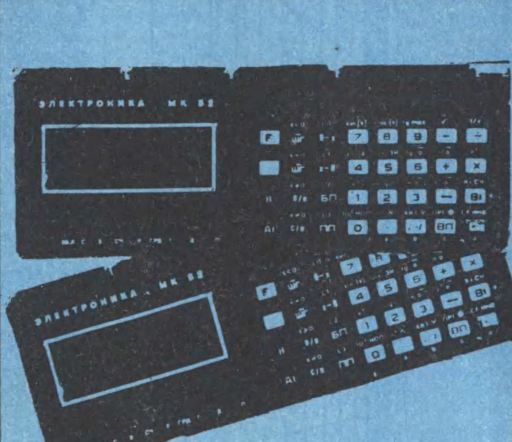
СТАТИСТИКА И ИНФОРМАТИКА

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ НА ПРОГРАММИРУЕМЫХ МИКРОКАЛЬКУЛЯТОРАХ



СТАТИСТИКА И ИНФОРМАТИКА

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ НА ПРОГРАММИРУЕМЫХ МИКРОКАЛЬКУЛЯТОРАХ

Справочное пособие

Под редакцией
профессора В.В.Шуракова

МОСКВА
"ФИНАНСЫ И СТАТИСТИКА"
1991

ББК 22.172
М34

Авторы: *А.М.Дубров, В.С.Мхитарян, Л.И.Трошин, И.В.Масленченко*

Рецензенты: *В.С.Проскуров, А.Н.Романов*

ISBN 5-279-00528-2

ПРЕДИСЛОВИЕ

Существует большой класс математико-статистических задач, для решения которых целесообразно использовать программируемые микрокалькуляторы. Они просты в работе и не требуют от пользователя специальной подготовки. Малые габариты и масса, питание от автономного источника или сети переменного тока, возможность выполнения всех операций с клавишного пульта, а также низкая стоимость делают эти микрокалькуляторы доступными для широкого круга пользователей. Программируемые микрокалькуляторы, обладая всеми возможностями обычных микрокалькуляторов, позволяют значительно сократить время вычисления в результате использования режима работы по программе. Особенно ощутимый эффект имеет место при решении многомерных задач статистического анализа.

Современные программируемые микрокалькуляторы второго поколения – "Электроника БЗ-34" и ее аналоги "Электроника МК-56" и "Электроника МК-54" – обладают функциональными возможностями, присущими ранее лишь большим ЭВМ: прямой и косвенной адресацией, организацией программ и циклов, безусловных и условных переходов, микропрограммным вычислением элементарных и отдельных специальных функций и т.д. Таким образом, эти программируемые микрокалькуляторы удовлетворяют важнейшим требованиям, предъявляемым к микроЭВМ.

Развитие программируемых микрокалькуляторов привело к созданию новых моделей типа "Электроника МК-52" и "Электроника МК-61", имеющих расширенный набор микропрограмм, встроенные постоянное и оперативное запоминающие устройства. Эти модели могут подключаться к различным периферийным устройствам. Последние модели программируемых микрокалькуляторов американской фирмы Hewlett Packard имеют ёмкость программной памяти на 448 шагов, решают системы из 2–7 линейных уравнений, сохраняют данные в регистрах памяти и после отключения питания. Такую же энергозащищенную память на 500 шагов имеет отечественный микрокалькулятор "Электроника МК-52". При работе с "Электроникой МК-52" в энергозащищенную память можно внести всю систему программ соответствующего раздела справочника и в дальнейшем решать задачи этого раздела, вызывая из энергозащищенной памяти в оперативную нужную программу. Использование блока расширения памяти (БРП-3) при моделировании рядов динамики позволяет существенно упростить процедуру вычисления трендов различных видов.

Программируемые микрокалькуляторы имеют простой символьно-кодовый язык программирования. Причем все модели отечественных микрокалькуляторов полностью совместимы между собой по языку программирования и системе команд. Кроме того, программы, составленные для отечественных программируемых микрокалькуляторов, могут быть после несложного перевода использованы и для ряда зарубежных микрокалькуляторов.

Отечественные программируемые микрокалькуляторы второго поколения работают в двух режимах: автоматическом и программирования. Функциональные возможности каждого режима подробно излагаются в "Техническом описании и инструкции по эксплуатации", прилагающемся к каждому микрокалькулятору.

При решении математико-статистических задач программируемые микрокалькуляторы позволяют выполнять громоздкие и трудоемкие вычисления в автоматическом режиме по программе. Однако не следует думать, что решение задач сводится, таким образом, к бездумному нажиманию клавиш. Напротив, акцент переносится с преодоления вычислительных трудностей на интерпретацию полученных результатов и более глубокое понимание изучаемого процесса. Применение программируемых микрокалькуляторов повышает эффективность изучения прикладной статистики и позволяет расширить круг рассматриваемых задач с реальными экономическими данными.

