $D_z = 0$. А это, с учетом равенства $\bar{z} = 0$ означает, что все значения z_{ij} величины z равны нулю. То есть никаких отклонений точек корреляционного поля от прямой регрессии нет, все точки поля находятся на этой прямой. Значит, в этом случае между переменными x и y имеется точная (функциональная) линейная зависимость вида

$$y - \bar{y} = k(x - \bar{x}) \quad (2.36)$$

Но если коэффициент ρ линейной корреляции отличен от ± 1 , то $D_z \neq 0$, и она будет тем больше, чем ближе будет ρ к нулю. А это значит, что появляется разброс точек корреляционного поля вокруг прямой регрессии. И он будет тем больше, чем ближе ρ к нулю. При $\rho = 0$ дисперсия D_z принимает свое максимально возможное значение, а значит и разброс точек корреляционного поля вокруг прямой линии регрессии становится максимальным. А сами уравнения (2.27) и (2.19) этой прямой принимают вид

$$\bar{y}_x^* = \bar{y} \quad (2.37)$$

Усредненное линейное приближение величины y перестаёт, таким образом, зависеть от x . То есть при $\rho=0$ линейная зависимость y от x отсутствует. Это значит, что или между величинами x и y вообще нет никакой связи (они независимы друг от друга), или они связаны, но какой-то нелинейной связью (параболической, гиперболической, и т.д.).

Рассмотрим этот вопрос подробнее. То есть поставим вопрос об оценке *любой*, а не только линейной, корреляционной связи между величинами x и y . Эту оценку осуществляют с помощью так называемого *корреляционного отношения*.

Чтобы ввести это понятие, рассмотрим функцию $\bar{y}_x = f(x)$, которая при $x = x_i (i = 1, 2, \dots, n)$ принимает экспериментальные средние значения $\bar{y}_{x_i} = f(x_i)$ величины y . Среднее значение экспериментальных средних \bar{y}_{x_i} , как показано в (2.10), равно общей средней \bar{y} величины y . А дисперсия $D_{\bar{y}_x}$ величины \bar{y}_x составляет лишь часть общей дисперсии D_y величины y :

$$D_y = D_{\bar{y}_x} + \overline{(y - \bar{y}_x)^2}. \quad (2.38)$$

Величины, входящие в (2.38), подсчитываются по формулам:

$$D_y = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})^2 n_{ij}}{N} = \frac{\sum_{j=1}^m (y_j - \bar{y})^2 m_j}{N}; \quad (2.39)$$

$$D_{\bar{y}_x} = \frac{\sum_{i=1}^n (\bar{y}_{x_i} - \bar{y})^2 n_i}{N}; \quad \overline{(y - \bar{y}_x)^2} = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij}}{N}.$$

Докажем равенство (2.38):

$$D_y = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y})^2 n_{ij}}{N} = \frac{\sum_{i=1}^n \sum_{j=1}^m [(y_j - \bar{y}_{x_i}) + (\bar{y}_{x_i} - \bar{y})]^2 n_{ij}}{N} =$$

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij}}{N} + 2 \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})(\bar{y}_{x_i} - \bar{y}) n_{ij}}{N} + \frac{\sum_{i=1}^n \sum_{j=1}^m (\bar{y}_{x_i} - \bar{y})^2 n_{ij}}{N}. \quad (2.40)$$

Первое из трех слагаемых (2.40) – это, согласно (2.39), $\overline{(y - \bar{y}_x)^2}$. Второе из них равно нулю. Действительно:

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})(\bar{y}_{x_i} - \bar{y})n_{ij}}{N} = \frac{\sum_{i=1}^n (\bar{y}_{x_i} - \bar{y}) \sum_{j=1}^m (y_j - \bar{y}_{x_i})n_{ij}}{N} = 0, \text{ ибо}$$

$$\sum_{j=1}^m (y_j - \bar{y}_{x_i})n_{ij} = \sum_{j=1}^m y_j n_{ij} - \bar{y}_{x_i} \sum_{j=1}^m n_{ij} = |\text{см. (2.2) и (2.3)}| = \bar{y}_{x_i} n_i - \bar{y}_{x_i} n_i = 0.$$

А третье слагаемое в (2.40) – это дисперсия $D_{\bar{y}_x}$:

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (\bar{y}_{x_i} - \bar{y})^2 n_{ij}}{N} = \frac{\sum_{i=1}^n (\bar{y}_{x_i} - \bar{y})^2 \sum_{j=1}^m n_{ij}}{N} = \frac{\sum_{i=1}^n (\bar{y}_{x_i} - \bar{y})^2 n_i}{N} = D_{\bar{y}_x}.$$

Равенство (2.38) доказано.

Согласно формуле (2.38), разброс значений величины y вокруг её общего среднего значения \bar{y} раскладывается на разброс вокруг \bar{y} узловых точек \bar{y}_{x_i} ($i = 1, 2, \dots, n$) экспериментальной линии регрессии (первое слагаемое), и на разброс значений y_j величины y (разброс облака точек корреляционного поля) вокруг экспериментальной линии регрессии (второе слагаемое).

Равенство (2.38) является одним из важнейших в регрессионном анализе. И оно является основным в родственном регрессионному дисперсионном анализе.

Введем теперь отношение

$$\eta_{yx} = \sqrt{\frac{D_{\bar{y}_x}}{D_y}} = \sqrt{1 - \frac{(y - \bar{y}_x)^2}{D_y}}, \quad (2.41)$$

которое будем называть *корреляционным отношением* y к x . Очевидно, что всегда

$$0 \leq \eta_{yx} \leq 1 \quad (2.42)$$

Если $\eta_{yx} = 0$, то $D_{\bar{y}_x} = 0$. А это означает, что $\bar{y}_x = \text{Const}$, причем эта константа, естественно, равна \bar{y} . То есть $\bar{y}_{x_i} = \bar{y}$ для всех x_i ($i = 1, 2, \dots, n$). Экспериментальная ломаная регрессии (рис. 2.1) становится в этом случае горизонтальной прямой. Следовательно, в этом случае величина y корреляционно (в среднем) от x не зависит.

А если $\eta_{yx} = 1$, то

$$\frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij}}{(y - \bar{y}_x)^2} = \frac{\sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij}}{N} = 0$$

То есть отсутствует разброс значений y_j вокруг их средних \bar{y}_{x_i} : каждому значению x_i соответствуют n_i одинаковых значений y_j , из которых образуется такое же их среднее значение \bar{y}_{x_i} . Иначе говоря, каждому значению x_i величины x соответствует одно и то же значение величины y . Но эти значения y , вообще говоря, меняются с изменением x . Таким образом, величины x и y связаны некоторой жесткой (функциональной) зависимостью $y = f(x)$. Причем форма этой зависимости может оказаться любой.

Таким образом, корреляционное отношение η_{yx} является мерой тесноты *любой*, а не только линейной, корреляционной зависимости величины y от величины x . И чем η_{yx} больше (ближе к 1), тем теснее эта зависимость, то есть тем ближе она к некоторой функциональной зависимости. А чем η_{yx} меньше (ближе к нулю), тем эта зависимость слабее. При $\eta_{yx} = 0$ этой зависимости нет вообще.

Можно доказать (на этом не останавливаемся), что всегда $|\rho| \leq \eta_{yx}$. При это в случае равенства $|\rho| = \eta_{yx}$ имеет место *точная линейная* корреляционная зависимость y от x . Это значит, что точки $(x_i; \bar{y}_{x_i})$ ($i = 1, 2, \dots, n$) выстраиваются на корреляционном поле строго по прямой. А чем больше различаются между собой ρ и η_{yx} , тем в большей степени ломаная регрессии отличается от прямой, а корреляционная зависимость y от x - от линейной зависимости.

Если в формуле (2.41) заменить реальные средние \bar{y}_{x_i} на сглаживающие средние $\bar{y}_{x_i}^* = f^*(x_i)$ и возвести полученное выражение в квадрат, то получим так называемый *коэффициент детерминации* d_{yx} :

$$d_{yx} = \frac{D_{\bar{y}_x^*}}{D_y}, \quad \text{где} \quad D_{\bar{y}_x^*} = \frac{\sum_{i=1}^n (\bar{y}_{x_i}^* - \bar{y})^2 n_i}{N}; \quad D_y = \frac{\sum_{j=1}^m (y_j - \bar{y})^2 m_j}{N} \quad (2.43)$$

Он показывает долю, которую составляет дисперсия $D_{\bar{y}_x^*}$ сглаживающих средних $\bar{y}_{x_i}^*$ по отношению к общей дисперсии D_y значений y_j величины y . То есть он показывает долю общего изменения (вариации) величины y , объясняемую подобранным сглаживающим уравнением регрессии $\bar{y}_x^* = f^*(x)$. Его обычно выражают в процентах. В компьютерных программах (например, в Excel) коэффициент детерминации обозначается символом R^2 . А величина $R = \sqrt{d_{yx}}$ называется *индексом корреляции*.

Корреляционное отношение η_{yx} не зависит, очевидно, от формы выбранного сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$, ибо его величина определяется исключительно экспериментальными данными. А вот коэффициент детерминации $d_{yx} = R^2$ от этой формы зависит. Как можно доказать,

$$(d_{yx})_{\max} = \eta_{yx}^2 \quad (2.44)$$

И чем больше d_{yx} (чем ближе он к η_{yx}^2), тем лучше построенное сглаживающее уравнение регрессии объясняет вариацию (изменение) зависимой величины y . А следовательно, тем удачнее это уравнение построено. При $d_{yx} = \eta_{yx}^2$ сглаживающая линия регрессии $\bar{y}_x^* = f^*(x)$ точно пройдет через все экспериментальные точки $(x_i; \bar{y}_{x_i})$ ($i = 1, 2, \dots, n$) корреляционного поля, то есть через все узлы ломаной, изображенной на рис. 2.1. Это – идеальный вариант для сглаживающей линии. Правда, уравнение $\bar{y}_x^* = f^*(x)$ такой идеальной сглаживающей линии при большом числе узлов выборочной ломаной линии регрессии, как правило, слишком сложно. Поэтому на практике идут на существенное упрощение подбираемого сглаживающего уравнения регрессии, жертвуя при этом неизбежным снижением коэффициента детерминации.

Кстати, если сглаживающее уравнение регрессии строить в линейной (наиболее простой) форме (2.5), то будем иметь:

$$d_{yx} = \rho^2 \quad (2.45)$$

Действительно, для этого случая на основании (2.16)

$$D_{\bar{y}_x^*} = \frac{\sum_{i=1}^n [k(x_i - \bar{x})]^2 n_i}{N} = k^2 D_x = k^2 \sigma_x^2, \quad (2.46)$$

и поэтому

$$d_{yx} = \frac{D_{\bar{y}_x^*}}{D_y} = \frac{k^2 \sigma_x^2}{\sigma_y^2} = \left(k \frac{\sigma_x}{\sigma_y} \right)^2 = \rho^2 \quad (2.47)$$

2.3 Исследование уравнения парной регрессии на адекватность при наличии повторных опытов

Построив сглаживающее уравнение регрессии $\bar{y}_x^* = f^*(x)$ в нескольких различных формах (линейное, гиперболическое, параболическое и т.д.) и выбрав из них лучшее (по коэффициенту детерминации), мы тем не менее еще не мо-

жем быть уверены в пригодности такого уравнения для приближения им истинного (искомого) уравнения регрессии $\bar{y}_x = f(x)$. Дело в том, что построенная сглаживающая линия регрессии может на некоторых своих участках выходить за пределы корреляционного поля, особенно если полоса точек этого поля узкая (корреляционная зависимость y от x близка к функциональной). Тогда на этих участках сглаживающая линия не будет соответствовать корреляционному полю (будем неадекватна ему), а значит, будет неадекватно ему и уравнение $\bar{y}_x^* = f^*(x)$ этой сглаживающей линии. Такое уравнение не может быть использовано для всех x , входящих в корреляционную таблицу, а значит, его применение чревато грубыми ошибками, если им мы будем приближать искомого уравнения регрессии $\bar{y}_x = f(x)$. В этом случае построенное уравнение регрессии $\bar{y}_x^* = f^*(x)$ считается *неадекватным опытными данными* и применяться не должно.

Таким образом, после построения сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$ его еще нужно проверить на адекватность опытными данными. Адекватность этого уравнения будет тем выше, чем лучше соответствующая ему сглаживающая линия регрессии будет вписываться в полосу точек корреляционного поля. То есть чем меньше будет разброс этих точек вокруг указанной линии.

Оценим величину этого разброса. Для этого подсчитаем сумму квадратов отклонений ординат y_j всех точек корреляционного поля от сглаживающей линии регрессии $\bar{y}_x^* = f^*(x)$:

$$Q_0 = \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i}^*)^2 n_{ij} \quad (2.48)$$

Проведем следующее преобразование этой суммы:

$$\begin{aligned} Q_0 &= \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i}^*)^2 n_{ij} = \sum_{i=1}^n \sum_{j=1}^m [(y_j - \bar{y}_{x_i}) + (\bar{y}_{x_i} - \bar{y}_{x_i}^*)]^2 n_{ij} = \\ &= \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij} + 2 \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})(\bar{y}_{x_i} - \bar{y}_{x_i}^*) n_{ij} + \sum_{i=1}^n \sum_{j=1}^m (\bar{y}_{x_i} - \bar{y}_{x_i}^*)^2 n_{ij} = \\ &= \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij} + 2 \sum_{i=1}^n (\bar{y}_{x_i} - \bar{y}_{x_i}^*) \sum_{j=1}^m (y_j - \bar{y}_{x_i}) n_{ij} + \sum_{i=1}^n (\bar{y}_{x_i} - \bar{y}_{x_i}^*)^2 \sum_{j=1}^m n_{ij} \end{aligned}$$

Учитывая, согласно (5.2) и (5.3), что

$$\sum_{j=1}^m n_{ij} = n_i; \quad \sum_{j=1}^m (y_j - \bar{y}_{x_i}) n_{ij} = \sum_{j=1}^m y_j n_{ij} - \bar{y}_{x_i} \sum_{j=1}^m n_{ij} = \bar{y}_{x_i} n_i - \bar{y}_{x_i} n_i = 0,$$

получим окончательно:

$$Q_0 = Q_{повт} + Q_{адекват} \quad (2.49)$$

Здесь

$$Q_{повт} = \sum_{i=1}^n \sum_{j=1}^m (y_j - \bar{y}_{x_i})^2 n_{ij}; \quad Q_{адекват} = \sum_{i=1}^n (\bar{y}_{x_i}^* - \bar{y}_{x_i})^2 n_i \quad (2.50)$$

Сумма $Q_{повт}$ характеризует разброс экспериментальных значений y_j вокруг их средних \bar{y}_{x_i} при проведении повторных опытов для всех x_i , поэтому она так и обозначена: $Q_{повт}$. Она определяет степень влияния на величину y различных неучтенных факторов (помех), не связанных с величиной x . Кстати, сумма $Q_{повт}$ не зависит, очевидно, от вида сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$, так что уменьшить или увеличить ее нельзя – она определяется исключительно экспериментальными данными. А вот вторая сумма $Q_{адекват}$ зависит от вида уравнения $\bar{y}_x^* = f^*(x)$. Она характеризует меру отклонений сглаживающих средних $\bar{y}_{x_i}^*$ от реальных (экспериментальных) средних \bar{y}_{x_i} . И чем эта сумма меньше, тем более адекватным будет, очевидно, сглаживающее уравнение регрессии $\bar{y}_x^* = f^*(x)$. Поэтому эта сумма так и обозначена: $Q_{адекват}$. Кстати, сумма $Q_{адекват}$ – это как раз та сумма Q (см. (2.9)), на минимизации которой основано построение сглаживающего уравнения регрессии.

Естественно, что если $Q_{адекват} = 0$, то сглаживающее уравнение регрессии полностью адекватно экспериментальным данным (корреляционной таблице (2.1)). А если $Q_{адекват} \neq 0$, что обычно и бывает на самом деле, то сравнивая $Q_{адекват}$ с $Q_{повт}$ выясняют, достаточно ли мала сумма $Q_{адекват}$, чтобы можно было согласиться с утверждением об адекватности построенного сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$ экспериментальным данным.

Такое сравнение $Q_{адекват}$ с $Q_{повт}$ и соответствующие выводы из такого сравнения делаются с помощью науки, называемой *математическая статистика*.

В соответствии с этой наукой, сначала выдвигается основная (*нулевая*) гипотеза H_0 : построенное сглаживающее уравнение регрессии $\bar{y}_x^* = f^*(x)$ адекватно экспериментальным данным. Одновременно выдвигается противоречащая ей *альтернативная* гипотеза H_1 , что оно им не адекватно. Принимая одну из этих гипотез, мы тем самым отвергаем другую.

Так как судить о справедливости той или иной из этих двух гипотез мы будем на основе своих экспериментальных данных, а они содержат в себе и погрешности измерительных приборов, и влияние различного рода случайных факторов (помех), то какой бы вывод в итоге мы ни сделали, мы можем ошибиться.

Причем ошибка может быть двух родов. Если мы отвергнем нулевую гипотезу H_0 , а она на самом деле правильная, то мы совершим *ошибку первого рода*. Вероятность такой ошибки принято обозначать буквой α . А если мы

примем нулевую гипотезу H_0 , а она на самом деле неправильная, то мы совершим ошибку второго рода, вероятность которой обозначается буквой β . При этом сразу отметим, что невозможно полностью избежать этих ошибок в силу ограниченности и случайности экспериментальных данных.

В таблице 2.3 представлены все возможные ситуации с принятием и непринятием гипотезы H_0 (в скобках указаны вероятности соответствующих ситуаций)

Таблица 2.3

Заключение о гипотезе H_0	На самом деле гипотеза H_0	
	Верна	Неверна
Отвергается	Ошибка 1-го рода (α)	Правильное решение ($1-\beta$)
Принимается	Правильное решение ($1-\alpha$)	Ошибка 2-го рода (β)

Естественно пожелать, чтобы обе возможные ошибки 1-го и 2-го родов имели как можно меньшие вероятности. Но добиться этого трудно, ибо очевидно, что уменьшение α автоматически ведет к увеличению β , и наоборот.

Проиллюстрируем это утверждение на таком житейском примере. Пусть гипотеза H_0 состоит в том, что переход пешеходом перекрестка на красный свет окончится для него плохо. Тогда отвергнуть эту гипотезу – это значит переходить перекресток, не обращая внимания на красный свет светофора. Принять эту гипотезу – это значит стоять на месте (ждать зеленый сигнал).

1) Если гипотеза H_0 верна, то переходить перекресток – это отвергнуть правильную гипотезу, то есть совершить ошибку 1-го рода (трагическую).

2) Если гипотеза H_0 неверна, то оставаться на месте – это принять неправильную гипотезу, т.е. совершить ошибку 2-го рода (потерять время на ожидание зеленого сигнала светофора)

Очевидно, что тяжесть последствий каждой из этих двух ошибок совершенно разная. И еще очевидно, что уменьшая вероятность α ошибки 1-го рода, пешеход автоматически будет увеличивать вероятность β ошибки 2-го рода. Действительно, уменьшая ошибку 1-го рода, пешеход не будет переходить на красный свет даже в тех случаях, когда машины от перекрестка далеко и переход ему практически ничем бы не грозил.

Можно дать понятиям ошибки 1-го и 2-го родов и юридическую трактовку. Пусть гипотеза H_0 состоит в том, что подсудимый невиновен (это исходная юридическая гипотеза, вытекающая из принципа презумпции невиновности). Тогда ошибка 1-го рода будет состоять в том, что подсудимому будет вынесен обвинительный приговор, несмотря на то, что он невиновен. А ошибка 2-го рода будет состоять в том, что подсудимому будет вынесен оправдательный приговор, несмотря на то, что он виновен. Эти ошибки тоже разные по тяжести. Стараясь уменьшить одну из них (например, ошибку первого рода), судья чаще должен выносить оправдательные приговоры, что автоматически ведет к увеличению ошибки 2-го рода.

При практической проверке гипотез, выдвигая нулевую гипотезу H_0 и альтернативную ей гипотезу H_1 , сразу задают допустимую вероятность α ошибки 1-го рода – вероятность отвергнуть нулевую гипотезу H_0 , если она верна. Конкретное числовое значение для этой ошибки выбирают, исходя из реальной тяжести последствий такой ошибки. Кстати, принятое значение α называют еще *уровнем значимости*. Таким образом, *уровень значимости α* – это вероятность отвергнуть нулевую гипотезу H_0 при условии, что она верна. При решении инженерных, сельскохозяйственных, экономических задач по статистической проверке гипотез обычно в качестве уровня значимости α принимают значения $\alpha = 0,1$; или $\alpha = 0,05$; или $\alpha = 0,01$. Часто, особенно в литературе прикладного характера, уровень значимости выражается еще в процентах. При этом только что названные уровни значимости называют 10% - ым, 5% - ым, 1% - ым уровнями значимости. Чем меньше мы выберем значение α , тем меньше будет риск отвергнуть нулевую гипотезу H_0 , если она правильная. Но зато тем больше будет риск принять ее, если она неправильная (то есть если правильной будет альтернативная гипотеза H_1). Брать $\alpha=0$ смысла нет, так как при $\alpha=0$ риск отклонить проверяемую гипотезу H_0 , если она верна, должен быть сведен к нулю. А это значит, что при $\alpha=0$ гипотезу H_0 , как бы мало вероятной она ни была, заведомо нужно принимать. То есть при $\alpha=0$ теряется сам смысл статистической проверки гипотез.

Приняв гипотезу H_0 , тем самым соглашаются с тем, что те результаты исследования, которые этой гипотезе противоречат, являются несущественными (незначимыми) и являются лишь следствием помех в эксперименте или недостаточностью экспериментальных данных. А отвергнув гипотезу H_0 , тем самым признают, что противоречащие гипотезе H_0 результаты исследования трудно объяснить лишь случайными помехами, они слишком существенны (значимы) для того, чтобы можно было согласиться со справедливостью гипотезы H_0 . Наше решение относительно значимости или незначимости экспериментальных данных, противоречащих гипотезе H_0 , зависит от принятого уровня значимости α . Отсюда и его название – *уровень значимости*.

Отметим еще одно важное и очевидное обстоятельство: принятие или непринятие гипотезы H_0 зависит не только от величины уровня значимости, но и от того, какова гипотеза H_1 – гипотеза, альтернативная гипотезе H_0 . Действительно, гипотеза H_1 противостоит гипотезе H_0 – они конкурируют друг с другом за принятие. И устоит ли гипотеза H_0 против гипотезы H_1 – это, естественно, зависит и от того, что эта гипотеза H_1 собой представляет.

В самом деле, если пешеход выбирает, переходить ли ему на красный свет или дождаться на этом же месте через пару минут зеленого – это одна ситуация. А если он выбирает между переходом на красный свет и необходимостью перехода в другом месте, далеко от данного – это ситуация совсем другая. Во втором случае пешеходы будут рисковать, переходя улицу, чаще. И тем самым будут увеличивать вероятность собственной ошибки первого рода (ведущей под колеса автомобиля).