Справочник рассчитан в первую очередь на читателей, знакомящихся с современными методами многомерного статистического анализа. Такие методы широко используются для решения многообразных социально-экономических, технико-экономических и других задач. Для более полного овладения этими методами можно рекомендовать читателю трехтомное справочное издание "Прикладная статистика" [1, 2, 3].

Справочник предполагает наличие знаний читателя по математическим дисциплинам в объеме первых курсов экономических и технических вузов и не требует специальной подготовки по программированию. Учитывая возможность использования справочника в целях обучения, в работе предусмотрены по каждой теме 25 вариантов заданий для самостоятельной работы (приложение 2).

Наряду с подробно рассмотренными контрольными примерами, самостоятельное решение варианта задания будет способствовать углубленному пониманию как математической, так и содержательной сторон соответствующего статистического метода.

Перечень операций, используемых в программах для микрокалькуляторов

№ п/п	Клавиши	Код операции	Содержание операции
1	2	3	4
1	0 ,..., 9	00,..., 09	Занесение цифр от 0 до 9 в регистр x
2	x → П 0 ,..., x → П 9	40,..., 49	Запись содержимого регистра x в регистр 0,..., 9
3	x → П a	4 –	Запись содержимого регистра x в регистр a
4	x → П b	4*L*	Запись содержимого регистра x в регистр b
5	x → П c	4*C*	Запись содержимого регистра x в регистр c
6	x → П d	4*Г*	Запись содержимого регистра x в регистр d
7	П → x 0,..., П → x 9	60,..., 69	Вызов содержимого регистра 0,..., 9 в регистр x
8	П → x a	6 –	Вызов содержимого регистра a в регистр x
9	П → x b	6*L*	Вызов содержимого регистра b в регистр x
10	П → x c	6*C*	Вызов содержимого регистра c в регистр x
11	П → x d	6*Г*	Вызов содержимого регистра d в регистр x
12	+	10	Сложение операндов
13	–	11	Вычитание операндов
14	×	12	Умножение операндов
15	:	13	Деление операндов
16	·	0 –	Занесение десятичной запятой
17	/–/	0*L*	Изменение знака числа или порядка
18	ВП	0*C*	Ввод порядка числа
19	F Вx	0	Восстановление предыдущего результата
20	Cx	0*Г*	Сброс содержимого регистра x
21	В↑	0*E*	Передвижение информации в стеке
22	↔	14	Обмен операндами в регистрах x и y
23	F 10^x	15	Вычисление функции 10^x
24	F e^x	16	Вычисление e^x
25	F lg	17	Вычисление десятичного логарифма $\lg x$
26	F ln	18	Вычисление натурального логарифма $\ln x$
27	F sin	1*C*	Вычисление функции синуса $\sin x$

1	2	3	4
28	F cos	1*Г*	Вычисление косинуса $\cos x$
29	F tg	1E	Вычисление тангенса $\mathrm{tg}\,x$
30	F $\sin^{-1}$	19	Вычисление арксинуса $\arcsin x$
31	F $\cos^{-1}$	1 –	Вычисление $\arccos x$
32	F tg^{-1}	1*L*	Вычисление $\mathrm{arctg}\,x$
33	F π	20	Занесение в регистр x константы $\pi = 3{,}1415926$
34	F $\sqrt{}$	21	Вычисление корня квадратного $\sqrt{x}$
35	F x^2	22	Возведение x в квадрат
36	F $1/x$	23	Вычисление обратной величины $1/x$
37	F x^y	24	Возведение числа x в степень y
38	F ○	25	Кольцевые передвижения информации в стеке
39	С/П	50	1. Прекращение прохождения программы в режиме программирования и фиксация содержимого регистра X на индикаторе 2. Вычисления по программе в режиме автоматической работы и прекращение вычислений при зацикливании
40	БП	51	Безусловный переход
41	В/О	52	1. Возврат из программы в режиме программирования 2. Переход на нулевой адрес в режиме автоматической работы
42	ПП	53	1. Переход на программу в режиме программирования 2. Потактовое прохождение программы в режиме автоматической работы
43	К НОП	54	Исключение команды
44	F $x \neq 0$	57	Прямой переход по условию $x \neq 0$
45	F $x \geqslant 0$	59	Прямой переход по условию $x \geqslant 0$
46	F $x = 0$	5*E*	Прямой переход по условию $x = 0$
47	F $x < 0$	5*C*	Прямой переход по условию $x < 0$
48	F L0	5*Г*	Организация циклов с регистром 0
49	F L1	5*L*	Организация циклов с регистром 1

Продолжение

1	2	3	4
50	F L2	58	Организация циклов с регистром 2
51	F L3	5 –	Организация циклов с регистром 3
52	$\overrightarrow{\text{ШГ}}$	–	Потактовое прохождение программы в режиме программирования в порядке увеличения адресов
53	$\overleftarrow{\text{ШГ}}$	–	Потактовое прохождение программы в режиме программирования в порядке уменьшения адресов
54	F ПРГ	–	Переход в режиме программирования
55	F АВТ	–	Переход в режиме автоматической работы

1. РЕГРЕССИОННЫЙ И КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

1.1. АЛГОРИТМ РЕГРЕССИОННОГО И КОРРЕЛЯЦИОННОГО АНАЛИЗА

1.1.1. Корреляционный анализ

Корреляционный анализ является одним из методов статистического анализа взаимозависимости нескольких признаков. Он применяется тогда, когда данные наблюдений можно считать случайными и выбранными из генеральной совокупности, распределенной по многомерному нормальному закону. Основная задача корреляционного анализа состоит в оценке корреляционной матрицы генеральной совокупности по выборке и определении на ее основе оценок частных и множественных коэффициентов корреляции и детерминации. Дополнительная задача корреляционного анализа (являющаяся основной в регрессионном анализе) – оценка уравнения регрессии.

Парный и *частный коэффициенты корреляции* характеризуют *тесноту* линейной зависимости между двумя переменными соответственно на фоне действия и при исключении влияния всех остальных показателей, входящих в модель. Они изменяются в пределах от – 1 до +1, причем чем ближе коэффициент корреляции к ± 1, тем сильнее зависимость между переменными. Если коэффициент корреляции больше 0, то *связь положительная,* а если меньше нуля – *отрицательная.*

Множественный коэффициент корреляции характеризует тесноту связи между одной переменной (результативной) и остальными, входящими в модель; изменяется в пределах от 0 до 1. Квадрат множественного коэффициента корреляции называется *множественным коэффициентом детерминации.* Он характеризует долю дисперсии одной переменной (результативной), обусловленной влиянием остальных (аргументов), входящих в модель.

Исходной для анализа является матрица

$$\mathbf{X} = \begin{pmatrix} x_{11} \ldots x_{1j} \ldots x_{1k} \\ \cdots\cdots\cdots\cdots \\ x_{i1} \ldots x_{ij} \ldots x_{ik} \\ \cdots\cdots\cdots\cdots \\ x_{n1} \ldots x_{nj} \ldots x_{nk} \end{pmatrix}$$

размерности $(n \times k)$, которая представляет собой n наблюдений для каждого из k факторов.

Оцениваются параметры k-мерной генеральной совокупности, подчиняющейся нормальному закону распределения, а именно: вектор средних ($\bar{x}$), вектор среднеквадратических отклонений (s) и корреляционная матрица (**R**) порядка k:

$$\bar{\mathbf{x}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_k \end{pmatrix}, \quad \mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \end{pmatrix},$$

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \dots & r_{1k} \\ r_{21} & 1 & \dots & r_{2k} \\ \cdot & \cdot & \dots & \cdot \\ r_{k1} & r_{k2} & \dots & 1 \end{pmatrix}$$

Матрица **R** является симметрической и положительно определенной, где

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}, \quad s_j = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2},$$

$$r_{jl} = \frac{\frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{il} - \bar{x}_l)}{s_j s_l}; \quad j, l = 1, 2, \dots, k,$$

x_{ij} – значение i-го наблюдения j-го фактора.

Кроме того, находятся точечные оценки частных и множественных коэффициентов корреляции любого порядка. Например, частный коэффициент корреляции (k – 2)-го порядка между факторами X_1 и X_2 равен:

$$r_{12/3,4,\dots,k} = -\frac{R_{12}}{\sqrt{R_{11} R_{22}}},$$

где R_{jl} – алгебраическое дополнение элемента r_{jl} корреляционной матрицы **R**.

Множественный коэффициент корреляции (k – 1)-го порядка фактора (результативного признака) X_1 определяется по формуле:

$$r_{1/2,3,\dots,k} = r_1 = \sqrt{1 - \frac{|\mathbf{R}|}{R_{11}}},$$

где $|\mathbf{R}|$ – определитель матрицы **R**.

Значимость частных и парных коэффициентов корреляции проверяется по t-критерию Стьюдента. Наблюдаемое значение критерия находится по формуле:

$$t_{\text{набл}} = \frac{r}{\sqrt{1-r^2}} \sqrt{n-l-2},$$

где r – соответственно оценка частного или парного коэффициента корреляции; l – порядок коэффициента корреляции, т.е. число фиксируемых факторов.