Заметим еще, что *отклонение нулевой гипотезы H_0* – акт гораздо более обоснованный, чем *принятие этой гипотезы*. Действительно, отклоняя гипотезу

зу H_0 , мы знаем вероятность того, что ошибаемся - ведь это вероятность ошибки первого рода. А она равна принятому уровню значимости α , то есть она известна и мала. А принимая гипотезу H_0 , мы не знаем вероятности возможной собственной ошибки (ошибки второго рода) - эта вероятность не контролируется, и она может быть не мала, особенно если слишком мал задаваемый уровень значимости α . Отклонив гипотезу H_0 , мы при малых значениях α практически можем быть уверены, что эта гипотеза действительно не верна. А приняв ее, мы отнюдь не можем утверждать, что подтвердили ее правильность. Мы можем лишь позволить себе осторожно сказать, что экспериментальные данные не противоречат этой гипотезе. Но ведь точно так же при малых α они могут не противоречить и многим другим, отличным от H_0 , гипотезам. А при $\alpha = 0$ любые противоречия гипотезе H_0 будут признаны несущественными (незначимыми) и гипотеза H_0 , даже самая невероятная, будет принята!

Для борьбы с принятием неверных гипотез H_0 нужно или увеличивать α , уменьшая тем самым ошибку β принятия неверной нулевой гипотезы, или, что гораздо продуктивнее, не увеличивая α , увеличивать объем экспериментальных данных, что будет повышать и качество, и надежность всех выводов, следующих из обработки этих данных.

В качестве критерия, в соответствии с которым решают, принять или не принять выдвинутую нулевую гипотезу H_0 , используют специально подобранную случайную величину K , распределение которой, при условии справедливости гипотезы H_0 , известно (величина K может иметь нормальное распределение, или распределение Стьюдента, или распределение Фишера-Снедекора, и т.д.). С учетом выдвинутой гипотезы H_0 , альтернативной гипотезы H_1 и принятого уровня значимости α множество всех возможных этой величины разбивают на два подмножества:

a) Множество тех значений k величины K , при которых нулевая гипотеза H_0 принимается (это множество называется *областью принятия гипотезы*);

b) Множество тех значений k величины K , при которых нулевая гипотеза отвергается (а, следовательно, принимается альтернативная гипотеза H_1). Это подмножество называется *критической областью*.

Точки $k_{кр}$ (их может быть одна или несколько), разделяющие область принятия гипотезы H_0 и критическую область, называются *критическими*.

Используя экспериментальные данные, вычисляют экспериментальное значение $k_{эксп}$ величины K . И если оно попадет в область принятия гипотезы H_0 , гипотезу H_0 принимают. А если в критическую область - то отвергают.

Заметим, что $k_{эксп}$ в силу случайности экспериментальных данных может попасть в критическую область и, следовательно, гипотеза H_0 будет отвергнута несмотря на то, что на самом деле она верна. Тогда будет допущена ошибка 1-го рода - отвергнута правильная гипотеза. Значит, вероятность такой ошибки, то есть *вероятность попадания $k_{эксп}$ в критическую область, должна быть равна принятому уровню значимости α* . Из этого условия, а также из соображений, касающихся возможной симметрии критической области, и находятся критические точки $k_{кр}$.

Наоборот, $k_{\text{эксп}}$ может попасть в область принятия гипотезы H_0 , когда она неверна. Тогда будет допущена ошибка 2-го рода – принята неправильная гипотеза. Вероятность β такой ошибки обычно неизвестна (она, в отличие от ошибки α , не задается и не контролируется). Не обращая внимание на эту ошибку второго рода, при попадании $k_{\text{эксп}}$ в область принятия гипотезы H_0 эту гипотезу принимают.

Для критических точек различных критериев K составлены готовые таблицы. Они содержатся во многих книгах по математической статистике. Есть такого рода таблицы и в Приложении к данному пособию.

Но можно получать указанные критические точки и с помощью компьютера. В частности, если в качестве критерия K избрана случайная величина, имеющая нормальное распределение, или распределение ХИ-квадрат, или распределение Стьюдента, или распределение Фишера-Снедекора (изучавшие математическую статистику должны быть знакомы с этими распределениями), то можно воспользоваться программами НОРМОБР, ХИ2ОБР, СТЬЮДРАСПОБР, ФРАСПОБР, содержащимися в Excel.

В свете всего сказанного выше вернемся к вопросу о проверке значимости (пригодности для всех значений x) сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$. Эту проверку делают по критерию Фишера-Снедекора, если заведомо известно (или подтверждено экспериментально), что зависимая величина y при любом значении независимой величины x распределена по нормальному закону (что обычно имеет место) и имеет не зависящую от x постоянную дисперсию.

Схема проверки уравнения регрессии на его адекватность экспериментальным данным

1. Выдвигаем основную (нулевую) гипотеза H_0 : построенное сглаживающее уравнение регрессии $\bar{y}_x^* = f^*(x)$ адекватно экспериментальным данным. Одновременно выдвигаем альтернативную гипотеза H_1 , что оно им не адекватно.

2. По формулам (2.50) подсчитываем суммы $Q_{\text{новт}}$ и $Q_{\text{адекв}}$ и сравниваем их между собой. Отметим только, что сравнивать их непосредственно друг с другом некорректно, так как в их образовании участвует разное число слагаемых. Перед сравнением этих сумм каждую из них рассчитывают на одно её независимое слагаемое, то есть на одну её *степень свободы*.

По определению, число степеней свободы некоторой вычисляемой по опытным данным величины – это количество данных, участвующих в её формировании, за вычетом количества независимых друг от друга числовых характеристик, вычисляемых по тем же опытным данным и тоже участвующих в формировании указанной величины.

У суммы $Q_{\text{новт}}$, в соответствии с её выражением (2.50), число степеней свободы $k_{\text{новт}} = N - n$. Действительно, в её формировании участвуют данные

всех N опытов. При этом она содержит в себе n величин \bar{y}_{x_i} , вычисляемых по тем же опытными данным. Поэтому $k_{новт} = N - n$. А у суммы $Q_{адекват}$, в соответствии с её выражением (2.50), число степеней свободы $k_{адекват} = n - q$, где q – число коэффициентов уравнения регрессии. Действительно, она содержит в себе n независимых экспериментальных средних \bar{y}_{x_i} и столько же подсчитанных по уравнению регрессии сглаженных средних $\bar{y}_{x_i}^*$. Но коэффициенты этого уравнения подсчитаны по экспериментальным данным. Поэтому в сумме $Q_{адекват}$ не n независимых слагаемых, а $n - q$. И, значит, $k_{адекват} = n - q$.

Делением сумм $Q_{новт}$ и $Q_{адекват}$ на их числа степеней рассчитываем каждую из них на одну её степень свободы. В итоге получаем *дисперсию повторности* $s_{новт}^2$ и *дисперсию адекватности* $s_{адекват}^2$:

$$s_{новт}^2 = \frac{Q_{новт}}{N - n}; \quad s_{адекват}^2 = \frac{Q_{адекват}}{n - q} \quad (2.51)$$

И вот их уже теперь сравнивать друг с другом корректно. Сравнивают их по критерию F Фишера-Снедекора. Сначала находим экспериментальное значение этого критерия:

$$f_{эксн} = \frac{s_{адекват}^2}{s_{новт}^2} \quad (2.52)$$

Потом по таблице критических точек распределения Фишера – Снедекора (см. таблицу 5 Приложения) находим критическое значение этого критерия

$$f_{кр} = f_{кр}(\alpha; n - q; N - n), \quad (2.53)$$

и сравниваем его с экспериментальным $f_{эксн}$. И если окажется, что $f_{эксн} > f_{кр}$, то при данном уровне значимости α гипотезу H_0 об адекватности сглаживающего уравнения регрессии отвергаем. То есть признаем построенное уравнение регрессии $\bar{y}_x^* = f^*(x)$ непригодным для приближения истинного уравнения регрессии $\bar{y}_x = f(x)$. А если окажется, что $f_{эксн} < f_{кр}$, то у нас нет оснований отвергать гипотезу H_0 . И только такое (адекватное) уравнение регрессии можно использовать в дальнейшем.

Кстати, если для имеющихся экспериментальных данных построено несколько различных уравнений регрессии, и все они адекватны экспериментальным данным, то лучшим среди них считается то, которое, не являясь заметно сложнее прочих, обеспечивает наибольший коэффициент детерминации $d_{yx} = R^2$.

Наряду с проверкой адекватности сглаживающего уравнения регрессии в целом имеется возможность проверить и значимость каждого его коэффициента в отдельности. Это значит – имеется возможность установить, достаточ-

но ли велико по модулю подсчитанное значение интересующего нас коэффициента для статистически обоснованного вывода о том, что он отличен от нуля. И если окажется, что коэффициент не значим, то его можно положить равным нулю. Это приведет к упрощению сглаживающего уравнения регрессии без существенного ущерба для его качества.

Такое исследование производится автоматически, если сглаживающее уравнение регрессии строится с помощью компьютерной программы корреляционно-регрессионного анализа (например, с помощью инструмента анализа «Регрессия» из Пакета анализа Excel, или с помощью программы ЛИНЕЙН из того же Excel). В сглаживающем уравнении регрессии, выдаваемом компьютером, указывается погрешность определения каждого его коэффициента, что позволяет установить, значим он или незначим. Программа выдает также коэффициент детерминации построенного уравнения регрессии, а заодно и позволяет установить, адекватно или неадекватно все уравнение в целом. Обо всем этом подробнее будет сказано ниже.

Впрочем, что касается компьютерного исследования уравнений регрессии на их адекватность опытным данным, то тут необходимо отметить одно важное обстоятельство. Изложенная выше схема исследования, основанная на использовании формул (2.50)-(2.53), предполагает, что таблица экспериментальных данных (2.1) содержит результаты y_j повторных опытов хотя бы для некоторых значений x_i величины x . Если же при каждом экспериментальном значении x_i величины x величина y определялась лишь однократно (повторных опытов не проводилось), то у нас будет отсутствовать (окажется нулевой) сумма $Q_{повт}$, характеризующая разброс экспериментальных точек корреляционного поля вокруг ломаной регрессии, ибо все экспериментальные точки будут лежать на этой ломаной, составляя её узлы. А так как сумма $Q_{повт}$ дает оценку влияния неучтенных факторов (помех) на исследуемую величину y , то этой оценки мы иметь не будем. Повторив опыты, мы можем в принципе получить кардинально другие результаты. А, значит, вся изложенная выше схема исследования уравнений регрессии на их адекватность экспериментальным данным окажется невозможной.

В таком случае эту адекватность проверяют другим путем. И именно этот другой путь используют компьютерные программы Excel. В том числе и при наличии повторных опытов.

2.4 Парная корреляция при отсутствии повторных опытов

Компьютерная программа парного линейного регрессионного анализа зависимости y от x предполагает, что таблица экспериментальных данных значений x и y задана в следующем простом виде:

Таблица 2.4

x_i	x_1	x_2	x_n
y_i	y_1	y_2	y_n

То есть в виде двух строк или двух столбцов. В случае отсутствия повторных опытов все значения x_i величины x будут разными. В противном случае в таблице 2.4 будут встречаться пары $(x_i; y_i)$ с одинаковыми значениями x_i . Число n - это число проведенных опытов. При отсутствии повторных опытов таблица (2.4) является частным случаем таблицы (2.1), когда

$$n_{ij} = \begin{cases} 1, & \text{если } i = j \\ 0, & \text{если } i \neq j \end{cases}; \quad n_i = 1; \quad m_j = 1; \quad N = n = m. \quad (2.54)$$

Экспериментальная линия регрессии для данных таблицы 2.4 будет представлять собой ломаную, соединяющую n точек с координатами $(x_i; y_i)$. А искомая линия регрессии $y^* = f^*(x)$, будучи достаточно простой (например, прямой), должна наилучшим образом сглаживать эту ломаную. Параметры сглаживающей линии находятся методом наименьших квадратов.

В частности, если указанную экспериментальную ломаную сглаживать прямой $y^* = kx + b$, то параметры k и b этой прямой находятся по выведенным выше формулам (2.15):

$$k = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}; \quad b = \bar{y} - k\bar{x} \quad (2.55)$$

Только здесь

$$\bar{x} = \frac{\sum x_i}{n}; \quad \bar{y} = \frac{\sum y_i}{n}; \quad \overline{x^2} = \frac{\sum x_i^2}{n}; \quad \overline{y^2} = \frac{\sum y_i^2}{n}; \quad \overline{xy} = \frac{\sum x_i y_i}{n}; \quad (2.56)$$

При этом введено сокращение: $\sum = \sum_{i=1}^n$. Реализуя эти формулы, программа устанавливает уравнение регрессии $y^* = kx + b$.

Вычислив по этому уравнению сглаженные (линеаризованные) значения $y_i^* = kx_i + b$ ($i = 1, 2, \dots, n$) величины y и дополнив ими таблицу (2.4), получим таблицу (2.5):

Таблица 2.5

x_i	x_1	x_2	x_n
y_i	y_1	y_2	y_n
y_i^*	y_1^*	y_2^*	y_n^*

Она будет наглядно демонстрировать расхождение между экспериментальными значениями y_i величины y и соответствующими им значениями y_i^* , кото-

рые получаются из уравнения регрессии. Кстати, значения y_i и y_i^* имеют, как это было доказано ранее (см (2.10)), общую среднюю \bar{y} :

$$\frac{\sum y_i}{n} = \frac{\sum y_i^*}{n} = \bar{y} \quad (2.57)$$

Приведем выражения и для других важных величин, о которых говорилось выше:

$$\begin{aligned} D_x = \sigma_x^2 &= \frac{\sum (x_i - \bar{x})^2}{n} = \overline{x^2} - (\bar{x})^2; & D_y = \sigma_y^2 &= \frac{\sum (y_i - \bar{y})^2}{n} = \overline{y^2} - (\bar{y})^2; \\ D_{y^*} = \sigma_{y^*}^2 &= \frac{\sum (y_i^* - \bar{y})^2}{n} = \overline{y^{*2}} - (\bar{y})^2; & \rho &= k \frac{\sigma_x}{\sigma_y} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y}; \\ d_{yx} = R^2 &= \frac{D_{y^*}}{D_y} = \frac{\sum (y_i^* - \bar{y})^2}{\sum (y_i - \bar{y})^2}. \end{aligned} \quad (2.58)$$

Причем в случае линейной корреляции $y^* = kx + b$ коэффициент детерминации $d_{yx} = R^2 = \rho^2$, где ρ - коэффициент линейной корреляции.

Рассмотрим теперь вопрос, как проверить на адекватность экспериментальным данным построенное уравнение регрессии в случае отсутствия повторных опытов. Соображения здесь такие: если разброс экспериментальных точек y_i вокруг линии регрессии мал, то есть все эти точки лежат близко к этой линии, и этих точек достаточно много, то маловероятно, что такая близость имеет место чисто случайно. Гораздо вероятнее, что между величинами x и y действительно имеется связь, и эта связь близка к той связи, которую устанавливает построенное уравнение регрессии. То есть у нас есть весомые доводы в пользу того, что такое уравнение можно считать адекватным экспериментальным данным. И эти доводы будут тем весомее, чем меньше будет суммарная величина отклонений экспериментальных значений y_i от соответствующих им значений y_i^* , вычисленных по уравнению регрессии.

В качестве такой суммарной величины естественно взять сумму $\sum (y_i - y_i^*)^2$. А сравнивают её с суммой $\sum (y_i^* - \bar{y})^2$, характеризующей разброс выровненных значений y_i^* вокруг общей средней \bar{y} . Это сравнение основано на главном уравнении линейного регрессионного анализа, которое как для повторных, так и для бесповторных опытов выглядит так:

$$\sum (y_i - \bar{y})^2 = \sum (y_i - y_i^*)^2 + \sum (y_i^* - \bar{y})^2 \quad (2.59)$$

Докажем равенство (2.59) для линейного уравнения регрессии $y^* = kx + b$, имея в виду экспериментальную таблицу (2.5) и формулы (2.55):

$$\begin{aligned}\sum (y_i - \bar{y})^2 &= \sum [(y_i - y_i^*) + (y_i^* - \bar{y})]^2 = \\ &= \sum (y_i - y_i^*)^2 + 2\sum (y_i - y_i^*)(y_i^* - \bar{y}) + \sum (y_i^* - \bar{y})^2\end{aligned}$$

Осталось доказать, что второе из этих трех слагаемых равно нулю. И это действительно так:

$$\begin{aligned}\sum (y_i - y_i^*)(y_i^* - \bar{y}) &= \sum (y_i - y_i^*)y_i^* - \bar{y}\sum (y_i - y_i^*) = \sum (y_i - y_i^*)y_i^*, \\ \text{ибо, согласно (2.57), } \sum (y_i - y_i^*) &= \sum y_i - \sum y_i^* = n\bar{y} - n\bar{y} = 0.\end{aligned}$$

Подставим в оставшуюся сумму $\sum (y_i - y_i^*)y_i^*$ значения $y_i^* = kx_i + b$ и учтем, согласно (2.55), что $b = \bar{y} - k\bar{x}$:

$$\begin{aligned}\sum (y_i - y_i^*)y_i^* &= \sum [(y_i - \bar{y}) - k(x_i - \bar{x})] \cdot [\bar{y} + k(x_i - \bar{x})] = \\ &= \bar{y}\sum [(y_i - \bar{y}) - k(x_i - \bar{x})] + k\sum (y_i - \bar{y})(x_i - \bar{x}) - k^2\sum (x_i - \bar{x})^2 = \\ &= \bar{y}(\sum (y_i - \bar{y}) - k\sum (x_i - \bar{x})) + k(\sum x_i y_i - \bar{y}\sum x_i - \bar{x}\sum y_i + \sum \bar{x} \cdot \bar{y}) - k^2\sum (x_i - \bar{x})^2 = \\ &= \bar{y}(0 - k \cdot 0) + k(n \cdot \overline{xy} - \bar{y} \cdot n\bar{x} - \bar{x} \cdot n\bar{y} + n \cdot \bar{x} \cdot \bar{y}) - k^2 n \sigma_x^2 = \\ &= kn(\overline{xy} - \bar{x} \cdot \bar{y} - k\sigma_x^2) = \text{учтем выражение (2.55) для } k = kn \cdot 0 = 0.\end{aligned}\quad (2.60)$$

Равенство (2.59) доказано. Согласно этому равенству, разброс экспериментальных значений y_i величины y вокруг их общей средней \bar{y} раскладывается на сумму квадратов их отклонений от соответствующих им выровненных значений y_i^* величины y (первое слагаемое), и на разброс выровненных y_i^* вокруг их общей средней \bar{y} (второе слагаемое).

Отметим, что равенство (2.59) является верным не только в том доказанном нами случае, когда линией регрессии является прямая $y^* = kx + b$, но и когда ею является любая кривая $y^* = f^*(x)$, лишь бы параметры её уравнения входили в него линейно. В частности, это равенство справедливо для гиперболического уравнения регрессии

$y^* = \frac{k}{x} + b$, параболического

$y^* = a_0 + a_1x + a_2x^2$, и т.д. Оно справедливо и в многомерном регрессионном анализе, когда исследуется линейная корреляционная зависимость одной величины от нескольких других. Но об этом отдельно будет сказано позже.

2.5 Исследование уравнения парной регрессии на адекватность при отсутствии повторных опытов

А теперь перейдем к использованию равенства (2.59) для исследования построенного линейного по параметрам уравнения регрессии на его адекватность экспериментальным данным. Введем следующие обозначения для входящих в него сумм с указанием их степеней свободы:

$$\begin{aligned} S_{общ} &= \sum (y_i - \bar{y})^2 && \text{— общая сумма; } && k_{общ} = n - 1; \\ S_{рег} &= \sum (y_i^* - \bar{y})^2 && \text{— регрессионная сумма; } && k_{рег} = q - 1; \\ S_{ост} &= \sum (y_i - y_i^*)^2 && \text{— остаточная сумма; } && k_{ост} = n - q. \end{aligned} \quad (2.61)$$

Здесь, напомним, $\sum = \sum_{i=1}^n$, где n - число пар $(x_i; y_i)$ экспериментальных данных, q - число коэффициентов уравнения регрессии. В частности, для уравнений $y^* = kx + b$ и $y^* = \frac{k}{x} + b$ число $q = 2$. И $q = 3$ - для уравнения $y^* = a_0 + a_1x + a_2x^2$.

С учетом обозначений (2.61) уравнение (2.59) примет вид:

$$S_{общ} = S_{рег} + S_{ост} \quad (2.62)$$

Смысл уравнения (2.62) мы уже отмечали выше: общий разброс $S_{общ}$ экспериментальных значений y_i величины y вокруг их общей средней \bar{y} раскладывается на сумму квадратов $S_{ост}$ их отклонений от соответствующих им выровненных значений y_i^* величины y , и на разброс $S_{рег}$ выровненных y_i^* вокруг их общей средней \bar{y} .

Отметим, что величину $S_{рег}$ иногда еще обозначают как $S_{объясн}$, ибо через неё уравнение регрессии объясняет изменение выходной величины y изменением тех входных величин, которые в это уравнение входят. А величина $S_{ост}$ характеризует необъясненную уравнением регрессии часть общей суммы. Её наличие связано с действием случайных факторов (помех).

Чем больше окажется величина $S_{рег}$ по сравнению с $S_{ост}$, тем качественнее уравнение регрессии объясняет изменение исследуемой величины y . То есть тем адекватнее экспериментальным данным оно будет.

Но эти суммы не сравниваются непосредственно друг с другом. Предварительно каждую из них рассчитывают на одну их степень свободы, для чего эти суммы делят на число их степеней свободы. Тем самым вычисляются *регрессионная дисперсия* $D_{рег} = s_{рег}^2$ и *остаточная дисперсия* $D_{ост} = s_{ост}^2$:

$$D_{reg} = s_{reg}^2 = \frac{S_{reg}}{k_{reg}} = \frac{S_{reg}}{q-1}; \quad D_{ocm} = s_{ocm}^2 = \frac{S_{ocm}}{k_{ocm}} = \frac{S_{ocm}}{n-q} \quad (2.63)$$

А вот их уже сравнивают между собой. А именно, находят их отношение

$$f_{эксн} = \frac{s_{reg}^2}{s_{ocm}^2} \quad (2.64)$$

Это отношение должно быть достаточно большим (больше некоторого $f_{крит}$), чтобы при заданном уровне значимости α мы могли сделать вывод, что построенное уравнение регрессии значимо, то есть адекватно экспериментальным данным.

Доказано, что если при каждом значении x_i входной величины x выходная величина y имеет нормальное распределение, и если её дисперсия не зависит от x (в реальных задачах это, как правило, выполняется), то величина $f_{эксн}$ имеет распределение Фишера-Снедекора с $\{k_1 = k_{reg} = q-1\}$ и $\{k_2 = k_{ocm} = n-q\}$ степенями свободы. По таблице критических точек этого распределения определяют её критическое значение $f_{крит} = f_{крит}(\alpha; k_1; k_2)$ и сравнивают его с $f_{эксн}$. Если окажется, что $f_{эксн} \succ f_{крит}$, то уравнение регрессии признают значимым (то есть адекватным экспериментальным данным). А если окажется, что $f_{эксн} \prec f_{крит}$ - незначимым (неадекватным).

Отметим, что именно так это делается в Excel в программе *Регрессия*, входящей в *Пакет анализа*. А также в программе ЛИНЕЙН, которая тоже осуществляет линейный регрессионный анализ (на её основе работает и программа *Регрессия*).