Напомним, что проверяемый коэффициент корреляции считается *значимым*, т.е. гипотеза H_0: $\rho = 0$ отвергается с вероятностью ошибки α, если $t_{\text{набл}}$ по модулю будет больше, чем $t_{\text{кр}}$, определяемое по таблицам t-распределения (см. приложение) для заданного α и $\nu = n - l - 2$.

Значимость множественного коэффициента корреляции (или его квадрата – коэффициента детерминации) проверяется по F-критерию.

Наблюдаемое значение, например для $\rho_{1/2,\,3,\ldots,\,k}$, находится по формуле:

$$F_{\text{набл}} = \frac{\frac{1}{k-1} r^2_{1/2,\ldots,k}}{\frac{1}{n-k}(1 - r^2_{1/2,\ldots,k})}.$$

Множественный коэффициент корреляции считается значимым, т.е. имеет место линейная статистическая зависимость между X_1 и остальными факторами $X_2,\ldots, X_k$, если: $F_{\text{набл}} > F_{\text{кр}}$ $(\alpha, k-1, n-k)$, где $F_{\text{кр}}$ определяется по таблице F-распределения для заданных α, $\nu_1 = k - 1$, $\nu_2 = n - k$.

1.1.2. Регрессионный анализ

Регрессионный анализ – это статистический метод исследования зависимости случайной величины Y от переменных X_j $(j = 1, 2,\ldots, k)$, рассматриваемых в регрессионном анализе как неслучайные величины независимо от истинного закона распределения X_j.

Обычно предполагается, что случайная величина Y имеет нормальный закон распределения с условным математическим ожиданием $\widetilde{Y} = \varphi$ $(X_1, X_2,\ldots, X_k)$, являющимся функцией от аргументов x_j, и с постоянной, не зависящей от аргументов дисперсией σ^2.

Наиболее часто встречаются линейные уравнения регрессии вида:

$$\widetilde{Y} = \beta_0 + \beta_1 x_1 + \ldots + \beta_j x_j + \ldots + \beta_k x_k,$$

линейные относительно неизвестных параметров β_j $(j = 0, 1, \ldots, k)$ и аргументов x_j.

Коэффициент регрессии β_j показывает, на какую величину в среднем изменится результативный признак Y, если переменную X_j увеличить на единицу его измерения, т.е. является *нормативным коэффициентом*.

В матричной форме регрессионная модель имеет вид:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

где $\mathbf{Y}$ – случайный вектор-столбец размерности $(n \times 1)$ наблюдаемых значений результативности признака $(y_1, y_2,..., y_n)$; $\mathbf{X}$ – матрица размерности $[n \times (k + 1)]$ наблюдаемых значений аргументов. Элемент матрицы x_{ij} рассматривается как неслучайная величина ($i = 1, 2,..., n$; $j = 0, 1, 2,..., k$; $x_{0i} = 1$); $\boldsymbol{\beta}$ – вектор-столбец размерности $[(k + 1) \times 1]$ неизвестных, подлежащих оценке параметров (коэффициентов регрессии) модели; $\boldsymbol{\varepsilon}$ – случайный вектор-столбец размерности $(n \times 1)$ ошибок наблюдений (остатков). Компоненты вектора ε_i независимы между собой, имеют нормальный закон распределения с нулевым математическим ожиданием ($M\varepsilon_i = 0$) и неизвестной дисперсией σ^2 ($D\varepsilon_i = \sigma^2$).

На практике рекомендуется, чтобы n превышало k не менее, чем в три раза.

Находится оценка уравнения регрессии вида:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k.$$

Согласно методу наименьших квадратов, вектор оценок коэффициентов регрессии $\mathbf{b}$ получается по формуле:

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\,\mathbf{X}^T\mathbf{Y},$$

где

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} \ldots x_{1k} \\ \vdots & \vdots \quad\quad \vdots \\ 1 & x_{i1} \ldots x_{ik} \\ \vdots & \vdots \quad\quad \vdots \\ 1 & x_{n1} \ldots x_{nk} \end{pmatrix}; \quad \mathbf{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}; \quad \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_j \\ \vdots \\ b_k \end{pmatrix};$$

$\mathbf{X}^T$ – транспонированная матрица $\mathbf{X}$; $(\mathbf{X}^T\mathbf{X})^{-1}$ – матрица, обратная матрице $\mathbf{X}^T\mathbf{X}$.

Оценка ковариационной матрицы коэффициентов регрессии вектора $\mathbf{b}$ определяется из выражения:

$$\mathbf{S}(b) = \hat{S}^2(\mathbf{X}^T\mathbf{X})^{-1},$$

где $\hat{S}^2 = \dfrac{1}{n-k-1}(\mathbf{Y}-\mathbf{Xb})^T(\mathbf{Y}-\mathbf{Xb})$.

Учитывая, что на главной диагонали ковариационной матрицы находятся дисперсии коэффициентов в регрессии, имеем:

$$\hat{S}^2_{b_{(j-1)}} = \hat{S}^2\left[(\mathbf{X}^T\mathbf{X})^{-1}\right]_{jj} \text{ для } j = 1, 2, \ldots, k, k+1.$$

Значимость уравнения регрессии, т.е. гипотеза H_0: $\beta = 0$ ($\beta_0 = \beta_1 = \ldots = \beta_k = 0$), проверяется по F-критерию, наблюдаемое значение которого определяется по формуле:

$$F_{\text{набл}} = \frac{Q_R/(k+1)}{Q_{\text{ост}}/(n-k-1)},$$

где $Q_R = (\mathbf{Xb})^T(\mathbf{Xb})$, $Q_{\text{ост}} = (\mathbf{Y}-\mathbf{Xb})^T(\mathbf{Y}-\mathbf{Xb})$.

По таблице F-распределения для заданных α, $\nu_1 = k+1$, $\nu_2 = n-k-1$ находят $F_{\text{кр}}$.

Гипотеза H_0 отклоняется с вероятностью α, если $F_{\text{набл}} > F_{\text{кр}}$. Из этого следует, что уравнение является значимым, т.е. хотя бы один из коэффициентов регрессии отличен от нуля.

Для проверки значимости отдельных коэффициентов регрессии, т.е. гипотез H_0: $\beta_j = 0$, где $j = 1, 2, \ldots, k$, используют t-критерий и вычисляют: $t_{\text{набл}}(b_j) = b_j/\hat{s}_{b_j}$. По таблице t-распределения для заданного α, $\nu = n-k-1$ находят $t_{\text{кр}}$.

Гипотеза H_0 отвергается с вероятностью ошибки α, если $|t_{\text{набл}}| > t_{\text{кр}}$. Из этого следует, что соответствующий коэффициент регрессии β_j значим, т.е. $\beta_j \neq 0$. В противном случае коэффициент регрессии незначим и соответствующая переменная в модель не включается. Тогда реализуется алгоритм пошагового регрессионного анализа, состоящий в том, что исключается одна из незначимых переменных, которой соответствует минимальное по абсолютной величине значение $t_{\text{набл}}$. После этого вновь проводят регрессионный анализ с числом факторов, уменьшенным на единицу. Алгоритм заканчивается получением уравнения регрессии со значимыми коэффициентами.

Существуют и другие алгоритмы пошагового регрессионного анализа, например с последовательным включением факторов.

1.2.1. Назначение программ

Система включает в себя три программы, которые позволяют получать следующие результаты:

матрицу $(\mathbf{X}^T\mathbf{X})$ и вектор $(\mathbf{X}^T\mathbf{Y})$ – программа 1.2.2 (табл. 1.2 и 1.3);

обратную матрицу $(\mathbf{X}^T\mathbf{X})^{-1}$ и вектор b – программа 1.2.3 (табл. 1.4 и 1.5);

показатели точности модели: векторы $\mathbf{e}$, δ, а также $\hat{\mathbf{y}}$; суммы квадратов $Q_R = \sum_{i=1}^{n} \hat{y}_i^2$; $Q_{\text{ост}} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ – программа 1.2.4 (табл. 1.6 и 1.7).

В программе 1.2.3 для списывания первой строки обратной матрицы $(\mathbf{X}^T\mathbf{X})^{-1}$ предусмотрен останов машины.

Входными данными для программ 1.2.2 и 1.2.4 являются столбцы x_{i1}, x_{i2} и y_i табл. 1.1.

Таблица 1.1. **Исходная информация для анализа и результаты расчета**

№ п/п	x_1	x_2	y	$e=y-\hat{y}$	$\delta=e/\hat{y}$
1	x_{11}	x_{12}	y_1	e_1	δ_1
2	x_{21}	x_{22}	y_2	e_2	δ_2
.	.	.	.	.	.
.	.	.	.	.	.
i	x_{i1}	x_{i2}	y_i	e_i	δ_i
.	.	.	.	.	.
.	.	.	.	.	.
n	x_{n1}	x_{n2}	y_n	e_n	δ_n

1.2.2. Программа вычисления матрицы $(\mathbf{X}^T\mathbf{X})$ и вектора $(\mathbf{X}^T\mathbf{Y})$