Отметим, что величину $f_{эксн}$ можно находить не только по формуле (2.64), но и выразить её через коэффициент детерминации $d_{yx} = R^2$, который подсчитывает и выдает программа. Действительно, согласно (2.58) и (2.61),

$$d_{yx} = R^2 = \frac{S_{reg}}{S_{общ}} \quad (2.64^*)$$

Поэтому, учитывая (2.63),

$$f_{эксн} = \frac{s_{reg}^2}{s_{ocm}^2} = \frac{S_{reg}(n-q)}{(q-1)S_{ocm}} = \frac{S_{reg}(n-q)}{(q-1)(S_{общ} - S_{reg})} = \frac{R^2(n-q)}{(q-1)(1-R^2)} \quad (2.65)$$

В частности, для линейного уравнения регрессии $y^* = kx + b$ число его коэффициентов $q = 2$, и поэтому для него

$$f_{\text{эксн}} = \frac{R^2(n-2)}{(1-R^2)}, \text{ а } f_{\text{крит}} = f_{\text{крит}}(\alpha; 1; n-2). \quad (2.65^*)$$

Примечание. Коэффициент детерминации R^2 характеризует близость построенной регрессии к исходным данным, которые содержат «нежелательную» случайную составляющую. Очевидно, что, взяв в качестве уравнения регрессии полином $n-1$ порядка, содержащего n коэффициентов, мы получим выравнивающую линию регрессии, проходящую через все n узлов экспериментальной ломаной регрессии. То есть получим $y_i^* = y_i$ для всех $i = 1, 2, \dots, n$. А, значит, получим $S_{\text{ост}} = 0$ и «идеальное» значение $R^2 = 1$. Но такое уравнение будет содержать в себе не только независимую переменную x , но и целиком включит в себя случайную составляющую, а это снижает точность использования построенного уравнения регрессии для прогноза. Да и выражения (2.64), (2.65) и (2.65*) для $f_{\text{эксн}}$ потеряют смысл из-за деления в них на ноль. Поэтому при выборе уравнения регрессии надо учитывать не только величину R^2 , но и «сложность» регрессионного уравнения, определяемое количеством его коэффициентов. Такой учет реализован в так называемом *приведенном коэффициенте детерминации* R_*^2 :

$$R_*^2 = 1 - \frac{n-1}{n-q}(1-R^2) \quad (2.65^{**})$$

Очевидно, что увеличение числа q коэффициентов уравнения регрессии уменьшает значение R_*^2 . Если количество коэффициентов у сравниваемых уравнений регрессии одинаково (например, $q = 2$), то отбор наилучшей регрессии можно осуществлять по величине R^2 . Если же в уравнениях регрессии меняется число коэффициентов, то отбор целесообразно вести по величине R_*^2 .

2.6 Пример построения парной регрессии вручную

Пример 1. На некотором предприятии выборочным путем исследовалась зависимость себестоимости y единицы продукции (в условных единицах) от объема x произведенной за день продукции. Данные за 30 рабочих дней, попавших в выборку, приведены в таблице 2.:

Таблица 2.6

$x_i \backslash y_j$	5	10	15	20	25	$m_j = \sum_{i=1}^5 n_{ij}$
10	–	–	–	1	4	5
11	–	3	6	4	1	14
12	1	3	2	–	1	7
13	3	–	1	–	–	4
$n_i = \sum_{j=1}^4 n_{ij}$	4	6	9	5	6	$N = 30$

Требуется подобрать подходящую форму сглаживающего уравнения регрессии $\bar{y}_x^* = f^*(x)$, оценивающего корреляционную зависимость себестоимости единицы продукции от объема продукции, произведенной за день, и построить это уравнение. Кроме того, требуется оценить степень тесноты указанной корреляционной зависимости, а также качество и адекватность построенного сглаживающего уравнения регрессии.

Решение. Прделаем сначала всю вычислительную работу вручную, чтобы наглядно проиллюстрировать приведенные выше теоретические формулы.

Сначала по данным корреляционной таблицы (2.6) построим корреляционное поле (рис.2.2):

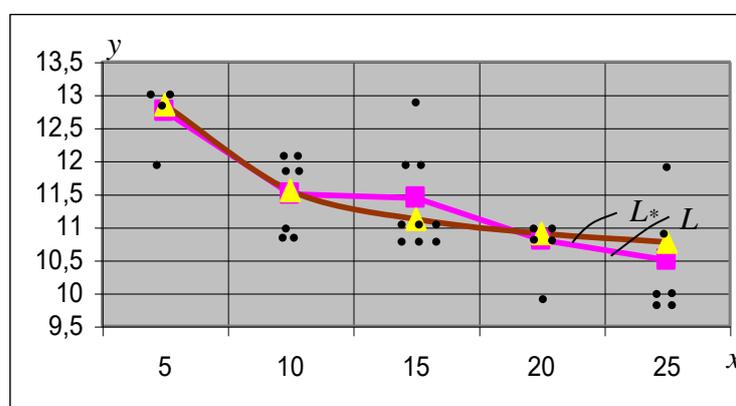


Рис.2.2

Используя формулу (2.3), вычислим для каждого x_i выборочную (экспериментальную) среднюю \bar{y}_{x_i} :

$$\bar{y}_{x_1=5} = \frac{12 \cdot 1 + 13 \cdot 3}{4} = 12,75; \quad \bar{y}_{x_2=10} = \frac{11 \cdot 3 + 12 \cdot 3}{6} = 11,50;$$

$$\bar{y}_{x_3=15} = 11,44; \quad \bar{y}_{x_4=20} = 10,80; \quad \bar{y}_{x_5=25} = 10,50;$$

По точкам $(x_i; \bar{y}_{x_i})$ строим на корреляционном поле экспериментальную линию регрессии – ломанную L (ее узлы на рис. 2.2 обозначены квадратиками).

Теперь встает очередная задача - решить, в какой форме $\bar{y}_x^* = f^*(x)$ искать сглаживающее уравнение этой экспериментальной линии регрессии.

Обратим внимание на то, что с увеличением x экспериментальные средние \bar{y}_x убывают, причем это убывание затухает. Так и должно быть (по смыслу рассматриваемых величин x и y). Это дает нам основание строить сглаживающее экспериментальное уравнение регрессии в гиперболической форме (2.6):

$$\bar{y}_x^* = \frac{k}{x} + b$$

Заметим, что такое уравнение – это фактически линейное уравнение $\bar{y}_x^* = kx + b$, если в последнем заменить x на $1/x$. Поэтому коэффициенты k и b этого уравнения можем найти с помощью формул (2.15) с заменой в них x на $1/x$:

$$k = \frac{\left(\overline{\frac{y}{x}}\right) - \left(\overline{\frac{1}{x}}\right) \cdot \bar{y}}{\left(\overline{\frac{1}{x^2}}\right) - \left[\left(\overline{\frac{1}{x}}\right)\right]^2}; \quad b = \bar{y} - k \cdot \left(\overline{\frac{1}{x}}\right) \quad (2.66)$$

Здесь

$$\begin{aligned} \left(\overline{\frac{1}{x}}\right) &= \frac{\sum_{i=1}^n \frac{1}{x_i} n_i}{N}; & \left(\overline{\frac{1}{x^2}}\right) &= \frac{\sum_{i=1}^n \frac{1}{x_i^2} n_i}{N}; \\ \bar{y} &= \frac{\sum_{j=1}^m y_j m_j}{N}; & \left(\overline{\frac{y}{x}}\right) &= \frac{\sum_{i=1}^n \sum_{j=1}^m \frac{y_j}{x_i} n_{ij}}{N} \end{aligned} \quad (2.67)$$

Реализуем формулы (2.67) и (2.66):

$$\left(\overline{\frac{1}{x}}\right) = \frac{\frac{1}{5} \cdot 4 + \frac{1}{10} \cdot 6 + \frac{1}{15} \cdot 9 + \frac{1}{20} \cdot 5 + \frac{1}{25} \cdot 6}{30} = 0,083$$

$$\left(\overline{\frac{1}{x^2}}\right) = \frac{\frac{1}{5^2} \cdot 4 + \frac{1}{10^2} \cdot 6 + \frac{1}{15^2} \cdot 9 + \frac{1}{20^2} \cdot 5 + \frac{1}{25^2} \cdot 6}{30} = 0,0094$$

$$\bar{y} = \frac{10 \cdot 5 + 11 \cdot 14 + 12 \cdot 7 + 13 \cdot 4}{30} = 11,33$$

$$\left(\frac{y}{x}\right) = \frac{\frac{12}{5} \cdot 1 + \frac{13}{5} \cdot 3 + \frac{11}{10} \cdot 3 + \dots + \frac{11}{25} \cdot 1 + \frac{12}{25} \cdot 1}{30} = 0,973$$

$$k = \frac{0,973 - 0,083 \cdot 11,33}{0,0094 - (0,083)^2} \approx 12,8; \quad b = 11,33 - 12,8 \cdot 0,083 \approx 10,27$$

Итак, сглаживающее экспериментальное уравнение регрессии в гиперболической форме таково:

$$\bar{y}_x^* = f^*(x) = \frac{12,8}{x} + 10,27 \quad (2.68)$$

Вычислим сглаживающие средние $\bar{y}_{x_i}^* = f^*(x_i)$ для всех x_i и сравним их с реальными (экспериментальными) средними \bar{y}_{x_i} :

x_i	5	10	15	20	25
n_i	4	6	9	5	6
\bar{y}_{x_i}	12,75	11,50	11,44	10,80	10,50
$\bar{y}_{x_i}^* = f^*(x_i)$	12,83	11,55	11,12	10,91	10,78

(2.69)

Впрочем, сначала убедимся, что и те, и другие средние подсчитаны правильно. Используя в качестве контроля формулы (2.10) убеждаемся, что обе суммы (2.10) дают один и тот же результат – общую среднюю $\bar{y} = 11,33$. То есть и реальные, и сглаживающие средние подсчитаны верно. И они весьма близки друг к другу. Это демонстрирует и рис. 2.2, где изображена гипербола (2.68) с указанием на ней точек $(x_i; \bar{y}_{x_i}^*)$, помеченных треугольниками.

А теперь перейдем к получению ответов на остальные вопросы – о степени тесноты корреляционной зависимости y от x и о качестве построенного уравнения (2.58).

При нелинейной корреляционной зависимости (а уравнение (2.68) нелинейно относительно x) степень тесноты корреляционной зависимости y от x оценивает корреляционное отношение η_{yx} . Подсчитывая его по формуле (2.39), получим: $\eta_{yx} \approx 0,75$. Величина η_{yx} оказалась весьма значительной (гораздо ближе к 1, чем нулю), что указывает на определенную и достаточно тесную корреляционную зависимость y от x .

Подсчитаем еще, используя формулу (2.41), коэффициент детерминации d_{yx} : $d_{yx} = R^2 \approx 0,502 \approx 50,2\%$. При этом, согласно (2.42),

$$(d_{yx})_{\max} = \eta_{yx}^2 = 0,5625 = 56,25 \% \quad (2.70)$$

Значение $d_{yx} \approx 50,2\%$ указывает, что построенное сглаживающее экспериментальное уравнение регрессии (2.68) объясняет 50,2% общего объема вариации (изменения) величины y и лишь немного не дотягивает до своего максимально возможного значения в 56,25%. И это имеет место при весьма простом виде уравнения (2.68). Чтобы окончательно убедиться в высоком качестве этого уравнения, следует проверить его на адекватность экспериментальным данным.

Так как корреляционная таблица 2.6 содержит результаты повторных опытов, то имеем возможность исследовать построенное уравнение регрессии (2.68) на адекватность, используя формулы (2.50) – (2.54):

$$\begin{aligned} Q_{\text{повт}} &= (12-12,75)^2 \cdot 1 + (13-12,75)^2 \cdot 3 + \dots + (11-10,50)^2 \cdot 1 + (12-10,50)^2 \cdot 1 \approx 10,77 \\ Q_{\text{адекв}} &= (12,85-12,75)^2 \cdot 4 + (11,55-11,50)^2 \cdot 6 + \dots + (10,77-10,50)^2 \cdot 6 \approx 1,464 \\ s_{\text{повт}}^2 &= \frac{10,77}{30-5} \approx 0,431; \quad s_{\text{адекв}}^2 = \frac{1,464}{5-2} \approx 0,488; \quad f_{\text{эксн}} = \frac{0,488}{0,431} = 1,13 \end{aligned} \quad (2.71)$$

Далее задаем уровень значимости α (например, $\alpha=0,05$) и по таблице критических точек распределения Фишера – Снедекора находим:

$$f_{кр} = f_{кр}(0.05; 3; 25) = 2,99 \quad (2.72)$$

И так как оказалось, что $f_{\text{эксн}} < f_{кр}$, то у нас нет оснований отвергать гипотезу H_0 об адекватности уравнения (2.68) экспериментальным данным. В пользу этого свидетельствует и рис. 2.2: гипербола L^* нигде не выходит за пределы корреляционного поля.

Все задания примера 1 выполнены.

Кстати, если бы мы искали сглаживающее уравнение регрессии в линейной форме (2.5), то, используя (2.13) и (2.15), получили бы:

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^5 x_i n_i}{N} = \frac{5 \cdot 4 + 10 \cdot 6 + 15 \cdot 9 + 20 \cdot 5 + 25 \cdot 6}{30} = 15,5 \\ \overline{x^2} &= \frac{\sum_{i=1}^5 x_i^2 n_i}{N} = \frac{5^2 \cdot 4 + 10^2 \cdot 6 + 15^2 \cdot 9 + 20^2 \cdot 5 + 25^2 \cdot 6}{30} = 282,5 \\ \bar{y} &= \frac{\sum_{j=1}^4 y_j m_j}{N} = \frac{10 \cdot 5 + 11 \cdot 14 + 12 \cdot 7 + 13 \cdot 4}{30} \approx 11,33 \\ \overline{y^2} &= \frac{\sum_{j=1}^4 y_j^2 m_j}{N} = \frac{10^2 \cdot 5 + 11^2 \cdot 14 + 12^2 \cdot 7 + 13^2 \cdot 4}{30} \approx 129,27 \end{aligned}$$

$$\sigma_x^2 = \overline{x^2} - (\bar{x})^2 = 42,25; \quad \sigma_x = 6,50; \quad \sigma_y^2 = \overline{y^2} - (\bar{y})^2 = 0,823; \quad \sigma_y = 0,907$$

$$\overline{xy} = \frac{\sum_{i=1}^5 \sum_{j=1}^4 x_i y_j n_{ij}}{N} = 171,5; \quad k = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x^2} = -0,0986; \quad b = \bar{y} - k\bar{x} = 12,86$$

И тогда вместо гиперболического (2.68) мы бы получили линейное уравнение

$$\bar{y}_x^* = f^*(x) = -0,0986x + 12,86 \quad (2.73)$$

Если подсчитать по этому уравнению сглаживающие средние $\bar{y}_{x_i}^*$ и сравнить их с экспериментальными средними \bar{y}_{x_i} , то получим следующую таблицу:

x_i	5	10	15	20	25
n_i	4	6	9	5	6
\bar{y}_{x_i}	12,75	11,50	11,44	10,80	10,50
$\bar{y}_{x_i}^* = f^*(x_i)$	12,37	11,87	11,38	10,89	10,40

(2.74)

Если для полученных сглаживающих средних $\bar{y}_{x_i}^*$ провести, на основе формулы (2.10), контроль, то он (проверьте это) сходится - опять получаем общую среднюю $\bar{y} = 11,33$.

Как видим, в таблице (2.74), как и в таблице (2.69), расхождение средних \bar{y}_{x_i} и $\bar{y}_{x_i}^*$ невелико. То есть линейное уравнение (2.73), как и гиперболическое уравнение (2.68), тоже представляется достаточно качественным.

Выясним все же, какое из них лучше. Для этого подсчитаем коэффициент детерминации d_{yx} и для линейного уравнения (2.73). Используя формулу (2.41) получим: $d_{yx} \approx 0,498 = 49,8\%$. Впрочем, в линейном случае его можно было бы найти и по формуле (2.43), если предварительно найти экспериментальный коэффициент линейной корреляции ρ :

$$\rho = \frac{\sigma_x}{\sigma_y} k = \frac{6,50}{0,907} (-0,0986) \approx -0,706; \quad d_{yx} = (\rho)^2 \approx 0,498 = 49,8\%$$

Итак, линейное сглаживающее уравнение регрессии (2.73) объясняет 49,8% всей вариации зависимой величины y . Гиперболическое же уравнение (2.68) объясняло чуть больше: 50,2% этой вариации. То есть по этому показателю гиперболическое уравнение чуть лучше линейного. И оно, по существу, так же просто, как и линейное.

Выше мы показали, что гиперболическое уравнение адекватно экспериментальным данным. Покажем, что и линейное уравнение им адекватно. Используя опять формулы (2.50) – (2.54), получаем:

$$Q_{новт} = 10,77 \text{ (см. (2.71))}$$

$$Q_{адекват} = (12,37 - 12,75)^2 \cdot 4 + (11,87 - 11,50)^2 \cdot 6 + \dots + (10,40 - 10,50)^2 \cdot 6 \approx 1,53 \quad (2.75)$$

$$s_{новт}^2 = \frac{10,77}{30 - 5} \approx 0,431; \quad s_{адекват}^2 = \frac{1,53}{5 - 2} \approx 0,511; \quad f_{эксп} = \frac{0,511}{0,431} \approx 1,18$$

При том же уровне значимости $\alpha = 0,05$, который был принят при проверке гипотезы об адекватности гиперболического уравнения регрессии, в соответствии с (2.55) получаем для линейного уравнения (2.73) то же самое критическое значение критерия Фишера – Снедекора, что было указано в (2.72): $f_{кр} = 2,99$. И так как опять $f_{эксп} < f_{кр}$, то у нас нет оснований отвергнуть гипотезу H_0 и об адекватности линейного сглаживания уравнения регрессии (2.73) экспериментальным данным.

В общем, оба сглаживающие уравнения регрессии – гиперболическое (2.68) и линейное (2.73) – адекватны экспериментальным данным и оба достаточно хороши. Из них несколько лучшим является гиперболическое уравнение (2.68).

2.7 Построение парной регрессии с помощью Excel

Все результаты (и даже сверх того), что мы получили выше вручную, с помощью Excel можно получить автоматически простым нажатием нескольких клавиш. Опишем, в частности, эту процедуру при использовании функции ЛИНЕЙН из комплекта статистических функций Excel. Эта функция на основе метода наименьших квадратов строит линейное уравнение регрессии $y^* = mx + b$ (используем обозначения программы), а также рассчитывает и выводит на экран (если её об этом попросить) дополнительную регрессионную статистику.

Реализация этой программы осуществляется с помощью следующих шагов:

Шаг 1. Вносим на лист Excel исходные экспериментальные значения $(x_i; y_i)$ ($i = 1, 2, \dots, n$) в виде двух столбцов (или двух строк)

Шаг 2. Выделяем на листе рядом с указанными выше столбцами таблицу из 5 строк и двух столбцов для вывода массива результатов работы программы ЛИНЕЙН.

Шаг 3. Открываем окно программы ЛИНЕЙН. В нем расположены 4 окошка.

- В первое из них «известные значения y » вносим столбец y_i .
- Во второе «известные значения x » вносим столбец x_i .
- Третье окошко «конст» предназначено для внесения в него логического значения ИСТИНА или ЛОЖЬ. Это значение указывает, требуется ли, чтобы константа b была равна 0.

Если аргумент **конст** имеет значение ИСТИНА или опущен, то параметр b уравнения регрессии $\bar{y}_x^* = mx + b$ вычисляется обычным путем по формуле $b = \bar{y} - m\bar{x}$. Если **конст** имеет значение ЛОЖЬ, то b изначально полагается

равным 0, и уравнение регрессии строится в виде $\bar{y}_x^* = mx$ с одним параметром m . Последний вариант используется, если заведомо ясно, что при $x = 0$ и $y = 0$. Это будет иметь место, например, если x - количество реализованной продукции, а y - доход от её продаж.

• Четвертое окошко «статистика» (как и третье) тоже предназначено для внесения в него логического значения ИСТИНА или ЛОЖЬ. Здесь это значение указывает, требуется ли рассчитывать и выводить на экран дополнительную статистику по регрессии. Если аргумент статистика имеет значение ИСТИНА, то функция ЛИНЕЙН выводит на экран (возвращает) регрессионную статистику в её полном объеме. А именно, выводит на месте, выделенном на шаге 2, массив из 5 строк и двух столбцов вида:

Таблица 2.7

m	b
se_m	se_b
R^2	se_y
F	df
SS_{reg}	SS_{resid}

Если аргумент **статистика** имеет значение ЛОЖЬ или опущен, то функция ЛИНЕЙН возвращает только коэффициент m и постоянную b . И тогда на втором шаге достаточно выделить под них массив лишь из двух горизонтально расположенных ячеек.

Шаг 4. Нажимаем комбинацию **Shift-Ctrl-Enter** и получаем массив результатов работы программы ЛИНЕЙН. Её работа на этом закончена.

Прокомментируем получаемые программой данные, приведенные в таблице 2.7.

- $\{m; b\}$ - это коэффициенты регрессии;
- $\{se_m; se_b\}$ - их среднеквадратические ошибки;
- $R^2 = \frac{SS_{reg}}{SS_{reg} + SS_{resid}}$ - коэффициент детерминации (см. (2.64*)); Здесь $SS_{reg} = S_{reg}$ - регрессионная сумма; $SS_{resid} = S_{ocm}$ - остаточная сумма; $SS_{reg} + SS_{resid} = S_{обц}$ - общая сумма квадратов отклонений экспериментальных y_i от их общей средней \bar{y} (см.(2.61)).
- se_y - оценка среднего квадратического отклонения экспериментальных y_i от $y_i^* = mx_i + b$, рассчитанных по уравнению регрессии. Эта оценка осуществляется по формуле $se_y = \sqrt{(se_y)^2}$, где

$$(se_y)^2 = \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n-2} = \frac{ss_{resid}}{n-2} \quad (2.76)$$

- дисперсия отклонений y_i от y_i^* . При её подсчете остаточная сумма $ss_{resid} = \sum_{i=1}^n (y_i - y_i^*)^2$ делится на число её степеней свободы $n-2$, ибо две степени свободы теряются при подсчете двух параметров m и b , входящих в выражение для $y_i^* = mx_i + b$.

- $$F = f_{эксн} = \frac{R^2(n-2)}{(1-R^2)} \quad (2.77)$$

(см (2.65)) – экспериментальное значение F - критерия Фишера-Снедекора, по которому выясняется, можно ли считать уравнение регрессии адекватным экспериментальным данным.

- $df = k_2 = n - q$ - второе из двух чисел степеней свободы F - критерия (первое из них $k_1 = q - 1$), где q - число коэффициентов уравнения регрессии; $q = 2$ для уравнения $\bar{y}_x^* = mx + b$. По заданному уровню значимости α и числам k_1 и k_2 находится табличное критическое значение $f_{крит}(\alpha; k_1; k_2)$ этого критерия. И если при сравнении $f_{эксн}$ с $f_{крит}$ окажется, что $f_{эксн} \succ f_{крит}$, то гипотеза об адекватности уравнения регрессии отвергается. А если $f_{эксн} \prec f_{крит}$, то принимается.

Отметим еще, что найденные программой среднеквадратические (стандартные) ошибки $\{se_m; se_b\}$ коэффициентов регрессии m и b позволяют ответить на вопрос, следует ли считать их значимыми (отличными от нуля), или они незначимы, а, значит, их можно положить равными нулю без существенного ущерба для уравнения регрессии. Если модули этих коэффициентов существенно (в два и более раз) превосходят их стандартные ошибки, то такие коэффициенты при обычном уровне значимости $\alpha = 0,05$ признаются существенными (значимыми). А если это не так – незначимыми.