Таблица 1.2. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 1.3)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
4	Занести нули в регистры памяти 1, 2,..., 9, d	Сх x → П 1 x → П 2 ...... x → П 9 x → П d	
5	Ввести в машину очередные наблюдения x_{i1}, x_{i2} и y_i	Набрать x_{i1} В↑ x_{i2} В↑ y_i	
6	Запустить программу с шага 00	В/О С/П	При зацикливании нажать клавишу С/П и перейти к п. 10
7	Если введены не все данные, т.е. (x_{i1}, x_{i2}, y_i), то вновь возвращаемся к п. 5, иначе переходим к п. 8		
8	Списать элементы матрицы $(\mathbf{X}^T\mathbf{X})$	П → x 1 П → x 2 П → x 3 П → x 4 П → x 5 П → x 6	На табло: n Σx_{i1} Σx_{i2} Σx_{i1}^2 $\Sigma x_{i1}x_{i2}$ Σx_{i2}^2
9	Списать элементы матрицы $(\mathbf{X}^T\mathbf{Y})$	П → x 7 П → x 8 П → x 9	На табло: Σy_i $\Sigma y_i x_{i1}$ $\Sigma y_i x_{i2}$
10	Проверить правильность ввода исходных данных	П → x a П → x b П → x c	На табло: x_{i1} x_{i2} y_i
11	Списать значение	П → x d	На табло: Σy_i^2
12	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора программы 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции	F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ . . . . $\overrightarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл. 1.3) Набрать правильную команду

Таблица 1.3. **Текст программы вычисления матрицы $(X^T X)$ и вектора $(X^T Y)$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x → П c	4C	27	x → П 6	46
01	F ○	25	28	П → x c	6C
02	x → П b	4L	29	П → x 7	67
03	F ○	25	30	+	10
04	x → П a	4–	31	x → П 7	47
05	F x^2	22	32	П → x a	6–
06	П → x 4	64	33	П → x c	6C
07	+	10	34	x	12
08	x → П 4	44	35	П → x 8	68
09	П → x a	6–	36	+	10
10	П → x 2	62	37	x → П 8	48
11	+	10	38	П → x b	6L
12	x → П 2	42	39	П → x c	6C
13	П → x b	6L	40	x	12
14	П → x 3	63	41	П → x 9	69
15	+	10	42	+	10
16	x → П 3	43	43	x → П 9	49
17	П → x a	6–	44	П → x 1	61
18	П → x b	6L	45	1	01
19	x	12	46	+	10
20	П → x 5	65	47	x → П 1	41
21	+	10	48	П → x c	6C
22	x → П 5	45	49	F x^2	22
23	П → x b	6L	50	П → x d	6Г
24	F x^2	22	51	+	10
25	П → x 6	66	52	x → П d	4Г
26	+	10	53	С/П	50

1.2.3. Программа вычисления обратной матрицы $(X^T X)^{-1}$ и вектора b

Таблица 1.4. **Инструкция по работе с программой вычисления обратной матрицы $(X^T X)^{-1}$ и вектора b**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/0 F ПРГ	
2	Записать программу в память	По тексту программы (табл. 1.5)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную программу

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
3	Перевести машину в автоматический режим работы	F АВТ	
4	Ввести в машину элементы матрицы $(X^T X)$ и вектора $(X^T Y)$	Набрать n x → П 1 Σx_{i1} x → П 2 Σx_{i2} x → П 3 Σx_{i1}^2 x → П 4 $\Sigma x_{i1} x_{i2}$ x → П 5 Σx_{i2}^2 x → П 6 Σy_i x → П 7 $\Sigma y_i x_{i1}$ x → П 8 $\Sigma y_i x_{i2}$ x → П 9	
5	Ввести в машину команды	1 x → П 0 П → x 5 В↑ П → x 4 П → x 6	
6	Запустить программу с шага 00	В/О С/П	При зацикливании нажать клавишу С/П и перейти к п. 9
7	Списать значения элементов первой строки обратной матрицы $(X^T X)^{-1}$	П → x a П → x b П → x c	a_{11} a_{12} a_{13}
8	Запустить программу с шага 41	С/П	
9	После окончания счета списать остальные элементы обратной матрицы $(X^T X)^{-1}$, вектора **b**, определитель матрицы $(X^T X)$	П → x d П → x 3 П → x 5 П → x a П → x b П → x c П → x 0	a_{22} a_{23} a_{33} b_0 b_1 b_2 $\|X^T X\|$
10	При ошибочном наборе номера подпрограммы, например 78, необходимо вернуться на два шага назад и набрать правильный номер подпрограммы	$\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ ПП 78	

Примечания: 1. П. 4 выполняется, если после проведения расчетов по программе 1.2.2 микрокалькулятор выключался или в регистрах памяти 1–9 не сохранены результаты расчетов по этой программе.

2. Перед выполнением п. 8 инструкции необходимо набрать П → x c.