Впрочем, такой простой вывод оправдан, если количество n экспериментальных пар значений $(x_i; y_i)$ достаточно велико (два-три десятка и более). При меньшем n значимость коэффициентов регрессии признается, когда

$$\frac{|m|}{se_m} \succ t_{крит}(\alpha; n - q); \quad \frac{|b|}{se_b} \succ t_{крит}(\alpha; n - q), \quad (2.78)$$

где $t_{крит}(\alpha; k = n - q)$ - критическое значение T - критерия Стьюдента (его можно взять в таблице 4 Приложения). В противном случае коэффициенты регрессии m и b признаются незначимыми.

Отметим ещё, что величины $\{se_m; se_b\}$ стандартных ошибок коэффициентов m и b линейного уравнения регрессии $y^* = mx + b$ рассчитываются по формулам:

$$se_m = \frac{se_y}{\sqrt{n} \cdot \sigma_x}; \quad se_b = \frac{se_y}{\sqrt{n}} \sqrt{1 + \frac{(\bar{x})^2}{\sigma_x^2}} \quad (2.79)$$

Выведем эти формулы. Сначала выведем выражение для se_m . Как следует из (2.55) и (2.56),

$$m = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2} = \frac{\frac{\sum x_i y_i}{n} - \bar{x} \cdot \frac{\sum y_i}{n}}{\sigma_x^2} = \frac{\sum (x_i - \bar{x}) y_i}{n \sigma_x^2} \quad (2.80)$$

Таким образом, коэффициент m линейно выражается через случайные значения y_i зависимой переменной y и неслучайные значения независимой переменной x . В регрессионном анализе изначально предполагается, что значения y_i распределены нормально с дисперсией, не зависящей от номера измерения i . Согласно уравнению регрессии,

$$y_i = mx_i + b + \varepsilon_i = y_i^* + \varepsilon_i \quad (i = 1, 2, \dots, n) \quad (2.81)$$

Здесь $\varepsilon_i = y_i - y_i^*$ - погрешности приближения значений экспериментальных значений y_i расчетными значениями $y_i^* = mx_i + b$, найденными методом наименьших квадратов. Статистическая оценка $s(\varepsilon_i)$ среднеквадратических значений погрешностей ε_i величин y_i и одновременно самих величин y_i дается формулой (2.76): $s(\varepsilon_i) = se_y$. С учетом этого дисперсия $D(m)$ коэффициента m равна:

$$D(m) = (se_m)^2 = \frac{\sum (x_i - \bar{x})^2 D(y_i)}{n^2 \sigma_x^4} = \frac{(se_y)^2 \cdot \sum (x_i - \bar{x})^2}{n^2 \sigma_x^4} = \frac{(se_y)^2 \cdot n \sigma_x^2}{n^2 \sigma_x^4} = \frac{(se_y)^2}{n \sigma_x^2},$$

откуда и следует выражение (2.79) для se_m .

Теперь подтвердим выражение (2.79) для se_b . Так как

$$b = \bar{y} - m\bar{x} = \frac{\sum y_i}{n} - \bar{x} \cdot \frac{\sum (x_i - \bar{x})y_i}{n\sigma_x^2} = \frac{1}{n} \sum \left(1 - \frac{\bar{x}(x_i - \bar{x})}{\sigma_x^2} \right) y_i,$$

то

$$\begin{aligned} D(b) &= (se_b)^2 = \frac{1}{n^2} \sum \left(1 - \frac{\bar{x}(x_i - \bar{x})}{\sigma_x^2} \right) (se_y)^2 = \\ &= \frac{(se_y)^2}{n^2} \cdot \sum \left(1 - \frac{2\bar{x}(x_i - \bar{x})}{\sigma_x^2} + \frac{(\bar{x})^2(x_i - \bar{x})^2}{\sigma_x^4} \right) = \\ &= \frac{(se_y)^2}{n^2} \left(\sum 1 - \frac{2\bar{x}}{\sigma_x^2} \sum (x_i - \bar{x}) + \frac{(\bar{x})^2}{\sigma_x^4} \sum (x_i - \bar{x})^2 \right) = \frac{(se_y)^2}{n^2} \left(n - \frac{2\bar{x}}{\sigma_x^2} \cdot 0 + \frac{(\bar{x})^2}{\sigma_x^4} \cdot n\sigma_x^2 \right) = \\ &= \frac{(se_y)^2}{n} \left(1 + \frac{(\bar{x})^2}{\sigma_x^2} \right), \end{aligned}$$

откуда и следует выражение (2.79) для se_b .

Добавим к этому ещё, что так как, по условию, величины y_i распределены нормально, а коэффициенты регрессии m и b линейно выражаются через y_i , то они тоже распределены нормально вокруг своих истинных значений. А их среднеквадратические отклонения от этих истинных значений (стандартные ошибки) $\{se_m; se_b\}$ нами только что вычислены.

А теперь покажем, что выдает программа ЛИНЕЙН, если ей предложить провести регрессионный анализ зависимости y от x в случае, когда таблица экспериментальных данных – это таблица 2.6.

Если оформить экспериментальные данные, представленные в этой таблице, в виде таблицы 2.4, содержащей $n = 30$ пар значений $(x_i; y_i)$ величин x и y , то, используя функцию ЛИНЕЙН и исходя из таблицы вида $(\frac{1}{x_i}; y_i)$, получим уравнение регрессии (2.68). А исходя из таблицы $(x_i; y_i)$ – уравнение (2.73). Точнее, получим следующих два массива, оформленных в виде таблицы 2.7 и содержащих всю регрессионную статистику для соответствующих уравнений регрессии:

Таблица 2.8

$y^* = \frac{m}{x} + b$		$y^* = mx + b$	
$m = 12,81541$	$b = 10,26966$	$m = -0,09862$	$b = 12,86193$
$se_m = 2,411129$	$se_b = 0,233809$	$se_m = 0,018646$	$se_b = 0,313401$
$R^2 = 0,502226$	$se_y = 0,662205$	$R^2 = 0,49976$	$se_y = 0,663843$
$F = 28,25039$	$df = 28$	$F = 27,97315$	$df = 28$
$ss_{reg} = 12,38823$	$ss_{resid} = 12,27843$	$ss_{reg} = 12,32742$	$ss_{resid} = 12,33925$

Согласно результатам, выданным программой, коэффициенты m и b в обоих уравнениях регрессии значимы, ибо их модули в несколько раз превосходят их среднеквадратические погрешности se_m и se_y соответственно. Адекватность обоих уравнений тоже подтверждается, ибо $F = f_{эксн} = 28,25$ для гиперболического уравнения (2.68), и $F = f_{эксн} = 27,97$ для линейного уравнения (2.73). И оба этих значения $f_{эксн}$ больше табличного критического значения

$$f_{крит} = f_{крит}(\alpha; q - 1; n - q) = f_{крит}(0.05; 1; 28) = 4,20.$$

2.8 Парные нелинейные регрессии и их построение в Excel

С помощью программ Excel можно строить не только линейные уравнения регрессии вида $y^* = kx + b$, но и различные нелинейные уравнения – как по объясняющей переменной x , так и по коэффициентам регрессии. Например, команда «Добавить линию тренда» позволяет построить уравнения регрессии в следующих формах:

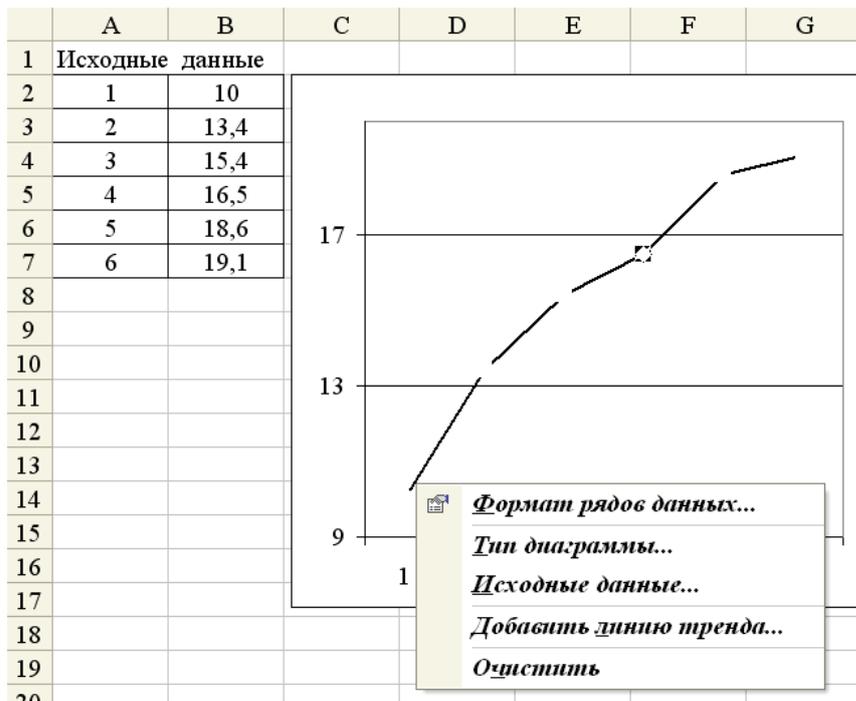
- линейное $y^* = b_0 + b_1x$;
- полиномиальное $y^* = b_0 + b_1x + b_2x^2 + \dots + b_kx^k \quad (k \leq 6)$;
- логарифмическое $y^* = b_0 + b_1 \ln x$;
- степенное $y^* = b_0x^{b_1}$;
- экспоненциальное $y^* = b_0e^{b_1x}$.

Заметим, что первое из этих уравнений линейно и по переменной x , и по своим коэффициентам. Второе и третье – только по коэффициентам. Четвертое и пятое нелинейно и по переменной x , и по своим коэффициентам b_0 и b_1 . Но какова бы ни была выбранная форма уравнения регрессии, его коэффициенты находятся на основе общего универсального метода – метода наименьших квадратов.

Построение любой из перечисленных регрессий с помощью электронных таблиц Excel осуществляется с помощью следующих шагов:

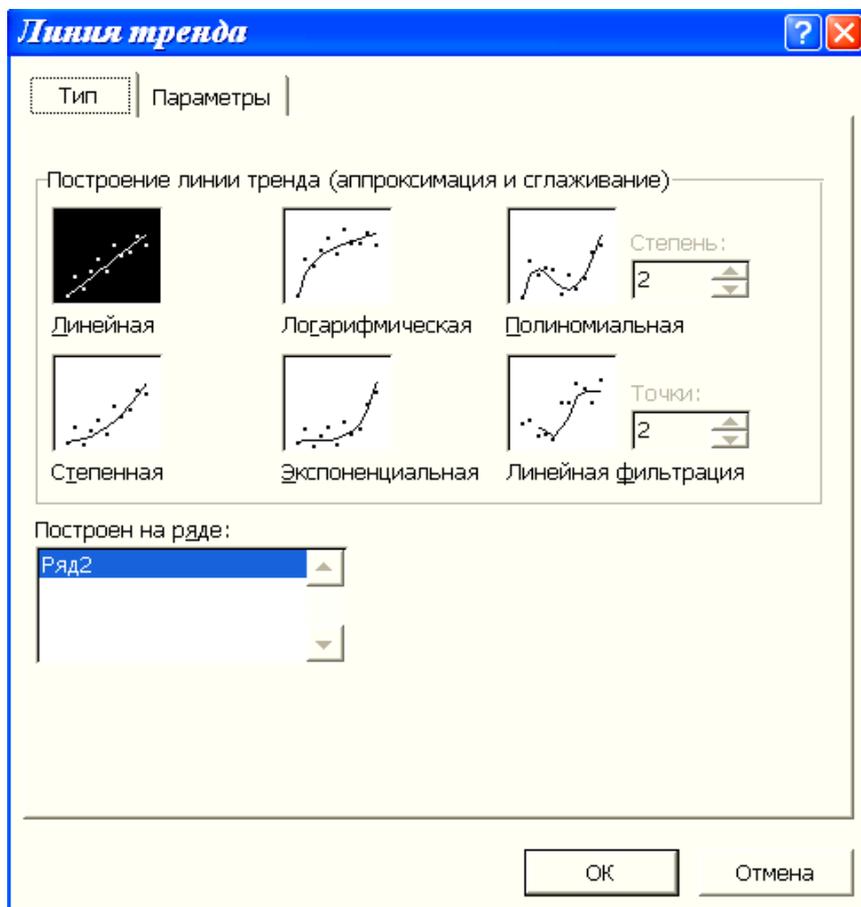
Шаг 1. На лист Excel вводим по столбцам исходные данные $(x_i; y_i) \quad (i = 1, 2, \dots, n)$.

Шаг 2. По этим данным строим в декартовой системе координат точечную диаграмму ломаной регрессии (как на рис. ниже).



Шаг 3. Устанавливаем курсор на построенном графике, делаем щелчок правой кнопкой мыши и в появившемся контекстном меню выполняем команду «Добавить линию тренда».

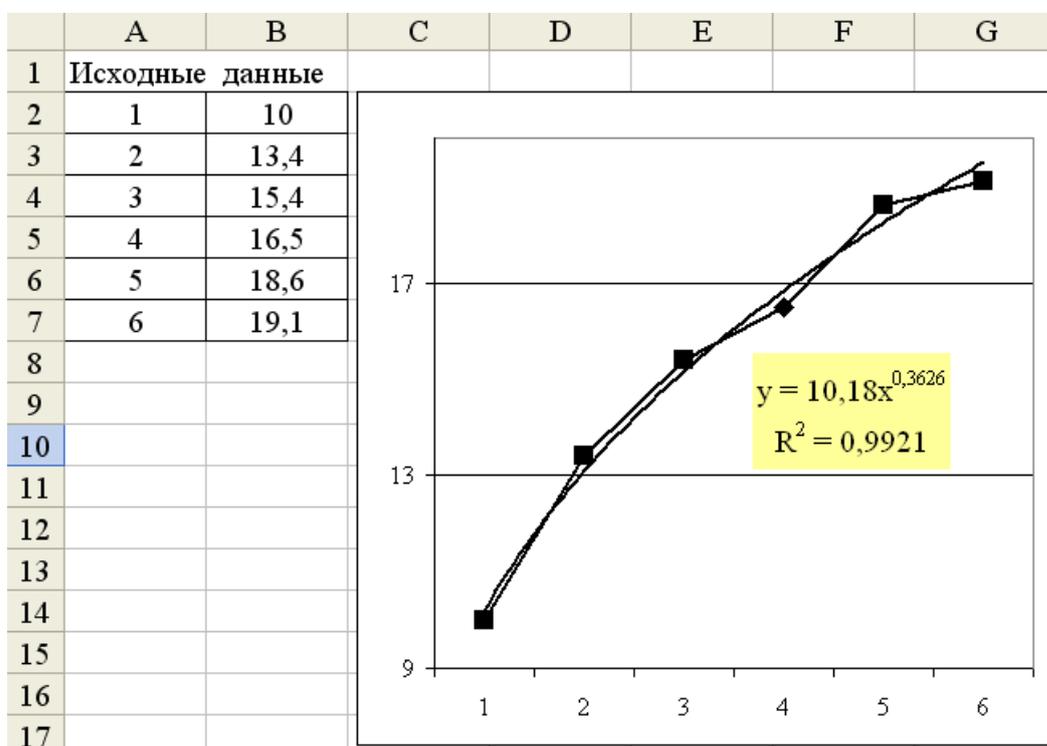
Шаг 4. В появившемся диалоговом окне активизируем закладку «Тип» и выбираем нужное уравнение регрессии.



Шаг 5. Активизируем закладку «Параметры» и включаем необходимые для нас опции:

- «Показать уравнение на диаграмме» - на диаграмме будет показано выбранное уравнение регрессии с вычисленными коэффициентами.
- «Поместить на диаграмму величину достоверности аппроксимации (R^2)» - на диаграмме будет показано значение коэффициента детерминации $d_{yx} = R^2$, которое можно использовать для проверки значимости построенной регрессии с помощью F-теста, сравнивая $f_{эксн}$ (формула (2.65)) с $f_{крит}(\alpha; k_1 = q - 1; k_2 = n - q)$.

Шаг 6. После задания всех перечисленных опций щелкаем по кнопке «ОК», и на диаграмме появиться формула построенного уравнения регрессии и значение коэффициента детерминации $d_{yx} = R^2$ (на рисунке ниже в качестве примера построено степенное уравнение регрессии).



2.9 Множественная корреляция и множественная регрессия

Перейдем теперь к более сложному случаю, когда экспериментальным путем изучается корреляционная зависимость (зависимость в среднем) одной величины y от нескольких других величин $\{x_1; x_2; \dots x_n\}$ (множественная корреляция). Множественную корреляционную зависимость характеризует *множественное уравнение регрессии* $y^* = f(x_1; x_2; \dots x_n)$, где y^* - усредненное (выровненное) значение величины y , соответствующее заданным значениям

переменных $\{x_1; x_2; \dots x_n\}$. Как правило, в качестве множественного уравнения регрессии используют линейное уравнение вида

$$y^* = m_n x_n + m_{n-1} x_{n-1} + \dots + m_1 x_1 + b \quad (2.82)$$

Построение такого уравнения по опытными данным с последующим его анализом представляет собой *множественный линейный регрессионный анализ*. Как и в случае парной регрессии, коэффициенты $\{m_n; m_{n-1}; \dots m_1; b\}$ множественной регрессии (2.82) находятся на основе метода наименьших квадратов.

В случае множественной регрессии этот метод, естественно, требует использования гораздо более сложных формул, чем в случае парной регрессии. Эти формулы, записанные на языке матричной алгебры, мы приводить не будем, так как на практике множественный линейный регрессионный анализ проводится с помощью компьютерных программ с уже введенными в них расчетными формулами. Например, с помощью той же экселевской программы (функции) ЛИНЕЙН, с помощью которой проводится и парный линейный регрессионный анализ. И схема осуществления множественного анализа практически повторяет схему осуществления парного анализа:

Шаг 1. Вводим на лист Excel столбцы опытных данных в виде таблицы 2.9. Здесь N - число проведенных опытов; n - число входных величин $\{x_1; x_2; \dots x_n\}$ ($n \leq 15$); x_{ij} - значение независимой переменной x_j в i -ом опыте; y_i - значение зависимой переменной y в i -ом опыте.

Таблица 2.9

Номер опыта	x_1	x_2	----	x_n	y
1	x_{11}	x_{12}	----	x_{1n}	y_1
2	x_{21}	x_{22}	----	x_{2n}	y_2
----	----	----	----	----	----
N	x_{N1}	x_{N2}	----	x_{Nn}	y_N

Шаг 2. Выделяем на листе Excel массив ячеек для вывода на него массива результатов работы программы ЛИНЕЙН (таблица 2.10).

Таблица 2.10

m_n	m_{n-1}	m_2	m_1	b
se_n	se_{n-1}	se_2	se_1	se_b
R^2	se_y				
F	df				
SS_{reg}	SS_{resid}				

ШАГ 3. Открываем окно программы и заполняем содержащиеся в нем 4 окошка.

- В первое из них «известные значения y » вносим столбец y_i .
- Во второе «известные значения x » вносим, согласно таблице 2.9, массив значений величин $\{x_1; x_2; \dots x_n\}$.
- Третье окошко «конст» предназначено для внесения в него логического значения ИСТИНА или ЛОЖЬ. Это значение указывает, требуется ли, чтобы константа b была равна 0.

Если аргумент **конст** имеет значение ИСТИНА или опущен, то параметр b уравнения регрессии (2.82) отыскивается наряду с остальными коэффициентами этого уравнения. Если **конст** имеет значение ЛОЖЬ, то b изначально полагается равным 0.

- Четвертое окошко «статистика» тоже предназначено для внесения в него логического значения ИСТИНА или ЛОЖЬ. Здесь это значение указывает, требуется ли рассчитывать и выводить на экран дополнительную статистику по регрессии. Если аргумент статистика имеет значение ИСТИНА или опущен, то функция ЛИНЕЙН выводит на экран (возвращает) полную регрессионную статистику, указанную в таблице 2.10. А если мы в это окошко вводим ЛОЖЬ, то программа возвращает только первую строку таблицы 2.10, то есть только коэффициенты регрессии.

Шаг 4. Нажимаем клавиатурную комбинацию **Shift-Ctrl-Enter**, и получаем то, что заказывали – полный или усеченный массив 2.10 результатов работы программы.

Смысл выводимых чисел тот же, что и для парной регрессии:

- $\{se_n; se_{n-1}; \dots se_2; se_1; se_b\}$ - средние квадратические отклонения коэффициентов регрессии $\{m_n; m_{n-1}; \dots m_1; b\}$ от их истинных значений (стандартные ошибки коэффициентов). По их величине этих ошибок можно судить о значимости или незначимости коэффициентов, сравнивая абсолютную величину коэффициентов с их стандартными ошибками – точно так же, как это делается для коэффициентов уравнения парной регрессии.

- R^2 - коэффициент детерминации;
- se_y - оценка среднего квадратического отклонения экспериментальных значений y_i величины y от вычисленных по уравнению регрессии значений y_i^* этой величины. Она подсчитывается по формуле:

$$se_y = \sqrt{\frac{\sum_{i=1}^N (y_i - y_i^*)^2}{N - q}}, \quad (2.83)$$

где q - количество коэффициентов уравнения регрессии: $q = n + 1$, если $b \neq 0$, и $q = n$, если $b = 0$.

- $$ss_{reg} = S_{reg} = \sum_{i=1}^N (y_i^* - \bar{y})^2 \text{ и } ss_{resid} = S_{ocm} = \sum_{i=1}^N (y_i - y_i^*)^2 \quad (2.84)$$

- регрессионная и остаточная суммы. Они те же, что и в (2.61), и нужны для того же - для исследования уравнения регрессии на адекватность. Через них по формуле (2.64*) подсчитывается коэффициент детерминации R^2 , а через него, по формуле (2.65), находится экспериментальное значение $F = f_{эксн}$ критерия Фишера-Снедекора.

- $df = k_2 = n - q$ - второе из двух чисел степеней свободы F - критерия (первое из них $k_1 = q - 1$), где q - число коэффициентов уравнения регрессии. По заданному уровню значимости α и числам k_1 и k_2 находится табличное критическое значение $f_{крит}(\alpha; k_1; k_2)$ этого критерия. И если при сравнении $f_{эксн}$ с $f_{крит}$ окажется, что $f_{эксн} > f_{крит}$, то гипотеза об адекватности уравнения регрессии отвергается. А если $f_{эксн} < f_{крит}$, то принимается.

Примечание 1. Переменные $\{x_1; x_2; \dots x_n\}$, входящие в уравнение регрессии (2.82), могут иметь разную размерность. А тогда и коэффициенты $\{m_1; m_2; \dots m_n\}$ перед ними будут иметь разную размерность, что не позволяет сравнивать их между собой ни по величине, ни по степени влияния соответствующих им переменных на зависимую переменную y . Чтобы это стало возможным, разделим обе части уравнения (2.82) на среднеквадратичную погрешность se_y величины y . В результате обе части уравнения станут безразмерными, и это уравнение можно затем записать в виде:

$$\frac{y^*}{se_y} = m'_n \cdot \frac{x_n}{s_{x_n}} + m'_{n-1} \cdot \frac{x_{n-1}}{s_{x_{n-1}}} + \dots + m'_1 \cdot \frac{x_1}{s_{x_1}} + \frac{b}{se_y}. \quad (2.85)$$

Здесь s_{x_j} - среднеквадратическое отклонение переменной x_j от её среднего значения (эти характеристики переменных легко находятся из столбцов их значений в таблице (2.9)), а

$$m'_j = m_j \cdot \frac{s_{x_j}}{se_y} \quad (j = 1, 2, \dots n) \quad (2.86)$$

- безразмерные *стандартизованные коэффициенты регрессии*. Каждый из них, согласно равенству (2.85), показывает, на сколько величин se_y изменится в среднем зависимая переменная y , если переменная x_j изменится на одно своё s_{x_j} . По величине стандартизованных коэффициентов регрессии можно судить по степени влияния соответствующих им переменных на зависимую переменную y : чей стандартизованный коэффициент больше, та переменная и сильнее влияет на y .