Таблица 1.5. **Текст программы вычисления матрицы $(X^T X)^{-1}$ и вектора b**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	ПП	53	44	87	87
01	78	78	45	x → П a	4–
02	x → П a	4–	46	П → x 3	63
03	П → x 2	62	47	В↑	0Е
04	П → x 6	66	48	П → x 6	66
05	П → x 3	63	49	П → x 1	61
06	П → x 5	65	50	ПП	53
07	ПП	53	51	78	78
08	78	78	52	x → П d	4Г
09	x → П b	4*L*	53	П → x 5	65
10	П → x 3	63	54	П → x 1	61
11	П → x 4	64	55	П → x 2	62
12	П → x 5	65	56	П → x 3	63
13	П → x 2	62	57	ПП	53
14	ПП	53	58	78	78
15	78	78	59	x → П 3	43
16	x → П c	4*C*	60	П → x d	6Г
17	П → x 3	63	61	П → x b	6*L*
18	x	12	62	ПП	53
19	П → x 1	61	63	87	87
20	П → x a	6–	64	x → П b	4*L*
21	x	12	65	П → x 2	62
22	+	10	66	В↑	0Е
23	П → x 2	62	67	П → x 4	64
24	П → x b	6*L*	68	П → x 1	61
25	x	12	69	ПП	53
26	+	10	70	78	78
27	x → П 0	40	71	x → П 5	45
28	П → x a	6–	72	П → x 3	63
29	П → x 0	60	73	П → x c	6*C*
30	:	13	74	ПП	53
31	x → П a	4–	75	87	87
32	П → x b	6*L*	76	x → П c	4*C*
33	П → x 0	60	77	С/П	50
34	:	13	78	x	12
35	x → П b	4*L*	79	F ○	25
36	П → x c	6*C*	80	x	12
37	П → x 0	60	81	↔	14
38	:	13	82	F ○	25
39	x → П c	4*C*	83	–	11
40	С/П	50	84	П → x 0	60
41	П → x b	6*L*	85	:	13
42	П → x a	6–	86	В/0	52
43	ПП	53	87	П → x 7	67

Шаг	Клавиши	Код	Шаг	Клавиши	Код
88	x	12	93	↔	14
89	↔	14	94	П → x 9	69
90	П → x 8	68	95	x	12
91	x	12	96	+	10
92	+	10	97	В/О	52

1.2.4. Программа вычисления выборочных характеристик $\hat{y}_i$, e_i, δ_i, Σe_i^2, $\Sigma \hat{y}_i^2$

Таблица 1.6. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 1.7)	Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 6 и 7 нули	Сх x → П 6 x → П 7	
5	Ввести в регистры памяти a, b, c элементы вектора **b**	Набрать b_0 x → П a b_1 x → П b b_2 x → П c	
6	Ввести в машину очередные данные x_{i1}, x_{i2}, y_i и запустить программу с шага 00	Набрать x_{i1} В↑ x_{i2} В↑ y_i В/О С/П	
7	Записать значения $\hat{y}_i$, e_i и δ_i в i-ю строку табл. 1.1	П → x 3 П → x 4 П → x 5	На табло: $\hat{y}_i$ $e_i = y_i - \hat{y}_i$ $\delta_i = e_i/\hat{y}_i$
8	Если введены не все исходные данные табл. 1.1, то перейти к п. 6, иначе – к п. 9		
9	Записать в итоговую строку табл. 1.1 значения	П → x 6 П → x 7	На табло: Σe_i^2 $\Sigma \hat{y}_i^2$

Примечание. П. 5 выполняется, если после проведения расчетов по программе 1.2.3 микрокалькулятор выключался или в регистрах памяти a, b, c не сохранены результаты расчетов по этой программе.

Таблица 1.7. **Текст программы вычисления $\hat{y}_i$; e_i; $\delta_i = e_i/\hat{y}_i$; Σe_i^2; $\Sigma \hat{y}_i^2$ $(i = \overline{1, n})$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x → П 0	40	16	П → x 7	67
01	F ○	25	17	+	10
02	x → П 2	42	18	x → П 7	47
03	F ○	25	19	П → x 0	60
04	x → П 1	41	20	П → x 3	63
05	П → x a	6 –	21	–	11
06	П → x 1	61	22	x → П 4	44
07	П → x b	6*L*	23	П → x 3	63
08	x	12	24	:	13
09	+	10	25	x → П 5	45
10	П → x 2	62	26	П → x 4	64
11	П → x c	6*C*	27	F x^2	22
12	x	12	28	П → x 6	66
13	+	10	29	+	10
14	x → П 3	43	30	x → П 6	46
15	F x^2	22	31	С/П	50

1.2.5. Работа с системой программ

Работа с системой программ заключается в следующем:

1. Заготовить табл. 1.1, заполнив в ней столбцы x_{i1}, x_{i2} и y_i.

2. Выполнить инструкцию по работе с программой 1.2.2 (табл. 1.2).