Примечание 2. Программу ЛИНЕЙН, строящую линейное уравнение регрессии вида (2.82), можно использовать для построения линейных по коэффициентам, но нелинейных по входным переменным $\{x_1; x_2; \dots x_n\}$ уравнений регрессии вида

$$y^* = m_n f_n(x_n) + m_{n-1} f_{n-1}(x_{n-1}) + \dots + m_1 f_1(x_1) + b, \quad (2.87)$$

Для этого нужно положить $\{x_1^* = f_1(x_1); x_2^* = f_2(x_2); \dots x_n^* = f_n(x_n)\}$. И уравнение (2.87) станет полностью линейным.

2.10 Пример построения множественной регрессии в Excel

Пример. Данные о сменной добыче угля на одного рабочего (переменная y – измеряется в тоннах), мощности пласта (переменная x_1 – измеряется в метрах) и уровне механизации работ в шахте (переменная x_2 – измеряется в процентах), характеризующие процесс добычи угля в 10 шахтах, приведены в таблице 2.11:

Таблица 2.11

Номер шахты	x_1	x_2	y
1	8	5	5
2	11	8	10
3	12	8	10
4	9	5	7
5	8	7	5
6	8	8	6
7	9	6	6
8	9	4	5
9	8	5	6
10	12	7	8

Предполагая, что между переменными y, x_1, x_2 существует линейная зависимость, найти аналитическое выражение этой зависимости, т.е. построить уравнение линейной регрессии $y^* = m_2 x_2 + m_1 x_1 + b$.

Решение. Используя программу ЛИНЕЙН, получаем следующий массив выходных данных:

Таблица 2.12

$m_2 = 0,367$	$m_1 = 0,854$	$b = -3,539$
$se_2 = 0,243$	$se_1 = 0,221$	$se_b = 1,907$
$R^2 = 0,811$	$se_y = 0,951$	
$F = 15,08$	$df = 7$	
$ss_{reg} = 27,270$	$ss_{resid} = 6,330$	

То есть искомое уравнение регрессии имеет вид: $y^* = 0,854x_1 + 0,376x_2 - 3,539$. Коэффициент детерминации $R^2 = 0,811$ высок – построенное уравнение регрессии объясняет действием факторов x_1 и x_2 (мощностью пласта угля и уровнем механизации работ на шахтах) более 80% сменной добычи y угля на одного рабочего. Уравнение адекватно экспериментальным данным, ибо $F = f_{эксн} = 15,08$, а $f_{крит} = f(\alpha; k_1 = q - 1; k_2 = n - q) = f(0,05; 2; 7) = 4,74$, и $f_{эксн} \succ f_{крит}$.

Исследуем еще на значимость найденные коэффициенты регрессии. Так как число n проведенных опытов невелико ($n = 10$), то эту значимость будем проверять с помощью формул (2.78), то есть с использованием критерия Стьюдента. Согласно массиву 2.12,

$$\frac{|m_2|}{se_2} = 1,51; \quad \frac{|m_1|}{se_1} = 3,86; \quad \frac{|b|}{se_b} = 1,86$$

А $t_{крит} = t_{крит}(\alpha; n - q) = t_{крит}(0,05; 7) = 2,36$ - табличное критическое значение критерия Стьюдента. И так как лишь $\frac{|m_1|}{se_1} \succ t_{крит}$, то значимым при уровне значимости $\alpha = 0,05$ следует признать лишь коэффициент m_1 , связанный с фактором x_1 - мощностью пласта угля. А два других коэффициента m_2 и b можем считать незначимыми и положить их равными нулю. Тогда уравнение регрессии сильно упростится и примет вид: $y^* = 0,854x_1$.

Заметим, что при обнулении незначимого коэффициента перед переменной пропадает в уравнении регрессии и сама эта переменная: она становится для зависимой переменной y незначимой. В рассматриваемом примере при обнулении коэффициента m_2 перед переменной x_2 в уравнении регрессии пропадает и эта переменная – уровень механизации шахт. Тем самым она попадает в ранг одной из случайных помех.

Примечание. Принимая гипотезу о незначимости коэффициентов уравнения регрессии, мы можем совершить ошибку второго рода – принять эту гипотезу, когда на самом деле она неверна. То есть когда на самом деле эти коэффициенты значимы. Так как вероятность этой ошибки неизвестна, и при малом уровне значимости она может быть велика, то, обнуляя незначимые коэффициенты, мы можем ошибиться. Чтобы эту ошибку не допускать, обычно всё-таки оставляют в уравнении регрессии и незначимые коэффициенты, хотя они и утяжеляют уравнение. Например, оставляя в нашем уравнении незначимый коэффициент m_2 , мы тем самым оставляем в уравнении регрессии и переменную x_2 - уровень механизации шахт. И это более логично, чем исключать эту переменную, то есть считать, что уровень механизации шахт не имеет существенного влияния на сменную добычу угля на одного рабочего (рабочие с этим

вряд ли согласятся). Дополнительное исследование большего количества шахт значимость переменной x_2 , скорее всего, подтвердит. А вот незначимый свободный коэффициент b можно и обнулить. Тем более, что уравнение регрессии примет более логичный вид $y^* = 0,854x_1 + 0,376x_2$, чем с ненулевым коэффициентом, ибо в нем при $x_1 = 0$ и $x_2 = 0$ и расчетное значение y^* величины y будет равно нулю, что вполне естественно.

Упражнения

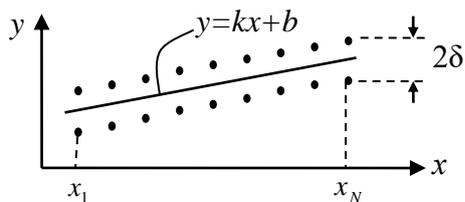


Рис. 2.3

1. Допустим, что при исследовании корреляционной зависимости y от x получено $2N$ точек корреляционного поля, расположенных симметрично относительно некоторой прямой $y = kx + b$ и удаленных от неё в направлении оси oy на $\pm\delta$ (см. рис. 2.3). Требуется:

- 1) Найти сглаживающее уравнение регрессии $\bar{y}_x^* = f^*(x)$;
- 2) Найти коэффициент линейной корреляции $\rho = \rho_{yx}$;
- 3) Найти корреляционное отношение η_{yx} ;
- 4) Найти коэффициент детерминации d_{yx} ;
- 5) Исследовать вопрос об адекватности экспериментальным данным найденного уравнения регрессии.

Ответ: 1) $\bar{y}_x^* = kx + b$ - точное сглаживающее уравнение регрессии;

$$2) \rho = \rho_{yx} = \frac{1}{\sqrt{1 + \left(\frac{\delta}{k\sigma_x}\right)^2}}; \quad 3) \eta_{yx} = \rho_{yx}; \quad 4) d_{yx} = \rho_{yx}^2 = \frac{1}{1 + \left(\frac{\delta}{k\sigma_x}\right)^2};$$

5) Уравнение регрессии $\bar{y}_x = kx + b$ адекватно экспериментальным данным при любом уровне значимости α .

2. При исследовании корреляционной зависимости урожайности некоторой культуры y (ц/га) от глубины орошения x (см) получена следующая корреляционная таблица:

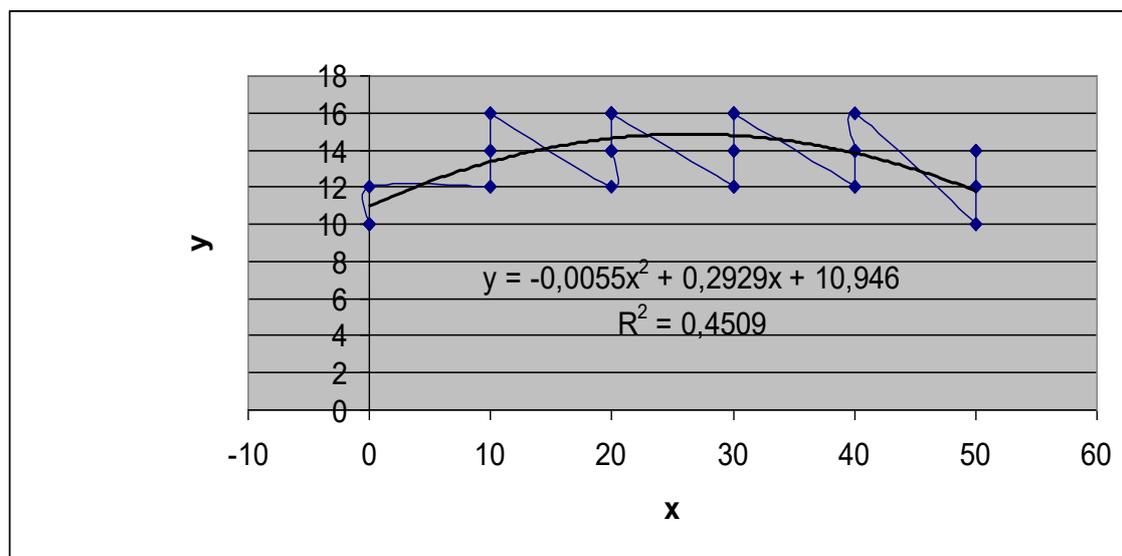
$x_i \backslash y_j$	0	10	20	30	40	50	$m_j = \sum_{i=1}^6 n_{ij}$
10	4	-	-	-	-	2	6
12	1	2	1	2	2	2	10
14	-	3	4	2	3	2	14
16	-	2	4	3	1	-	10
$n_i = \sum_{j=1}^6 n_{ij}$	5	7	9	7	6	6	$N=40$

Построить корреляционное поле, соответствующее данной корреляционной таблице. Из уравнений (2.5) - (2.7) выбрать наиболее подходящую форму для построения уравнения регрессии $\bar{y}_x^* = f^*(x)$ и построить это уравнение. Найти коэффициент детерминации $d_{yx} = R^2$. При уровне значимости $\alpha = 0,05$ проверить гипотезу об адекватности полученного уравнения регрессии.

Ответ.

С помощью функции «Линия тренда» построено параболическое уравнение регрессии $\bar{y}_x^* = f^*(x) = 10,95 + 0,293 \cdot x - 0,00553 \cdot x^2$. Коэффициент детерминации этого уравнения $R^2 = 0,450$. Уравнение адекватно экспериментальным данным, так как, согласно (2.65), $f_{эксн} = \frac{0,4509 \cdot 37}{2 \cdot (1 - 0,4509)} = 15,19$, а $f_{крит} = f_{крит}(\alpha; q - 1; n - q) = f_{крит}(0,05; 2; 37) = 3,28 < f_{эксн}$.

Примечание 1. Решение о построении параболического уравнения регрессии было принято после построения точечной диаграммы, изображающей корреляционное поле, в которое в качестве линии тренда хорошо вписалась квадратичная парабола. Запрошены и выданы на диаграмму уравнение этой параболы и её коэффициент детерминации (см. рис. ниже).



Примечание 2. Так как исходная корреляционная таблица содержит результаты повторных опытов, то мы могли исследовать полученное программой уравнение регрессии на адекватность и по-другому, опираясь на формулы (2.50) - (2.53). Но и этот путь подтверждает адекватность уравнения регрессии, ибо

$$s_{ад}^2 = 2,25; \quad s_{повт}^2 = 2,41; \quad f_{эксн} = \frac{s_{ад}^2}{s_{повт}^2} = \frac{2,246}{2,411} = 0,93,$$

$$а \quad f_{кр} = f_{кр}(\alpha = 0,05; k_1 = 6 - 3 = 3; k_2 = N - n = 40 - 6 = 34) = 2,88.$$

И так как $f_{\text{эксн}} < f_{\text{кр}}$, то найденное уравнение регрессии можно считать адекватным экспериментальным данным.

3. Исследовалась корреляционная зависимость некоторой величины z от двух других величин x и y . Экспериментальные данные по 20 проведенным опытам приводятся ниже в таблице:

i	x_i	y_i	z_i		i	x_i	y_i	z_i
1	25,5	17,3	85,5		11	54,9	25,0	60,2
2	34,1	19,8	81,7		12	32,8	28,0	74,8
3	37,3	30,1	71,7		13	55,5	33,9	63,4
4	44,7	31,9	62,7		14	41,5	16,1	74,2
5	44,6	38,3	66,4		15	41,5	21,0	71,6
6	41,0	26,5	70,6		16	52,7	28,7	60,0
7	49,5	36,2	65,0		17	37,9	20,3	81,1
8	45,1	21,0	72,8		18	43,9	19,9	71,5
9	56,5	29,5	67,6		19	35,0	22,6	77,2
10	35,4	29,4	90,0		20	27,8	20,1	91,2
					Σ	837,2	515,6	1459,2

1) С помощью функции ЛИНЕЙН построить для этих данных уравнение регрессии $z^* = f^*(x; y)$ в линейной форме $z^* = ax + by + c$ и исследовать его на адекватность экспериментальным данным.

Ответ: Искомое уравнение имеет вид $z^* = 112,44 - 0,8045x - 0,2253y$. Оно адекватно экспериментальным данным, так как $f_{\text{эксн}} > f_{\text{крит}}$, ибо $F = f_{\text{эксн}} = 25,37$, а $f_{\text{крит}} = f(0,05; 2; 17) = 3,59$.

ГЛАВА 2 ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ

§1. Основы планирования экспериментов

Эксперимент занимает центральное место в науке. Однако для получения требуемого эффекта от эксперимента он должен быть правильно спланирован, а его результаты математически корректно и с наибольшим по информации выходом обработаны. Решение этой проблемы указывает **математическая теория планирования эксперимента**. Эту теорию изложим, опираясь, главным образом, на монографию [1].

Основой математической теории планирования эксперимента является **стохастическое моделирование**, приводящее к уравнениям регрессии, описывающим изучаемый объект или процесс. А конечной целью планируемых и осуществляемых по разработанным планам экспериментов является поиск условий, при которых исследуемый процесс будет протекать оптимально.

1. Описание объекта исследования

Для описания объекта исследования удобно пользоваться представлением о кибернетической системе, которую называют «черный ящик» (рис.1). Здесь

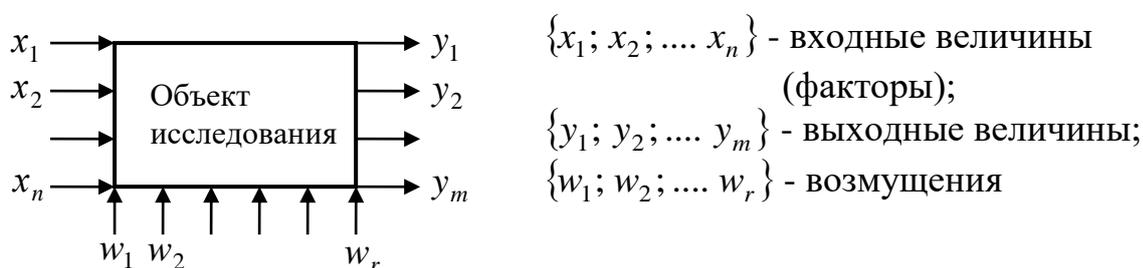


Рис.1

Факторы.

Факторами $\{x_1; x_2; \dots x_n\}$ называют воздействующие на ход процесса управляемые переменные величины. Каждый фактор x_i ($i = 1, 2, \dots n$) может принимать несколько (не менее двух) различных значений x_{ij} ($j = 1, 2, \dots$), которые называются уровнями фактора. Фактор считается заданным, если заданы все его возможные уровни. В частности, для каждого фактора должны быть указаны его предельные возможные значения.

Предполагается, что уровни всех факторов могут поддерживаться постоянными на всем протяжении эксперимента с высокой точностью. Ясно, что разница между двумя соседними уровнями фактора должна быть существенно больше ошибки фиксации уровня фактора, иначе соседние уровни фактора окажутся неразличимыми.

К факторам и их уровням предъявляются **следующие требования**:

1. Список всех существенных факторов должен быть полным. Если экспериментатор из-за боязни усложнения задачи составляет неполный список факторов, то это может привести к малоэффективному и даже бессмысленному исследованию. Поэтому лучше включить в список все, даже, на первый взгляд, несущественные факторы, чем пропустить один существенный. Попавшие же в список несущественные факторы в дальнейшем можно отсеять (как это делается – об этом будет сказано ниже).

2. Факторы должны допускать количественную оценку. Здесь оговаривается лишь принципиальная возможность этого. В частности, количественной мерой качественного фактора (например, вида сырья) могут быть порядковые номера, которые присваиваются разновидностям используемого сырья.

3. Факторы должны быть независимы друг от друга. Это значит, что любая комбинация уровней факторов может быть осуществлена в эксперименте.

Возмущения.

Возмущения $\{w_1; w_2; \dots w_r\}$ являются случайными неуправляемыми факторами, неизбежными в каждом эксперименте и влияющими на точность итоговых результатов. Их влияние оценивается при обработке результатов экспериментов.

Выходные величины.

Выходные величины $\{y_1; y_2; \dots y_m\}$ определяются по неизвестным заранее законам входными величинами $\{x_1; x_2; \dots x_n\}$ и возмущениями $\{w_1; w_2; \dots w_r\}$, и характеризуют параметры изучаемого объекта. Выходные величины должны удовлетворять **следующим требованиям:**

1. Выходные параметры должны быть *непосредственно измеряемыми или легко вычисляемыми* по данным эксперимента числовыми величинами.

2. Выходные величины должны *однозначно определяться* заданием значений входных величин (факторов). То есть при повторении эксперимента должна повториться и совокупность значений выходных переменных (в рамках неизбежного разброса, вызванного действием возмущений).

2. Цели эксперимента. Параметр оптимизации

Выходными параметрами $\{y_1; y_2; \dots y_m\}$ определяется по заданным математическим соотношениям еще один набор числовых величин $\{z_1; z_2; \dots z_s\}$, которые называются *параметрами оптимизации изучаемого объекта*. Значения параметров оптимизации и характеризуют качество объекта. Именно в рамках значений параметров оптимизации формулируются цели исследования.

Например, требуется добиться максимума по z_1 при минимуме по z_2 и при заданных ограничениях по z_3 , и т.д. При этом ставится задача установить, при каких значениях входных параметров $\{x_1; x_2; \dots x_n\}$ всё это будет достигаться.

К параметрам оптимизации предъявляются **следующие требования:**

1. Желательно, чтобы параметр оптимизации был *один*. Если же постановка задачи предусматривает оптимизацию сразу нескольких параметров, то

идут по одному из двух путей. Либо выбирают из них один, наиболее важный, в качестве главного параметра оптимизации, а на другие накладывают ограничения (рамки). Либо строят *обобщенный параметр оптимизации*, достигающий своего экстремума при условии достижения своих экстремумов каждым из параметров оптимизации. Например, если имеется три параметра оптимизации $\{z_1; z_2; z_3\}$, причем требуется добиться максимумов по z_1 и z_2 при условии минимума по z_3 , то в качестве обобщенного параметра оптимизации может быть взят параметр $z = \alpha_1 z_1 + \alpha_2 z_2 - \alpha_3 z_3$, где $\{\alpha_1; \alpha_2; \alpha_3\}$ - положительные весовые множители, характеризующие важность оптимизации соответствующего параметра $\{z_1; z_2; z_3\}$. В частности, в предположении одинаковой важности всех параметров оптимизации все α_i следует положить равными единице.

2. Параметр оптимизации должен быть *комплексным технико-экономическим критерием*, способным как можно более всесторонне характеризовать качество объекта.

3. Параметр оптимизации должен выражаться *одним числом*, однозначно определяемым набором значений входных величин (факторов). Указанная однозначность понимается в статистическом смысле. То есть каждому набору значений факторов должно соответствовать одно значение параметра оптимизации (с точностью до ошибки эксперимента, связанной с наличием озмущений).

4. Ошибка определения параметра оптимизации не зависит от значений факторов, то есть является одной и той же при любом наборе уровней факторов.

3. Целевая функция. Значимость и незначимость факторов

Ключевым моментом математической теории планирования эксперимента является получение с помощью специальной серии немногочисленных экспериментов аналитической зависимости параметра оптимизации z от существенных факторов $\{x_1; x_2; \dots x_n\}$, адекватной исследуемому объекту. Указанная зависимость строится статистическими методами в виде *уравнения регрессии* той или иной сложности, и называется эта зависимость *функцией отклика* или *целевой функцией*. По найденной целевой функции определяются численные значения факторов, при которых параметр оптимизации достигает своего оптимума (экстремума). При этом используется классическая схема нахождения экстремумов функций нескольких переменных.

В заключение проводится *серия контрольных опытов* в окрестности расчетной точки оптимума. Если такая контрольная серия показывает, что оптимум действительно достигнут, то исследование считается законченным, ибо его цели достигнуты. Если же контрольные опыты дают неудовлетворительный результат, то найденное уравнение регрессии (целевая функция) требует уточнения. Для чего делается новая серия опытов в окрестности наилучшей точки, найденной в предыдущей серии опытов. По результатам этой новой се-

рии опытов уточняется целевая функция, а уж затем по ней определяется искомая точка оптимума.

По степени влияния факторов на параметр оптимизации вводится понятие **значимости факторов**.

Фактор x_i считается значимым (существенным), если он *существенно* влияет на параметр оптимизации. Это значит, что изменение его уровней (варьирование) приводит к изменению параметра оптимизации, существенно большему разбросу значений параметра оптимизации, вызванного действием случайных факторов $\{w_1; w_2; \dots w_r\}$. Ясно, что несущественно влияющие на параметр оптимизации факторы можно не учитывать как факторы, заменив их постоянными усредненными значениями.

Задача выявления значимых факторов и их расстановка по степени влияния (ранжировка) является одной из первых в ряду задач, решение которых предусматривает математическое планирование эксперимента. Указанная задача является основной для так называемого *дисперсионного анализа*. Но её же можно решить и в рамках регрессионного анализа с помощью стандартизованных коэффициентов регрессии (см. §2, формула (2.86)). Дополнительным достоинством регрессионного анализа является то, что параллельно с отсеиванием незначимых факторов и ранжировкой значимых устанавливается вид целевой функции.

4. Подготовка к проведению экспериментов

Прежде, чем переходить к планированию опытов, следует собрать и изучить всю известную ранее (априорную) информацию о поведении изучаемого объекта. Например, из предыдущих опытов, из литературы, и т.д. При этом все собранные данные помогут яснее представить себе цели и пути решения поставленной задачи. Собранную априорную информацию удобно представить в виде анкеты, приводимой ниже.

Анкета для сбора априорной информации

1. Постановка задачи, выбор параметров оптимизации

1. Краткое описание процесса, объекта.
2. Формулировка цели исследования (если задач несколько - проранжировать их по степени важности).
3. Выбор параметров оптимизации. Следует заполнить приводимую ниже таблицу, включив в неё все необходимые параметры оптимизации и указав в примечании сравнительную важность каждого из них.

Таблица 1

Номер параметра оптимизации	Название	Размерность	Область определения	Точность	Примечание

4. Желаемый результат. Какой результат будет считаться отличным, хорошим, удовлетворительным, неудовлетворительным.

II. Выбор факторов

1. Список всех подозреваемых факторов, которые могут влиять на процесс.
2. Список факторов, включаемых в реальный эксперимент.
3. Существует ли возможность установления значения фактора на любом заданном уровне?
4. Сохраняются ли заданные значения уровней факторов в течение опыта?
5. Могут ли некоторые комбинации уровней факторов привести к остановке процесса (например, к поломке, нетехнологичности, и т.д.)?

III. Число опытов

1. Желаемое число опытов, ограничения на число опытов.
2. Желаемый срок проведения исследования.
3. Примерная длительность одного опыта.
4. Стоимость и затраты труда при проведении одного опыта.
5. Возможность выполнения параллельных (контрольных) опытов и их желаемое число.
6. Желаемая стратегия проведения опытов (например, по одному в день, и т.д.).

IV. Учет априорной информации

1. Условия и результаты, достигнутые ранее при изучении аналогичных процессов.

§2. Полные ортогональные планы эксперимента типа 2^n

1. Выбор целевой функции и уровней варьирования факторов

Если $\{x_1; x_2; \dots x_n\}$ - полный список факторов, то целевая функция есть функция этих факторов. Аналитический вид этой функции заранее, как правило, неизвестен. Предполагается лишь аналитичность этой функции, то есть разложимость её в степенной ряд (в ряд Тейлора):

$$z = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^n \sum_{j=1}^n b_{ij} x_i x_j + \dots \quad (2.1)$$

Приближенное выражение для искомой целевой функции строят в виде отрезка ряда (2.1), причем количество удерживаемых членов зависит от плана эксперимента.

На начальном этапе обычно используют двухуровневые планы эксперимента, когда каждый фактор варьируется на двух уровнях, симметричных относительно некоторого исходного (нулевого) уровня. Два этих уровня называют верхним и нижним. А именно, если x_i^0 - исходный (нулевой) уровень фактора x_i , то верхний x_i^+ и нижний x_i^- уровни находятся по формулам:

$$\begin{aligned} x_i^+ &= x_i^0 + \Delta x_i \\ x_i^- &= x_i^0 - \Delta x_i \end{aligned} \quad (i = 1, 2, \dots, n) \quad (2.2)$$

Величина Δx_i называется интервалом варьирования i - го фактора. Естественно, что нулевой уровень фактора и интервал варьирования должны быть выбраны таким образом, чтобы верхний и нижний уровни фактора не выходили за границы области определения фактора.

Точка $M_0(x_1^0; x_2^0; \dots x_n^0)$ n -мерного факторного пространства называется центром плана эксперимента, а n -мерный параллелепипед, определяемый неравенствами

$$x_i^- \leq x_i \leq x_i^+ \quad (i = 1, 2, \dots, n) \quad (2.3)$$

называется областью эксперимента.

Если объект исследовался ранее, то в качестве центра плана эксперимента выгоднее всего взять точку n -мерного факторного пространства, в которой параметр оптимизации имел наилучшее значение. Впрочем, это возможно, если только ни один из факторов в указанной точке не имеет своего предельного значения. В противном случае производят сдвиг от указанной точки к другой, близкой ей – такой, чтобы ни один фактор в этой точке (центре плана) не имел своего предельного значения.

Если же объект изучен слабо или вообще не изучен, то в качестве центра плана эксперимента обычно берут точку факторного пространства, имеющую своими координатами середины областей определения факторов.

Для интервалов варьирования принято условное разделение их на три класса: узкий, средний и широкий. Если интервал варьирования составляет менее 10% от области определения фактора, то интервал варьирования считается узким. Если больше 30% - широким. А если от 10% до 30% - средним.

Выбор того или иного интервала варьирования каждого фактора при планировании эксперимента – задача непростая, так как её решение зависит от многих причин – точности фиксирования факторов, степени нелинейности целевой функции, точности фиксирования значений параметра оптимизации, близости центра плана к искомой точке оптимума, и т.д. То есть удачный выбор интервалов варьирования факторов требует определённой предварительной информации об исследуемом объекте. Если же объект изучен слабо, то обычно по всем факторам выбирают средний интервал варьирования. В дальнейшем, после производства экспериментов и их обработки, делают выводы о пригодности выбранных интервалов варьирования. И при необходимости их увеличивают или, наоборот, уменьшают.

Для удобства планирования факторы *кодируют*. А именно, вместо натуральных значений факторов x_i вводят их кодированные значения x_i^* по формулам:

$$x_i^* = \frac{x_i - x_i^0}{\Delta x_i}; \quad x_i = x_i^0 + x_i^* \cdot \Delta x_i \quad (i = 1, 2, \dots, n) \quad (2.4)$$

Тогда верхнему уровню x_i^+ фактора x_i соответствует кодированное значение +1, а нижнему уровню значение -1. В дальнейшем будем полагать, что факторы x_i уже кодированы. При этом знак (*) будем опускать.

2. Матрица планирования

Перейдем теперь непосредственно к плану эксперимента. Для применения регрессионного анализа наиболее удобными оказываются так называемые полные ортогональные планы эксперимента типа 2^n и их дробные реплики.

Остановимся сначала на полном ортогональном плане эксперимента типа 2^n . Этот план называется полным, так как он предусматривает реализацию всех возможных сочетаний уровней факторов. Общее их число для n двухуровневых факторов, очевидно, равно 2^n , откуда и название плана. Понятие ортогональности плана будет выяснено ниже.

Если все возможные сочетания уровней факторов представить в виде таблицы, то эта таблица будет называться матрицей планирования эксперимента типа 2^n . Например, для плана 2^2 матрица планирования будет следующей (таблица 2.1):

Таблица 2.1

Номер опыта	Факторы		Целевая функция
	x_1	x_2	
1	+1	+1	z_1
2	-1	+1	z_2
3	+1	-1	z_3
4	-1	-1	z_4

А для плана типа 2^3 соответствующая матрица представлена в таблице 2.2:

Таблица 2.2

Номер опыта	Факторы			Целевая функция
	x_1	x_2	x_3	
1	+1	+1	+1	z_1
2	-1	+1	+1	z_2
3	+1	-1	+1	z_3
4	-1	-1	+1	z_4
5	+1	+1	-1	z_5
6	-1	+1	-1	z_6
7	+1	-1	-1	z_7
8	-1	-1	-1	z_8

И вообще, для построения матрицы планирования типа 2^n при заданном n поступают так:

- столбец значений первого фактора длиной 2^n составляют из чередующихся построчно значений $+1$ и -1 ;
- в столбце значений второго фактора той же длины чередуют два раза $+1$, два раза -1 ;
- в столбце значений третьего сначала 4 раза подряд пишут $+1$, затем 4 раза подряд -1 , и т.д.

Легко видеть, что столбец последнего, n -го фактора будет наполовину (верхнюю) состоять из $+1$, а наполовину (нижнюю) из -1 .

Каждая строка матрицы планирования определяет условия проведения соответствующего опыта. Нумерация опытов дана условно. На практике последовательность проведения опытов выбирается случайно – например, по таблице случайных чисел. Эта операция называется рандомизацией матрицы планирования. Рандомизация проводится с целью исключения возможного влияния опытов друг на друга.

3. Целевая функция

Двухуровневый эксперимент типа 2^n связан с поиском уравнения регрессии для целевой функции z в виде неполной квадратичной зависимости вида

$$z = b_0 + \sum_{i=1}^n b_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n b_{ij} x_i x_j \quad (2.5)$$

В частности, для двух факторов ($n = 2$) эта зависимость примет вид:

$$z = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 \quad (2.6)$$

Для трех факторов ($n = 3$) будем иметь:

$$z = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 \quad (2.7)$$

Таким образом, согласно (2.5), в ряду Тейлора (2.1) удерживаются лишь слагаемые первого и второго порядков, исключая квадратичные слагаемые вида $b_{ii} x_i^2$, которые не могут быть определены при планировании 2^n , когда факторы варьируются лишь на двух уровнях. Для более точного, чем (2.5), представления целевой функции применяются планирования экспериментов других, более сложных видов (ортогональное композиционное планирование второго порядка, планы типа 3^n , и др.). Но мы эти планы рассматривать не будем.

После построения целевой функции (2.5) проверяют её согласованность с результатами эксперимента, или, как говорят, её адекватность объекту исследования. Если уравнение (2.5) оказалось неадекватным, то приходится повторять эксперимент при других (меньших) интервалах варьирования факторов,

уменьшая тем самым область эксперимента. При этом особенно действенным является уменьшение интервала варьирования наиболее значимых факторов x_i - тех, при которых коэффициенты b_i велики. Если же это уравнение оказалось адекватным, то его используют для поиска искомого оптимума.

Коэффициенты b_i ($i = 1, 2, \dots, n$) линейной части функции (2.5) определяют характер и степень влияния факторов x_i на параметр оптимизации z . В частности, положительность коэффициента b_i указывает на то, что с ростом значений фактора x_i растет и параметр оптимизации. И наоборот, при отрицательном b_i с ростом x_i параметр оптимизации убывает. Чем больше по модулю коэффициент b_i , тем сильнее влияние фактора x_i на параметр оптимизации z . Или, как говорят, тем более *значим* фактор x_i . При достаточно малых по абсолютной величине значениях коэффициента b_i фактор x_i признается незначимым, то есть несущественно влияющим на параметр оптимизации, и слагаемое $b_i x_i$ из уравнения регрессии исключается. Значение незначимого фактора в последующих опытах фиксируется по его среднему значению.

Таким образом, найденные коэффициенты b_i при линейных членах уравнения регрессии (2.5), при условии значимости этого уравнения, позволяют произвести ранжировку факторов по степени их значимости. А оказавшиеся незначимыми факторы отбросить, переведя их в ранг помех $\{w_1; w_2; \dots, w_r\}$.

Коэффициенты b_{ij} характеризуют более тонкие эффекты парного взаимодействия факторов x_i и x_j , то есть влияние на параметр оптимизации z каждого из этих факторов в зависимости от того, какое значение имеет другой фактор. При значимости того или иного коэффициента b_{ij} соответствующее слагаемое $b_{ij} x_i x_j$ уравнения регрессии (2.5) учитывается наряду с линейными слагаемыми.

Применение регрессионного анализа при полном плане типа 2^n позволяет оценить эффекты всех факторов и эффекты любых их парных взаимодействий (и даже, при необходимости, эффекты тройных взаимодействий типа $x_i x_j x_k$ ($i \neq j \neq k$), четверных и т.д.). Что, впрочем, обычно не делается из-за отсутствия ясной интерпретации таких взаимодействий и обычной их незначимости.

Выявление тех или иных незначимых взаимодействий позволяет сократить необходимое количество экспериментов, проводя вместо $N=2^n$ в 2, или в 4, или в 8 и т.д. раз меньше этих опытов (так называемые дробные реплики типа 2^{n-p}). Об этом будет говориться ниже, в §5.

§3. Расширенная матрица планирования эксперимента

Статистическая обработка полных планов типа 2^n или их дробных реплик предусматривает оценку разброса значений параметра оптимизации z , возникающую при повторении экспериментов из-за действия случайных факторов. Эта оценка может быть известна заранее (например, из предыдущих опытов,

или из литературных данных, или определяется точностью измерительного прибора). В противном случае для её определения осуществляют повтор проводимых опытов.

Наиболее удобными при обработке являются опыты, повторяемые каждый одинаковое число раз (обычно 2-4 раза). Для повторных опытов, как и для основных, проводят рандомизацию.

Все результаты проведенных опытов вносят в расширенную матрицу планирования, содержащую не только уровни факторов в опытах, но и взаимодействия, значимость которых требуется оценить. Кроме того, расширенная матрица содержит столбец рандомизации, определяющий порядок проведения опытов, и столбцы для записи значений параметра оптимизации.

Для случая трех факторов при учете всех возможных их взаимодействий она будет, например, такой (таблица 3.1):

Таблица 3.1

№ опыта в матрице	Рандомизация		Факторы			Взаимодействия				z			Примечание
	1-ый опыт	2-ой опыт	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	z_{i1}	z_{i2}	\bar{z}_i	
1	2	9	+1	+1	+1	+1	+1	+1	+1	z_{11}	z_{12}	\bar{z}_1	
2	6	13	-1	+1	+1	-1	-1	-1	-1	z_{21}	z_{22}	\bar{z}_2	
3	1	15	+1	-1	+1	-1	+1	-1	-1	z_{31}	z_{32}	\bar{z}_3	
4	7	10	-1	-1	+1	+1	-1	+1	+1	z_{41}	z_{42}	\bar{z}_4	
5	3	16	+1	+1	-1	+1	-1	-1	-1	z_{51}	z_{52}	\bar{z}_5	
6	8	14	-1	+1	-1	-1	+1	+1	+1	z_{61}	z_{62}	\bar{z}_6	
7	4	12	+1	-1	-1	-1	-1	+1	+1	z_{71}	z_{72}	\bar{z}_7	
8	5	11	-1	-1	-1	+1	+1	-1	-1	z_{81}	z_{82}	\bar{z}_8	

Матрица 3.1 составлена для случая двукратной повторности каждого опыта. При большей повторности опытов матрицу следует дополнить соответствующими столбцами. В матрице указан порядок проведения всех опытов (столбец «рандомизация»). Номера опытов взяты наугад из таблицы случайных чисел. Условия проведения каждого опыта указываются значениями факторов x_1, x_2, x_3 . Результаты измерений параметра оптимизации z в каждом опыте и средние его значения \bar{z}_i заносятся в соответствующие столбцы матрицы.

Расширенная матрица плана 2^n , пример которой для случая $n = 3$ дает таблица 3.1, обладает рядом важных свойств, одни из которых очевидны, а доказательство других опустим. Эти свойства следующие:

Свойства вектор-столбцов расширенной матрицы планирования

1. Каждый вектор-столбец факторов и их взаимодействий содержит одинаковое число компонент $+1$ и -1 .

2. Среди вектор-столбцов факторов и их взаимодействий нет одинаковых.

3. Вектор-столбцы всех факторов и всех их взаимодействий попарно ортогональны. Указанное свойство ортогональности означает, что перемножая построчно элементы любых двух столбцов факторов и взаимодействий и складывая получившиеся числа, в итоге получим нуль (скалярное произведение этих вектор-столбцов равно нулю). Именно в силу ортогональности столбцов факторов и взаимодействий планы типа 2^n называются ортогональными.

Ортогональность вектор-столбцов факторов и их взаимодействий означает еще их линейную независимость друг от друга. То есть ни один из этих вектор-столбцов нельзя выразить через линейную комбинацию остальных.

4. Ортогональная матрица планирования типа 2^n может содержать не более 2^n ортогональных вектор-столбцов.

5. Ортогональный план типа 2^n обладает свойством *ротатабельности*. Это значит, что точность предсказания значений параметра оптимизации для линейной модели (без учета взаимодействий) одинакова на равных расстояниях от центра эксперимента и не зависит от направления в факторном пространстве.

§4. Математическая обработка результатов экспериментов

Если p - количество повторных опытов, постоянное для каждой строки матрицы планирования (для каждого опыта), то коэффициенты уравнения регрессии (2.5) находятся независимо друг от друга по следующим простым формулам:

$$b_0 = \frac{\sum_{k=1}^N \bar{z}_k}{N}; \quad b_i = \frac{\sum_{k=1}^N \bar{z}_k x_{ik}}{N}; \quad b_{ij} = \frac{\sum_{k=1}^N \bar{z}_k x_{ik} x_{jk}}{N}; \quad (N = 2^n; 1 \leq i, j \leq n) \quad (2.8)$$

Здесь x_{ik} - значение фактора x_i в k -ом опыте, то есть значение фактора x_i , содержащееся в k -ой строке матрицы планирования; \bar{z}_k - среднее значение параметра оптимизации в k -ом опыте. Указанные в (2.8) суммы очень легко находить, пользуясь расширенной матрицей планирования.

Выведем формулы (2.8). Для этого подставим в выражение (2.5) для целевой функции вместо факторов x_i и их парных взаимодействий $x_i x_j$ столбцы их

значений x_{ik} и $x_{ik}x_{jk}$ из матрицы планирования, а вместо параметра оптимизации z - столбец его средних значений \bar{z}_k ($k = 1, 2, \dots, N$). Таким образом, из одного уравнения (2.5) мы получим систему из $N = 2^n$ уравнений

$$\bar{z}_k = b_0 + \sum_{i=1}^n b_i x_{ik} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n b_{ij} x_{ik} x_{jk} \quad (k = 1, 2, \dots, N = 2^n) \quad (2.9)$$

Если теперь сложить все эти N уравнений и учесть, что столбцы значений всех факторов и их взаимодействий в сумме дают нуль,

$$\sum_{k=1}^N x_{ik} = 0 \quad (i = 1, 2, \dots, n); \quad \sum_{k=1}^N x_{ik} x_{jk} = 0 \quad (i, j = 1, 2, \dots, n) \quad (2.10)$$

то придем к первому из равенств (2.8), то есть к выражению для b_0 .

Чтобы получить выражение (2.8) для b_i , умножим уравнения системы (2.9) на столбец значений x_{sk} фактора x_s и опять сложим все получившиеся уравнения. И так как скалярные произведения

$$\sum_{k=1}^N x_{ik} x_{sk} = \begin{cases} N & (i = s) \\ 0 & (i \neq s) \end{cases}; \quad \sum_{k=1}^N x_{ik} x_{jk} \cdot x_{sk} \equiv 0, \quad (2.11)$$

в силу ортогональности вектор-столбцов факторов и их взаимодействий, то в результате такой операции получим:

$$\sum_{k=1}^N \bar{z}_k x_{sk} = b_s \cdot N \Rightarrow b_i = \frac{\sum_{k=1}^N \bar{z}_k x_{ik}}{N} \quad (i = 1, 2, \dots, n),$$

Аналогично получается и последнее выражение (2.8) для коэффициентов b_{ij} .

Найдя коэффициенты $\{b_0, b_i, b_{ij}\}$ уравнения регрессии (2.5), мы тем самым получим аналитическое выражение зависимости параметра оптимизации z от факторов x_i ($i = 1, 2, \dots, n$).

Дальнейшая работа связана с с проверкой качества уравнения (2.5), то есть с его проверкой на адекватность опытным данным. А также с проверкой значимости коэффициентов $\{b_0, b_i, b_{ij}\}$ этого уравнения, то есть существенности их отличия от нуля. Тем самым проверяется существенность (значимость) влияния на параметр оптимизации тех факторов и тех взаимодействий, перед которыми эти коэффициенты стоят. Ибо если какие-то из коэффициентов уравнения (2.5) будут признаны незначимыми, то их без особого ущерба для качества уравнения (2.5) можно положить равными нулю (обнулить). В итоге уравнение (2.5) упростится: в нем исчезнут незначимые факторы и незначимые взаимодействия факторов.

Вся эта информация собирается с помощью дисперсионного анализа [7]. Этот анализ предполагает обязательное проведение повторных опытов, и желательно одинаковое их количество p для каждого опыта, что мы и предполагаем.

Приведем все необходимые формулы для такого анализа, которые, кстати, являются частными случаями формул, с которыми мы уже имели дело в пункте 2.3 параграфа 2 главы 1. Для каждой строки i расширенной матрицы планирования, то есть для каждого опыта, имеющего p повторностей, находим дисперсию s_i^2 параметра оптимизации z , характеризующую степень разброса значений z_{ik} ($k = 1, 2, \dots, p$) этого параметра в i -ом опыте:

$$s_i^2 = \frac{\sum_{k=1}^p (z_{ik} - \bar{z}_i)^2}{p-1} \quad (i = 1, 2, \dots, N) \quad (2.12)$$

Число $p-1$, стоящее в знаменателе выражения для дисперсии s_i^2 параметра оптимизации z в i -ом опыте, имеющем p повторностей, является числом степеней свободы этой дисперсии. Это число представляет собой количество независимых величин z_{ik} , участвующих в выражении для этой дисперсии. Оно на единицу меньше их общего числа p , ибо одна степень свободы тратится на вычисление их среднего значения \bar{z}_i .

Найдя дисперсии отдельных опытов s_i^2 , затем находим среднее значение этих дисперсий по всем опытам, то есть дисперсию параметра оптимизации во всём эксперименте. Её еще называют *дисперсией воспроизводимости опытов*:

$$s_{\{z\}}^2 = \frac{\sum_{i=1}^N s_i^2}{N} = \frac{\sum_{i=1}^N \sum_{k=1}^p (z_{ik} - \bar{z}_i)^2}{N(p-1)} \quad (2.13)$$

Число $N(p-1)$, стоящее в знаменателе выражения для дисперсии $s_{\{z\}}^2$ параметра оптимизации z во всем эксперименте, является числом степеней свободы этой дисперсии. Это число представляет собой количество всех независимо измеренных значений z_{ik} параметра оптимизации во всем эксперименте.

Применение дисперсионного анализа предполагает, что дисперсии s_i^2 ($i = 1, 2, \dots, N$) несущественно (незначимо) различаются между собой. В противном случае применять дисперсионный анализ для любых статистических оценок некорректно.

Для проверки этой однородности дисперсий s_i^2 применяется критерий Кочрена. Для этого находят величину

$$G = \frac{\{s_i^2\}_{\max}}{\sum_{i=1}^N s_i^2} \quad (2.14)$$

- экспериментальное значение критерия Кочрена. Выдвигается основная (нулевая) гипотеза H_0 : дисперсии s_i^2 различаются между собой незначимо. Принимается уровень значимости α , который представляет собой вероятность сделать ошибку первого рода (отвергнуть гипотезу H_0 о незначимости различий дисперсий s_i^2 , если она верная). Обычно принимают $\alpha = 0,01$ или $\alpha = 0,05$. Далее по таблице критических точек распределения Кочрена, содержащейся во многих руководствах по математической статистике, находят критическое значение $G_{кр} = G_{кр}(\alpha; k = p - 1; N = 2^n)$ критерия Кочрена. Входными параметрами этой таблицы являются числа $(\alpha; k = p - 1; N = 2^n)$. И если окажется, что $G < G_{кр}$, то дисперсии s_i^2 признаются однородными (различающимися незначимо). А если $G > G_{кр}$, то дисперсии s_i^2 различаются значимо, и тогда дальнейшая работа в рамках дисперсионного анализа будет некорректной, а потому не производится.

Будем считать, что всё в порядке: критерий Кочрена подтвердил незначимость различий дисперсий s_i^2 . Тогда по формуле (2.5) подсчитываем теоретические значения $(z_k)_{теор}$ параметра оптимизации z для всех $N = 2^n$ строк расширенной матрицы планирования эксперимента:

$$(z_k)_{теор} = b_0 + \sum_{i=1}^n b_i x_{ik} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n b_{ij} x_{ik} x_{jk} \quad (k = 1, 2, \dots, N) \quad (2.15)$$

Эти теоретические (вычисленные) значения параметра оптимизации z , вообще говоря, не будут совпадать со средними экспериментальными значениями \bar{z}_k ($k = 1, 2, \dots, N$). Находим разности между ними (невязки):

$$\Delta z_k = (z_k)_{теор} - \bar{z}_k \quad (k = 1, 2, \dots, N) \quad (2.16)$$

Чем они меньше по абсолютной величине – тем качественней будет выражение (2.5). Или, как говорят, тем адекватнее оно будет экспериментальным результатам. По невязкам Δz_k подсчитывается дисперсия адекватности уравнения регрессии (2.5) опытным данным:

$$s_{ад}^2 = \frac{\sum_{k=1}^N (\Delta z_k)^2}{N - N_0}; \quad N_0 = 1 + \frac{n(n+1)}{2} \quad (2.17)$$

Здесь N_0 - общее количество коэффициентов b_i и b_{ij} , входящих в уравнение регрессии, включая b_0 , а $N - N_0$ - число степеней свободы дисперсии s_{ad}^2 .

Затем сравниваются между собой дисперсия s_{ad}^2 и $s_{\{z\}}^2$. Это делается с помощью критерия Фишера

$$F = \frac{s_{ad}^2}{s_{\{z\}}^2}, \quad (2.18)$$

где F - случайная величина, имеющая распределение Фишера-Снедекора со степенями свободы $k_1 = N - N_0$ числителя и $k_2 = N(p - 1)$ знаменателя. По формуле (2.19) вычисляется экспериментальное значение $F_{эксн}$ критерия Фишера, а затем сравнивают его с критическим значением $F_{кр} = F_{кр}(\alpha; k_1; k_2)$ этого критерия, которое берется из таблицы критических точек этого критерия. И если окажется, что $F_{эксн} < F_{кр}$, уравнение (2.5) признают адекватным опытными данными. Следовательно, его можно использовать для аналитического исследования изучаемого объекта. А если $F_{эксн} > F_{кр}$, то уравнение (2.5) не является пригодным для этого исследования, ибо его использование может приводить к неверным оценкам.

Кроме того, подсчитываются дисперсии коэффициентов уравнения регрессии

$$s_{\{b_i\}}^2 = s_{\{b_{ij}\}}^2 = \frac{s_{\{z\}}^2}{N} \quad (1 \leq i, j \leq n), \quad (2.19)$$

имеющие, согласно (2.18), одно и то же значение, и имеющие число степеней свободы $f = N(p - 1)$, равное числу степеней свободы дисперсии $s_{\{z\}}^2$.

Заметим, что эти формулы являются следствием формул (2.79) параграфа 2 главы 1, так как для каждой переменной x_i и каждого их взаимодействия $x_i x_j$ их среднее значение $\bar{x} = 0$, а их среднее квадратическое отклонение $\sigma_x = 1$.

Наконец, подсчитываются величины

$$t_i = \frac{|b_i|}{s_{\{b_i\}}} = \sqrt{\frac{N}{s_{\{z\}}^2}} \cdot |b_i|; \quad t_{ij} = \frac{|b_{ij}|}{s_{\{b_{ij}\}}} = \sqrt{\frac{N}{s_{\{z\}}^2}} \cdot |b_{ij}|; \quad (1 \leq i, j \leq n) \quad (2.20)$$

- экспериментальные значения случайных величин t_i и t_{ij} , имеющих распределение Стьюдента с $k = N(p - 1)$ степенями свободы. Они нужны для оценки значимости коэффициентов регрессии b_i и b_{ij} уравнения (2.5). По таблице критических точек распределения Стьюдента находят критическое значение $t_{кр}(\alpha; k)$ этого распределения. И если окажется, что какое-то из эксперимен-

тальных значений (2.20) окажется меньше $t_{кр}$, то соответствующий этому значению коэффициент b_i или b_{ij} признается незначимым, и его в уравнении (2.5) можно положить равным нулю. А если величина (2.20) окажется больше $t_{кр}$, то соответствующий ей коэффициент b_i или b_{ij} признается значимым, и его в уравнении (2.5) следует сохранить. А вместе с ним в уравнении (2.5) сохранится соответствующий ему фактор x_i или взаимодействие $x_i x_j$. И чем больше по абсолютной величине окажутся сохранившиеся в уравнении (2.5) слагаемые, тем больше они и значимы.

Напомним еще раз: обоснованность применения изложенного выше регрессионного анализа к статистической обработке экспериментов обусловлена выполнением требований к факторам и параметру оптимизации, о которых говорилось выше (§1). В частности, выполнением следующего важнейшего условия: дисперсия параметра оптимизации s_i^2 (формула (2.12)) примерно одна и та же (однородна) во всех опытах.

И еще раз напомним последовательность статистической обработки расширенной матрицы планирования с помощью регрессионного анализа:

1. По формулам (2.12) производится подсчет дисперсий s_i^2 повторных опытов в каждой строке матрицы планирования ($i = 1, 2, \dots, N = 2^n$).

2. Производится проверка однородности дисперсий s_i^2 (незначимости различия между ними) с помощью критерия Кочрена.

3. Приступить к дальнейшей работе по регрессионному анализу можно лишь после того, как будет подтверждена однородность дисперсий s_i^2 . Если же дисперсии s_i^2 окажутся неоднородными, то задача усложняется. В таких случаях часто оказывается полезным изменение масштаба для параметра оптимизации z - вместо него берется некоторая его функция, например квадратный корень или логарифм.

4. Если дисперсии s_i^2 однородны, то по формуле (2.13) производится оценка дисперсии параметра оптимизации $s_{\{z\}}^2$ во всём эксперименте.

5. Определяются, согласно формулам (2.8), коэффициенты уравнения регрессии (2.5) и записывается это уравнение регрессии.

6. По уравнению регрессии вычисляются теоретические значения $(z_k)_{теор}$ ($k = 1, 2, \dots, N$) для всех $N = 2^n$ опытов (формулы (2.15)).

7. По формуле (2.17), с учетом формул (2.16), вычисляется дисперсия адекватности $s_{ад}^2$.

8. Проверяется адекватность уравнения регрессии исследуемому объекту. Для этого по формуле (2.19) вычисляется экспериментальное значение критерия Фишера $F_{эксн}$. А затем при выбранном уровне значимости α по степеням свободы $k_1 = N - N_0$ числителя и $k_2 = N(p - 1)$ знаменателя определяется по таблице критических точек критерия Фишера критическое значение его значе-

ние $F_{кр} = F_{кр}(\alpha; k_1; k_2)$. И если окажется, что $F_{эксн} \prec F_{кр}$, то уравнение регрессии (2.5) признается адекватным исследуемому объекту. А если уравнение регрессии оказывается неадекватным, то приходится проводить новую серию опытов, уменьшая интервалы варьирования факторов x_i , особенно тех, коэффициенты b_i при которых в уравнении регрессии являются наибольшими по абсолютной величине. При этом в новой серии опытов полезно взять новый центр эксперимента. А именно, в качестве нового центра следует выбрать точку факторного пространства, в которой экспериментальное значение параметра оптимизации оказалось максимальным (или минимальным, если оптимумом является минимум). При этом в новой серии опытов можно не проводить повторных опытов, которые нужны лишь для оценки дисперсии параметра оптимизации $s_{\{z\}}^2$, так как эта оценка уже была получена в первой серии опытов.

9. Если уравнение регрессии оказалось адекватным изучаемому объекту, то проверяется значимость коэффициентов уравнения регрессии. А именно, для каждого коэффициента по формулам (2.20) вычисляется соответствующее ему экспериментальное значение критерия Стьюдента $t_{эксн}$. Затем при выбранном уровне значимости и числе степеней свободы $f = N(p-1)$ определяется табличное критическое значение критерия Стьюдента $t_{кр}$. И если окажется, что $t_{эксн} \prec t_{кр}$, то соответствующий коэффициент уравнения регрессии (2.5) признается незначимым, и его можно положить равным нулю. А если $t_{эксн} \succ t_{кр}$, то значимым, и его в уравнении (2.5) оставляют. В окончательном виде уравнение регрессии (2.5) будет содержать только значимые факторы и взаимодействия этих факторов.

В заключение приведем выписки из таблиц критических значений критериев Кочрена, Фишера и Стьюдента:

Таблица 4.1.
Критические значения критерия Кочрена $G_{кр}(\alpha; k_1; k_2)$

Уровень значимости $\alpha = 0,05$			
$k_2 = N$	$k_1 = p - 1$		
	1	2	3
2	0,9985	0,9750	0,9392
4	0,9065	0,7679	0,6841
8	0,6798	0,5157	0,4377
16	0,4500	0,3190	0,2620

Таблица 4.2.

Критические значения критерия Фишера $F_{кр}(\alpha; k_1; k_2)$

Уровень значимости $\alpha = 0,05$								
$k_2 = N(p-1)$	$k_1 = N - N_0$							
	1	2	3	4	6	8	10	∞
2	18,51	19,00	19,16	19,25	19,33	19,37	19,39	19,50
4	7,71	6,94	6,59	6,39	6,16	6,04	5,96	5,63
6	5,99	5,14	4,76	4,53	4,28	4,15	4,06	3,67
8	5,32	4,46	4,07	3,84	3,58	3,44	3,34	2,93
12	4,75	3,88	3,49	3,26	3,00	2,85	2,76	2,30
16	4,49	3,63	3,24	3,01	2,74	2,59	2,49	2,01
24	4,26	3,40	3,01	2,78	2,51	2,36	2,26	1,73
48	4,04	3,19	2,80	2,56	2,30	2,14	2,03	1,45
∞	3,84	2,99	2,60	2,37	2,09	1,94	1,83	1,00

Таблица 4.3.

Критические значения критерия Стьюдента $t_{кр}(\alpha; k)$

Уровень значимости $\alpha = 0,05$								
$k = N(p-1)$	2	4	8	16	24	60	120	∞
$t_{кр}$	4,30	2,78	2,31	2,12	2,06	2,00	1,98	1,96

§5. Дробный факторный эксперимент

Из всех возможных двухуровневых планов полный ортогональный план эксперимента типа 2^n по своим свойствам, качеству добываемой информации и простоте обработки является наилучшим и применяется, если число n факторов $\{x_1; x_2; \dots x_n\}$ невелико ($n \leq 4$). Если же $n > 4$, то число $N = 2^n$ опытов полного плана 2^n с увеличением n очень быстро растет и становится чрезмерно большим.

В самом деле, для построения целевой функции (2.5) необходимо определить один коэффициент b_0 , n коэффициентов b_i и $\frac{n(n-1)}{2}$ коэффициентов b_{ij} . То есть их общее число N_* равно

$$N_* = 1 + n + \frac{n(n-1)}{2} = 1 + \frac{n(n+1)}{2} \quad (2.21)$$

А если у функции (2.5) отыскивать лишь её линейную часть (первые два слагаемых), то понадобится определить лишь коэффициенты b_0 и b_i общим числом $N_{**} = 1 + n$. Общее же число опытов $N = 2^n$, проводимых в полном факторном эксперименте, при $n > 2$ превосходит и N_* , и, тем более, N_{**} - см. таблицу 5.1:

Таблица 5.1

n	1	2	3	4	5	6	7	8	9	10
$N = 2^n$	2	4	8	16	32	64	128	256	512	1024
N_*	2	4	7	11	16	22	29	37	46	56
N_{**}	2	3	4	5	6	7	8	9	10	11

И превосходит тем значительнее, чем больше число факторов n . То есть в полном факторном эксперименте типа 2^n обнаруживается большая избыточность проводимых опытов по сравнению с количеством определяемых коэффициентов уравнения (2.5), представляющего собой аналитическое выражение для параметра оптимизации z . А так как необходимо ещё проводить и повторные опыты, то эта избыточность становится ещё больше.

В связи с этим возникает естественное желание сократить число проводимых опытов. Это и делается с помощью дробных реплик 2^{n-m} полного плана 2^n . Эти реплики представляют собой выдержки из полного плана 2^n , содержащие в 2^m меньше опытов, чем в плане 2^n . В частности, дробная реплика 2^{n-1} содержит в $2^1 = 2$ раза опытов меньше, чем план 2^n , и называется полу-репликой. Соответственно дробная реплика 2^{n-2} содержит в $2^2 = 4$ раза меньше опытов, чем полный план 2^n , и называется четверть-репликой. И т.д.

Особенно большой выигрыш при использовании дробных реплик получается при построении линейных моделей. То есть когда функция отклика (2.5) строится в линейном виде

$$z = b_0 + \sum_{i=1}^n b_i x_i \quad (2.22)$$

Для её построения требуется найти лишь $N_{**} = 1 + n$ коэффициентов b_i ($i = 0, 1, 2, \dots, n$).

Рассмотрение конкретных примеров начнем с самого простого – с полного факторного эксперимента типа 2^2 , то есть с эксперимента с двумя факторами x_1 и x_2 . Расширенная матрица планирования этого эксперимента будет иметь вид, представленный в таблице 5.2:

Таблица 5.2

Номер опыта	Факторы		Взаимодействия	Параметр оптимизации
	x_1	x_2		
			$x_1 x_2 (x_3)$	z
1	+	+	+	z_1
2	-	+	-	z_2
3	+	-	-	z_3
4	-	-	+	z_4

Пользуясь таким планированием, можно построить целевую функцию z в полном её виде:

$$z = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 \quad (2.23)$$

Данные таблицы 5.2 и использование формул (2.8) позволяют найти все четыре коэффициента $\{b_0; b_1; b_2; b_{12}\}$ этой функции. А если строить лишь линейную модель

$$z = b_0 + b_1x_1 + b_2x_2, \quad (2.24)$$

то достаточно определить лишь три коэффициента $\{b_0; b_1; b_2\}$. Для их определения достаточно провести три опыта. А оставшийся четвертый опыт можно использовать для введения и дальнейшего определения ещё одного фактора x_3 . В качестве столбца значений этого дополнительного фактора x_3 примем столбец значений взаимодействия x_1x_2 факторов x_1 и x_2 (см. таблицу 5.2). Тогда целевая функция (2.23) примет вид:

$$z = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \quad (2.25)$$

А коэффициенты $\{b_0; b_1; b_2; b_3\}$ этой функции найдутся по формулам (2.8), которые в данном случае примут вид:

$$b_0 = \frac{\sum_{k=1}^4 z_k}{4}; \quad b_i = \frac{\sum_{k=1}^4 z_k x_{ik}}{4}; \quad (i = 1, 2, 3) \quad (2.26)$$

Но здесь необходимо отметить следующее. В этой модели, когда мы взаимодействие x_1x_2 заменили фактором x_3 , парные взаимодействия $\{x_1x_2; x_1x_3; x_2x_3\}$ и тройное взаимодействие $x_1x_2x_3$ перестали быть независимыми, ибо

$$x_1x_2 = x_3; \quad x_1x_3 = x_1^2x_2 = x_2; \quad x_2x_3 = x_2^2x_1 = x_1; \quad x_1x_2x_3 = x_3^2 = 1 \quad (2.27)$$

То есть столбцы значений факторов $\{x_1; x_2; x_3\}$ матрицы эксперимента 5.2 отличимы от столбцов взаимодействий $\{x_2x_3; x_1x_3; x_1x_2\}$ соответственно. А это значит, что, определяя по формулам (2.26) коэффициенты целевой функции (2.25), мы не будем знать, чем вызвано их значение: отдельным действием соответствующих им факторов, или взаимодействием других факторов. То есть произойдет не отдельное определение действий факторов и всех возможных их взаимодействий на целевую функцию, как это происходит в полном факторном эксперименте, а определение *смешанного действия* факторов и их взаимодействий. Но если мы пренебрежем эффектом взаимодействия факторов, то коэф-

коэффициенты b_i ($i = 1, 2, 3$) линейной целевой функции (2.26) будут приближенно определять действие на неё отдельных факторов x_i ($i = 1, 2, 3$).

Конечно, эти коэффициенты $\{b_1; b_2; b_3\}$ будут точнее определять действие на целевую функцию факторов $\{x_1; x_2; x_3\}$, если их определять из полного факторного эксперимента 2^3 , содержащего 8 опытов, ибо в полном факторном эксперименте эффекты всех факторов и всех их взаимодействий определяются независимо друг от друга, они между собой не смешиваются. Но это требует вдвое больше опытов, чем те 4 опыта, которые указаны в таблице 5.2.

Отметим, что таблица 5.2 представляет собой выдержку из полного плана эксперимента, проводимого для трех факторов и представленного в таблице 3.1. План таблицы 5.2 – это полуреплика плана таблицы 3.1.

В свою очередь, полный ортогональный план типа 2^3 для трех факторов $\{x_1; x_2; x_3\}$, представленный в матрице (3.1), можно превратить в полуреплику полного ортогонального плана 2^4 , то есть в полуреплику плана для четырех факторов $\{x_1; x_2; x_3; x_4\}$. Для этого одно из взаимодействий плана 2^3 нужно принять за новый фактор x_4 . Чаще всего в качестве фактора x_4 принимают тройное взаимодействие $\{x_1x_2x_3\}$, которое обычно является наименее значимым и к тому же плохо поддается интерпретации. Но если известно, что какое-то из парных взаимодействий не имеет смысла, то новый фактор x_4 вводят вместо него. Из такого плана можно найти коэффициенты b_i ($i = 0, 1, 2, 3, 4$) линейной целевой функции

$$z = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4, \quad (2.28)$$

проведя вместо 16 лишь 8 опытов (полуреплику). Правда, коэффициенты b_i ($i = 0, 1, 2, 3, 4$) этой функции будут характеризовать влияние факторов x_i на целевую функцию не в чистом виде, а с примесью их взаимодействий. Но эта примесь – плата за существенное снижение необходимого для построения модели количества опытов.

Кстати, так как в полном ортогональном плане типа 2^3 , представленном в таблице 3.1, содержится 4 взаимодействия, то это позволяет, обозначая их за новые факторы, получить матрицу плана, содержащего до 7 факторов. При четырех факторах это будет полуреплика плана 2^4 ; при 5 факторах – четверть-реплика плана 2^5 ; при 6 – одна восьмая реплика плана 2^6 ; при 7 – одна шестнадцатая плана 2^7 . И все эти дробные реплики позволят построить целевую функцию в её линейном виде. При этом примеси, связанные со взаимодействием факторов, игнорируются.

§6. Крутое восхождение к оптимуму

Будем считать, что мы построили адекватную линейную модель исследуемого объекта или процесса (уравнение регрессии) в виде полинома первой степени

$$z = b_0 + \sum_{i=1}^n b_i x_i = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n, \quad (2.29)$$

все коэффициенты b_i которого значимы (а незначимые – отсеяны вместе с соответствующими им незначимыми факторами). Теперь это уравнение можно использовать для поиска тех значений факторов x_i , при которых параметр оптимизации z достигает своего оптимального значения. Если требуется достигнуть максимума этого параметра, то следует увеличивать значения тех факторов x_i , коэффициенты b_i при которых положительны, и уменьшать те из них, коэффициенты при которых отрицательны. И делать это выгоднее всего по направлению градиента функции z .

Как известно, градиент любой функции $z = z(x_1; x_2; \dots x_n)$ - это вектор ∇z n -мерного пространства $(x_1; x_2; \dots x_n)$, имеющий следующие координаты:

$$\nabla z = \left\{ \frac{\partial z}{\partial x_1}; \frac{\partial z}{\partial x_2}; \dots \frac{\partial z}{\partial x_n} \right\} \quad (2.30)$$

В направлении этого вектора функция z растет наиболее быстро (а в противоположном направлении – убывает наиболее быстро). Для целевой функции (2.29) её градиент – это вектор

$$\nabla z = \{b_1; b_2; \dots b_n\} \quad (2.31)$$

И если наша задача – обеспечить максимум целевой функции z , то двигаться факторном пространстве из центра эксперимента надо в направлении градиента этой функции. А для этого нужно изменять координаты x_i факторного пространства пропорционально коэффициентам b_i уравнения регрессии (2.29).

Впрочем, это оправдано, пока мы находимся в исследованной области $-1 \leq x_i \leq 1$ ($i = 1, 2, \dots n$) факторного пространства. Если, двигаясь по градиенту, мы выйдем за пределы этой области, то дальнейшее движение по нему уже не будет гарантировать нам увеличение параметра оптимизации z , ибо после выхода за пределы исследованной области направление градиента функции z может измениться. Поэтому, выйдя по градиенту за границы исследованной области, нужно сделать в направлении градиента два-три контрольных опыта, а затем строить и осуществлять новый план эксперимента в области факторного пространства, примыкающей к предыдущей. А в качестве центра нового экспериментального плана взять точку факторного пространства, в которой

экспериментальное значение z является наилучшим. А из неё по новому градиенту двигаться дальше – до тех пор, пока не попадем в ту область, которая содержит искомую точку экстремума целевой функции z .

Кстати, повторные опыты в новых экспериментальных планах можно не проводить, так как они нужны лишь для определения дисперсии воспроизводимости опытов $s_{\{z\}}^2$, которая уже была определена в первой серии опытов.

Этот метод называется крутым восхождением к экстремуму. Такое название метода связано с тем, что именно так, в частности, следует двигаться на горном склоне, чтобы кратчайшим путем достигнуть вершины горы.

Долог или короток будет этот путь – это зависит от многих обстоятельств: от того, как близко находится экстремальная точка факторного пространства (вершина горы) от центра первоначального плана эксперимента; как быстро меняется параметр оптимизации при движении по факторному пространству; удастся ли строить адекватные опытным данным уравнения регрессии, и т.д. И тут многое определяется опытом и навыками экспериментатора.

Кстати, достигнув крутым восхождением области факторного пространства, содержащую искомую точку экстремума параметра оптимизации, можно построить так называемый экспериментальный план второго порядка для более детального её исследования. Но на этом мы не останавливаемся, отсылая читателя, например, к книге [7].

Числовой пример

Для иллюстрации теоретического материала, изложенного в этой главе, рассмотрим числовой пример.

Пусть требуется найти наилучшие значения некоторых четырех физических факторов $\{\tilde{x}_1, \tilde{x}_2, \tilde{x}_3, \tilde{x}_4\}$, обеспечивающие максимум некоторого параметра оптимизации z . Области определения этих факторов в соответствующих им единицах измерения представлены в таблице 6.1:

Таблица 6.1

Фактор	Область определения	Область интереса
\tilde{x}_1	20 - 150	20 - 60
\tilde{x}_2	3 - 8	3 - 7
\tilde{x}_3	20 - 70	20 - 50
\tilde{x}_4	1 - 4	1 - 2

Столбец «область интереса» получен из анализа предыдущего исследования объекта.

Для эксперимента выбраны уровни варьирования факторов, представленные в таблице 6.2:

Таблица 6.2

Уровни	Факторы			
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4
Основной	30	4	30	1,5
Интервал варьирования	5	1	10	0,5
Верхний	35	5	40	2
Нижний	25	3	20	1

Интервал варьирования по первому фактору составляет примерно 4% от области его определения; по второму и третьему – 20%; по четвертому – 17%. То есть для первого фактора выбран узкий интервал варьирования, а по остальным трем средний.

Принято решение строить линейную модель исследуемого объекта, то есть модель вида

$$z = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4, \quad (2.32)$$

где

$$x_1 = \frac{\tilde{x}_1 - 30}{5}; \quad x_2 = \frac{\tilde{x}_2 - 4}{1}; \quad x_3 = \frac{\tilde{x}_3 - 30}{10}; \quad x_4 = \frac{\tilde{x}_4 - 1,5}{0,5} \quad (2.33)$$

- кодированные значения факторов. В соответствии с таким решением нет смысла строить полный ортогональный план типа 2^4 , содержащий 16 опытов (если не считать повторных). Можем ограничиться его полурепликой типа 2^{4-1} , взяв за основу полный план 2^3 , содержащий лишь 8 опытов и представленный в таблице 2.2 (в кратком виде) и в таблице 3.1 (в развернутом виде). При этом в качестве столбца для дополнительного (четвертого) фактора примем в плане, представленном в таблице 3.1, столбец $x_1x_2x_3$ тройного взаимодействия первых трех факторов. И проигнорируем столбцы парных взаимодействий факторов.

Пронумеруем планируемые опыты (таблица 6.3) с указанием условий их проведения для кодированных значений факторов:

Таблица 6.3

Номер опыта	x_1	x_2	x_3	x_4	Номер опыта	x_1	x_2	x_3	x_4
1	+	+	+	+	5	+	+	-	-
2	-	+	+	-	6	-	+	-	+
3	+	-	+	-	7	+	-	-	+
4	-	-	+	+	8	-	-	-	-

В натуральных переменных этот план представлен таблицей (6.4):

Таблица 6.4

Номер опыта	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4	Номер опыта	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4
1	35	5	40	2	5	35	5	20	1
2	25	5	40	1	6	25	5	20	2
3	35	3	40	1	7	35	3	20	2
4	25	3	40	2	8	25	3	20	1

А в таблице 6.5 представлены результаты, полученные при проведении этих опытов при двукратной повторности каждого из них.

Таблица 6.5.

Первая серия опытов

№ опыта	Рандомизация		Факторы				Параметр оптимизации z		
	1-ый опыт	2-ой опыт	x_1	x_2	x_3	x_4	z_{i1}	z_{i2}	\bar{z}_i
1	2	9	+1	+1	+1	+1	15,41	12,99	14,20
2	6	13	-1	+1	+1	-1	5,42	4,00	4,71
3	1	15	+1	-1	+1	-1	3,58	2,78	3,18
4	7	10	-1	-1	+1	+1	6,10	4,50	5,30
5	3	16	+1	+1	-1	-1	5,35	4,15	4,75
6	8	14	-1	+1	-1	+1	4,76	3,12	3,94
7	4	12	+1	-1	-1	+1	4,10	2,70	3,40
8	5	11	-1	-1	-1	-1	2,40	1,46	1,93

Для каждой строки i матрицы планирования (6.5), то есть для каждого опыта, имеющего $p = 2$ повторностей, по формуле (2.12) находим дисперсию s_i^2 параметра оптимизации z , характеризующего степень разброса значений z_{ik} ($k = 1, 2, \dots, p$) этого параметра в i -ом опыте ($i = 1, 2, \dots, N = 8$). То есть вычисляем величины

$$s_i^2 = (z_{i1} - \bar{z}_i)^2 + (z_{i2} - \bar{z}_i)^2 \quad (i = 1, 2, \dots, 8) \quad (2.34)$$

Их числовые значения приведены в таблице (6.6):

Таблица 6.6

i	1	2	3	4	5	6	7	8
s_i^2	2,928	1,008	0,320	1,280	0,720	1,345	0,980	0,442

Исследуем эти дисперсии на однородность (на незначимость их различия) по критерию Кочрена. Для этого сначала по формуле (2.15) найдем экспериментальное значение этого критерия:

$$G = G_{\text{эксн}} = \frac{\{s_i^2\}_{\max}}{\sum_{i=1}^N s_i^2} = \frac{2,928}{9,023} = 0,325$$

Затем по таблице (4.1) критических значений критерия Кочрена при уровне значимости $\alpha = 0,05$ и степенях свободы $f_1 = p - 1 = 1$ и $f_2 = N = 8$ находим критическое значение этого критерия: $G_{\text{кр}} = 0,680$. И так как $G_{\text{эксн}} < G_{\text{кр}}$, то у нас нет оснований считать дисперсии s_i^2 различающимися значимо. Они различаются незначимо. А, значит, мы можем продолжить статистический анализ результатов проведенных опытов.

Найдем среднее значение дисперсий s_i^2 , то есть дисперсию воспроизводимости опытов $s_{\{z\}}^2$ во всем эксперименте:

$$s_{\{z\}}^2 = \frac{\sum_{i=1}^N s_i^2}{N} = \frac{9,023}{8} = 1,13 \quad (2.35)$$

Она нам понадобится чуть позже – при оценке значимости уравнения регрессии (2.32), которое мы сейчас построим. Используя формулы (2.8), получаем коэффициенты уравнения (2.32):

$$b_0 = 5,176; \quad b_1 = 1,206; \quad b_2 = 1,724; \quad b_3 = 1,671; \quad b_4 = 1,534 \quad (2.36)$$

Таким образом, получено следующее аналитическое выражение для параметра оптимизации z :

$$z = 5,176 + 1,206x_1 + 1,724x_2 + 1,671x_3 + 1,534x_4 \quad (2.37)$$

Подсчитаем по этому уравнению значения z_i параметра оптимизации для каждого из 8 проведенных опытов, и сравним эти значения со средними опытными его значениями \bar{z}_i , содержащимися в таблице (6.5):

Таблица 6.7

i	1	2	3	4	5	6	7	8
z_i	11,31	5,83	4,80	5,45	4,90	5,56	4,52	-0,96
\bar{z}_i	14,20	4,71	3,18	5,30	4,75	3,94	3,40	1,93
$\Delta z_i = z_i - \bar{z}_i$	-2,89	1,12	1,62	0,15	0,15	1,62	1,12	-2,89

Контроль: суммы чисел z_i и \bar{z}_i должны совпадать. Или, что одно и то же, сумма всех невязок Δz_i должна быть равна нулю. У нас это (проверьте) выполняется.

Как видим из таблицы 6.7, опытные и расчетные значения параметра оптимизации не совпадают. Выясним, значимо или незначимо их различие. То есть адекватна или неадекватна опытными данным построенная линейная модель (2.37) исследуемого объекта. Для этого, используя формулу (2.18)

$$s_{ad}^2 = \frac{\sum_{k=1}^N (\Delta z_k)^2}{N - N_0}, \quad (2.38)$$

подсчитаем дисперсию адекватности s_{ad}^2 . Здесь $N - N_0$ - число её степеней свободы, где $N = 8$ число опытов в плане нашего эксперимента (без учета повторности), а $N_0 = 5$ - число коэффициентов b_i в построенном уравнении регрессии (2.37). Значит, $N - N_0 = 3$, а

$$s_{ad}^2 = \frac{\sum_{i=1}^8 (\Delta z_i)^2}{3} = 8.17$$

Теперь сравним дисперсию адекватности s_{ad}^2 с дисперсией воспроизводимости $s_{\{z\}}^2$ по критерию Фишера. Используя формулу (2.20), имеем следующее экспериментальное значение критерия Фишера:

$$F_{эксн} = \frac{s_{ad}^2}{s_{\{z\}}^2} = \frac{8,71}{1,13} = 7,71$$

А его критическое значение $F_{кр} = F_{кр}(\alpha; k_1; k_2)$ при уровне значимости $\alpha = 0,05$, числе степеней свободы $k_1 = 3$ числителя и $k_2 = 8$ знаменателя, согласно таблице 4.2, равно: $F_{кр} = F_{кр}(0,05; 3; 8) = 4,07$. И так как $F_{эксн} > F_{кр}$, то построенное уравнение регрессии (2.37) неадекватно опытными данным. Следовательно, оно не годится для теоретического исследования параметра оптимизации z .

Ну что ж, нам не повезло. Нужно будет строить новую, но уже адекватную, модель исследуемого объекта. А для этого придется строить новую серию опытов – с новым центром плана эксперимента и с новыми (уменьшенными) интервалами варьирования факторов. В качестве нового центра эксперимента выгодно взять параметры наилучшего опыта из проведенных экспериментов. Согласно таблице (6.5), наилучшее (максимальное) значение параметра оптимизации z имело место в первом опыте плана, когда все факторы устанавлива-

лись на своих верхних уровнях. То есть когда их натуральные значения были такими: $\tilde{x}_1 = 35$; $\tilde{x}_2 = 5$; $\tilde{x}_3 = 40$; $\tilde{x}_4 = 2$ (см. таблицу 6.4). Эти значения факторов и примем за центр нового плана эксперимента. И, по сравнению с предыдущим планом, сузим вдвое интервалы варьирования по каждому фактору. То есть условия нового плана эксперимента теперь будут определяться не таблицей (6.2), а таблицей (6.8):

Таблица 6.8

Уровни	Факторы			
	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4
Основной	35	5	40	2
Интервал варьирования	2,5	0,5	5	0,25
Верхний	37,5	5,5	45	2,25
Нижний	32,5	4,5	35	1,75

При этом вместо результатов экспериментов, представленных в таблице 6.5, получены новые результаты, представленные в таблице 6.9:

Таблица 6.9

Вторая серия опытов

№ опыта	Рандомизация		Факторы				Параметр оптимизации z		
	1-ый опыт	2-ой опыт	x_1	x_2	x_3	x_4	z_{i1}	z_{i2}	\bar{z}_i
1	6	16	+1	+1	+1	+1	12,85	14,15	13,50
2	5	7	-1	+1	+1	-1	13,20	11,80	12,50
3	8	9	+1	-1	+1	-1	8,70	7,70	8,20
4	3	14	-1	-1	+1	+1	9,84	7,16	8,50
5	2	12	+1	+1	-1	-1	12,10	13,30	12,70
6	10	15	-1	+1	-1	+1	15,21	12,79	14,00
7	1	13	+1	-1	-1	+1	9,25	10,75	10,00
8	4	11	-1	-1	-1	-1	12,15	10,85	11,50

Вместо таблицы 6.6 получаем новую таблицу 6.10 дисперсий s_i^2 параметра оптимизации в отдельных опытах:

Таблица 6.10

i	1	2	3	4	5	6	7	8
s_i^2	0,845	0,980	0,500	3,591	0,720	2,928	1,125	0,845

Проверка по критерию Кочрена опять подтверждает незначимость различия дисперсий s_i^2 . А, значит, можем продолжить статистический анализ результатов проведенных опытов.

Находим дисперсию воспроизводимости опытов $s_{\{z\}}^2$ во всём эксперименте:

$$s_{\{z\}}^2 = \frac{\sum_{i=1}^N s_i^2}{N} = \frac{11,534}{8} = 1,44$$

Новая серия опытов приводит к другому, чем (2.37), уравнению регрессии:

$$z = 11,363 - 0,263x_1 + 1,813x_2 - 0,688x_3 + 0,138x_4 \quad (2.39)$$

В соответствии с новым уравнением регрессии изменится и таблица 6.7. Её новый вид – таблица 6.11:

Таблица 6.11

i	1	2	3	4	5	6	7	8
z_i	12,363	12,613	8,461	9,263	13,463	14,265	10,113	10,363
\bar{z}_i	13,50	12,50	8,20	8,50	12,70	14,00	10,00	11,50
$\Delta z_i = z_i - \bar{z}_i$	-1,137	0,113	0,261	0,763	0,763	0,265	0,113	-1,137

Согласно этой таблице

$$s_{ad}^2 = \frac{\sum_{i=1}^8 (\Delta z_i)^2}{3} = \frac{3,914}{3} = 1,30; \quad F_{эксн} = \frac{s_{ad}^2}{s_{\{z\}}^2} = \frac{1,30}{1,44} = 0,90; \quad (2.40)$$

$$F_{кр} = F_{кр}(0,05; 3; 8) = 4,07$$

И так как $F_{эксн} < F_{кр}$, то уравнение (2.39), в отличие от уравнения (2.37), адекватно опытным данным.

Выясним теперь, все ли коэффициенты этого уравнения значимы. Для этого сначала по формуле (2.19) подсчитаем дисперсии этих коэффициентов:

$$s_{\{b_i\}}^2 = \frac{s_{\{z\}}^2}{N} = \frac{1,44}{8} = 0,18 \quad (0 \leq i \leq 4) \quad (2.41)$$

Далее, по формуле (2.21)

$$t_i = \frac{|b_i|}{s_{\{b_i\}}} = \sqrt{\frac{N}{s_{\{z\}}^2}} \cdot |b_i| \quad (0 \leq i \leq 4) \quad (2.42)$$

подсчитаем экспериментальные значения случайных величин t_i , имеющих распределение Стьюдента с $k = N(p-1) = 8(2-1) = 8$ степенями свободы:

$$t_i = \sqrt{\frac{8}{1,44}} \cdot |b_i| = 2,36 \cdot b_i \quad (0 \leq i \leq 4)$$

$$t_0 = 26,82; \quad t_1 = 0,62; \quad t_2 = 4,28; \quad t_3 = 1,62; \quad t_4 = 0,33$$

По таблице 4.3 критических точек распределения Стьюдента находим критическое значение $t_{кр}(\alpha; k)$ этого распределения: $t_{кр} = t_{кр}(0,05; 8) = 2,31$. Так как t_1, t_3 и t_4 оказались меньше $t_{кр}$, то коэффициенты b_1, b_3 и b_4 в уравнении (2.39) являются незначимыми, и их без особого ущерба для качества этого уравнения можно заменить нулями. В итоге это уравнение сильно упростится:

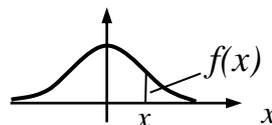
$$z = 11,363 + 1,813x_2$$

Полученное уравнение свидетельствует о том, что факторы x_1, x_3 и x_4 , отсутствующие в нем, являются для параметра оптимизации в области эксперимента незначимыми. Их можно зафиксировать на их средних значениях $\{x_1 = 0; x_3 = 0; x_4 = 0\}$. То есть, согласно таблице 6.8, на их натуральных значениях $\{\tilde{x}_1 = 35; \tilde{x}_3 = 40; \tilde{x}_4 = 2\}$. И достигать максимума параметра оптимизации z следует, понемногу увеличивая фактор \tilde{x}_2 и отслеживая получающийся эффект с помощью контрольных опытов.

(таблицы)

Таблица 1.

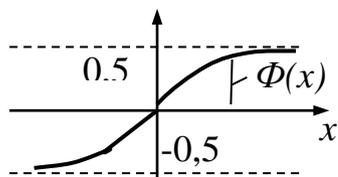
Значения функции $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$



x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0026
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002-	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблица 2.

Значения функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$

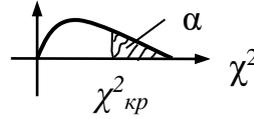


x	Φ(x)										
0,00	0,0000	0,39	0,1517	0,78	0,2823	1,17	0,3790	1,56	0,4406	1,95	0,4744
0,01	0,0040	0,40	0,1554	0,79	0,2852	1,18	0,3810	1,57	0,4418	1,96	0,4750
0,02	0,0080	0,41	0,1591	0,80	0,2881	1,19	0,3830	1,58	0,4429	1,97	0,4756
0,03	0,0120	0,42	0,1628	0,81	0,2910	1,20	0,3849	1,59	0,4441	1,98	0,4761
0,04	0,0160	0,43	0,1664	0,82	0,2939	1,21	0,3869	1,60	0,4452	1,99	0,4767
0,05	0,0199	0,44	0,1700	0,83	0,2967	1,22	0,3883	1,61	0,4463	2,00	0,4772
0,06	0,0239	0,45	0,1736	0,84	0,2995	1,23	0,3907	1,62	0,4474	2,02	0,4783
0,07	0,0279	0,46	0,1772	0,85	0,3023	1,24	0,3925	1,63	0,4484	2,04	0,4793
0,08	0,0319	0,47	0,1808	0,86	0,3051	1,25	0,3944	1,64	0,4495	2,06	0,4803
0,09	0,0359	0,48	0,1844	0,87	0,3078	1,26	0,3962	1,65	0,4505	2,08	0,4812
0,10	0,0398	0,49	0,1879	0,88	0,3106	1,27	0,3980	1,66	0,4515	2,10	0,4821
0,11	0,0438	0,50	0,1915	0,89	0,3133	1,28	0,3997	1,67	0,4525	2,12	0,4830
0,12	0,0478	0,51	0,1950	0,90	0,3159	1,29	0,4015	1,68	0,4535	2,14	0,4838
0,13	0,0517	0,52	0,1985	0,91	0,3186	1,30	0,4032	1,69	0,4545	2,16	0,4846
0,14	0,0557	0,53	0,2019	0,92	0,3212	1,31	0,4049	1,70	0,4554	2,18	0,4854
0,15	0,0596	0,54	0,2054	0,93	0,3238	1,32	0,4066	1,71	0,4564	2,20	0,4861
0,16	0,0636	0,55	0,2088	0,94	0,3264	1,33	0,4082	1,72	0,4573	2,22	0,4868
0,17	0,0675	0,56	0,2123	0,95	0,3289	1,34	0,4099	1,73	0,4582	2,24	0,4875
0,18	0,0714	0,57	0,2157	0,96	0,3315	1,35	0,4115	1,74	0,4591	2,26	0,4881
0,19	0,0753	0,58	0,2190	0,97	0,3340	1,36	0,4131	1,75	0,4599	2,28	0,4887
0,20	0,0793	0,59	0,2224	0,98	0,3365	1,37	0,4147	1,76	0,4608	2,30	0,4893
0,21	0,0832	0,60	0,2257	0,99	0,3389	1,38	0,4162	1,77	0,4616	2,32	0,4898
0,22	0,0871	0,61	0,2291	1,00	0,3413	1,39	0,4177	1,78	0,4625	2,34	0,4904
0,23	0,0910	0,62	0,2324	1,01	0,3438	1,40	0,4192	1,79	0,4633	2,36	0,4909
0,24	0,0948	0,63	0,2357	1,02	0,3461	1,41	0,4207	1,80	0,4641	2,38	0,4913
0,25	0,0987	0,64	0,2389	1,03	0,3485	1,42	0,4222	1,81	0,4649	2,40	0,4918
0,26	0,1026	0,65	0,2422	1,04	0,3508	1,43	0,4236	1,82	0,4656	2,42	0,4922
0,27	0,1064	0,66	0,2454	1,05	0,3531	1,44	0,4251	1,83	0,4664	2,44	0,4927
0,28	0,1103	0,67	0,2486	1,06	0,3554	1,45	0,4265	1,84	0,4671	2,46	0,4931
0,29	0,1141	0,68	0,2517	1,07	0,3577	1,46	0,4279	1,85	0,4678	2,48	0,4934
0,30	0,1179	0,69	0,2549	1,08	0,3599	1,47	0,4292	1,86	0,4686	2,50	0,4938
0,31	0,1217	0,70	0,2580	1,09	0,3621	1,48	0,4306	1,87	0,4693	2,60	0,4953
0,32	0,1255	0,71	0,2611	1,10	0,3643	1,49	0,4319	1,88	0,4699	2,70	0,4965
0,33	0,1293	0,72	0,2642	1,11	0,3665	1,50	0,4332	1,89	0,4706	2,80	0,4974
0,34	0,1331	0,73	0,2673	1,12	0,3686	1,51	0,4345	1,90	0,4713	2,90	0,4981
0,35	0,1368	0,74	0,2703	1,13	0,3708	1,52	0,4357	1,91	0,4719	3,00	0,4986
0,36	0,1406	0,75	0,2734	1,14	0,3729	1,53	0,4370	1,92	0,4726	3,50	0,4997
0,37	0,1443	0,76	0,2764	1,15	0,3749	1,54	0,4382	1,93	0,4732	4,00	0,4999
0,38	0,1480	0,77	0,2794	1,16	0,3770	1,55	0,4394	1,94	0,4738	5,00	0,5

Таблица 3.

Критические точки $\chi^2_{кр} = \chi^2_{кр}(\alpha; k)$ распределения «Хи-квадрат» с k степенями свободы, удовлетворяющие условию

$$P(\chi^2 > \chi^2_{кр}) = \alpha$$

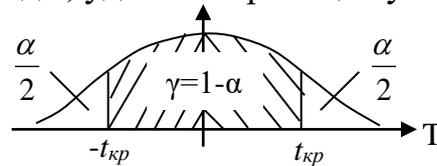


k	α			k	α		
	0,10	0,05	0,01		0,10	0,05	0,01
1	2,71	3,84	6,63	16	23,54	26,30	32,00
2	4,61	5,99	9,21	17	24,77	27,59	33,41
3	6,25	7,81	11,34	18	25,99	28,87	34,81
4	7,78	9,95	13,28	19	27,20	30,14	36,19
5	9,24	11,07	15,09	20	28,41	31,41	37,57
6	10,64	12,59	16,81	21	29,62	32,67	38,93
7	12,02	14,07	18,48	22	30,81	33,92	40,29
8	13,36	15,51	20,09	23	32,01	35,17	41,64
9	14,68	16,92	21,67	24	33,20	36,42	42,98
10	15,99	18,31	23,21	25	34,38	37,65	44,31
11	17,28	19,68	24,72	26	35,56	38,89	45,64
12	18,55	21,03	26,22	27	36,74	40,11	46,96
13	19,81	22,36	27,69	28	37,92	41,34	48,28
14	21,06	23,68	29,14	29	39,09	42,56	49,59
15	22,31	25,00	30,58	30	40,26	43,77	50,89

Таблица 4.

Критические точки $t_{кр} = t_{кр}(\alpha; k)$ распределения Стьюдента (Т - распределения) с k степенями свободы, удовлетворяющие условию

$$P(-t_{кр} < T < t_{кр}) = \gamma = 1 - \alpha$$

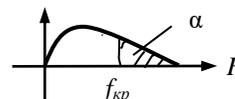


k	$\gamma (\alpha)$			k	$\gamma (\alpha)$		
	0,90 (0,10)	0,95 (0,05)	0,99 (0,01)		0,90 (0,10)	0,95 (0,05)	0,99 (0,01)
1	6,31	12,71	63,66	9	1,83	2,26	3,25
2	2,92	4,30	9,92	10	1,81	2,23	3,17
3	2,35	3,18	5,84	15	1,75	2,13	2,95
4	2,13	2,78	4,60	20	1,72	2,09	2,85
5	2,02	2,57	4,03	30	1,70	2,04	2,75
6	1,94	2,45	3,71	60	1,67	2,00	2,66
7	1,89	2,36	3,50	120	1,66	1,98	2,62
8	1,86	2,31	3,36	∞	1,64	1,96	2,58

Таблица 5.

Критические точки $f_{кр} = f_{кр}(\alpha; k_1; k_2)$ распределения Фишера–Снедекора (F – распределения) с $(k_1; k_2)$ степенями свободы, где k_1 – число степеней большей дисперсии, а k_2 – число степеней меньшей дисперсии, удовлетворяющие условию

$$p(F > f_{кр}) = \alpha$$



$k_1 \backslash k_2$	$\alpha = 0,05$																		
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	240	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
17	4,45	3,59	3,20	2,9	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	4,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2*155	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	0,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	3,60	2,37	2,21	2,10	2,01	1,94	1,83	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00
k_1	$\alpha = 0,01$																		
k_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Литература

1. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. М.: Наука, 1976.
2. Воскобойников Ю.Е. Эконометрика в Excel. Ч. 1. Парный и множественный регрессионный анализ: учебное пособие. Новосибирск: Изд-во Новосибирский государственный архитектурно-строительный университет, 2005.
3. Горелова Г.В. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel. Ростов н/Д: Феникс, 2006.
4. Математическое моделирование в компонентах природы (интерактивный курс) / Г.Н. Камышова, В.В. Корсак, А.С. Фалькович, О.Ю. Холуденева. Саратов: Изд-во Саратовского ГАУ, 2012.
5. Комогорцев В.Ф. Лекции по теории вероятностей и математической статистике. Брянск: Изд-во Брянская ГСХА, 2009.
6. Кремер Н.Ш. Теория вероятностей и математическая статистика. М., «Юнити», 2002.
7. Сидняев Н.И. Теория планирования экспериментов и анализ статистических данных. 2 изд., переработанное и дополненное: учебное пособие для магистров. М., Юрайт, 2015.
8. Введение в математическое моделирование / под ред. П.В. Трусова. М., Логос, 2005.

Учебное издание

Комогорцев Владимир Филиппович

**МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ ПРОЦЕССОВ
В КОМПОНЕНТАХ ПРИРОДЫ**

Учебное пособие

Редактор Павлютина И.П.

Подписано к печати 23.08.2018 г. Формат 60x84 ¹/₁₆.

Бумага офсетная. Усл. п. л. 7,61. Тираж 100 экз. Изд. №6200.

Издательство Брянского государственного аграрного университета
243365 Брянская обл., Выгоничский район, с. Кокино, Брянский ГАУ