После окончания вычислений по программе (табл. 1.3) в адресуемых регистрах памяти сохраняются значения следующих элементов симметрической матрицы $(\mathbf{X}^T\mathbf{X})$ и вектора $(\mathbf{X}^T\mathbf{Y})$:

$$(\mathbf{X}^T\mathbf{X}) = \begin{pmatrix} n & \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i2} \\ & \sum_{i=1}^{n} x_{i1}^2 & \sum_{i=1}^{n} x_{i1}x_{i2} \\ & & \sum_{i=1}^{n} x_{i2}^2 \end{pmatrix},$$

$$(\mathbf{X}^T\mathbf{Y}) = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_{i1}y_i \\ \sum_{i=1}^{n} x_{i2}y_i \end{pmatrix}.$$

3. Выполнить инструкцию по работе с программой 1.2.3 (табл. 1.4). В программе 1.2.3 (табл. 1.5) предусмотрен *останов машины* для списания первой строки обратной матрицы $(\mathbf{X}^T\mathbf{X})^{-1}$, т.е. a_{11}, a_{12}, a_{13}, где a_{ij} – элемент обратной матрицы, лежащей на пересечении i-й строки и

j-го столбца (i, j = 1, 2, 3). Входными данными программы являются 1 х→П 0, П→х 5, В↑, П→х 4, П→х 6. Перед пуском второй части программы 1.2.3 необходимо нажать клавиши П → х с. Вторая часть программы располагается в программной памяти с адреса 41. Результатом работы этой части программы являются значения элементов a_{22}, a_{23} и a_{33} обратной матрицы $(\mathbf{X}^T\mathbf{X})^{-1}$, а также элементы вектора коэффициентов регрессии $\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$. Все значения элементов обратной матрицы $(\mathbf{X}^T\mathbf{X})^{-1}$ и вектора $\mathbf{b}$ необходимо списать.

4. Выполнить инструкцию по работе с программой 1.2.4 (табл. 1.6).

Входными данными этой программы (табл. 1.7) являются значения x_{i1}, x_{i2} и y_i, содержащиеся в табл. 1.1.

Результатом работы программы 1.2.4 являются расчетные значения $\hat{y}_i$, абсолютные $e_i = y_i - \hat{y}_i$ и относительные $\delta_i = e_i/\hat{y}_i$ отклонения, которые заносятся в соответствующие столбцы табл. 1.1. После ввода последнего n-го наблюдения получаем суммы $\sum\limits_{i=1}^{n} e_i^2$ и $\sum\limits_{i=1}^{n} \hat{y}_i^2$.

5. После окончания вычислений по программе 1.2.4 в адресуемых регистрах памяти сохраняются значения величин $\sum\limits_{i=1}^{n} e_i^2$ и $\sum\limits_{i=1}^{n} \hat{y}_i^2$, которые можно использовать для дальнейших вычислений вручную при проверке значимости уравнения регрессии. На этом работа по *системе программ закончена. Следует помнить:* при ручных вычислениях содержимое адресуемых регистров памяти *не затрагивается,* поэтому в случае ошибки имеется возможность повторного счета.

1.3. СИСТЕМА ПРОГРАММ КОРРЕЛЯЦИОННОГО АНАЛИЗА

1.3.1. Программа вычисления оценок параметров трехмерного нормального закона распределения

Эта программа (табл. 1.8, 1.9) позволяет получать следующие оценки: вектор средних $\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{y} \end{pmatrix}$, вектор средних квадратических отклонений $\begin{pmatrix} s_1 \\ s_2 \\ s_y \end{pmatrix}$ и матрицу парных коэффициентов корреляции $\mathbf{R}$. Так как матрица $\mathbf{R}$ симметрическая с единицами на главной диагонали, то в регистрах памяти сохраняются значения, лежащие над главной диагональю:

$$\begin{pmatrix} 1 & r_{12} & r_{1y} \\ & 1 & r_{2y} \\ & & 1 \end{pmatrix}.$$

Таблица 1.8. **Инструкция по работе с программой корреляционного анализа**

№ п/п	Инструкция	Нажимаемые клавиши		Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ		
2	Записать программу в память	По тексту программы (табл. 1.9)		Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ		
4	Ввести в машину элементы матрицы $(\mathbf{X}^T\mathbf{X})$, вектора $(\mathbf{X}^T\mathbf{Y})$ и $\sum\limits_{i=1}^{n} y_i^2$	Набрать		
		n	x → П 1	
		$\sum\limits_{i=1}^{n} x_{i1}$	x → П 2	
		$\sum\limits_{i=1}^{n} x_{i2}$	x → П 3	
		$\sum\limits_{i=1}^{n} x_{i1}^2$	x → П 4	
		$\sum\limits_{i=1}^{n} x_{i1}x_{i2}$	x → П 5	
		$\sum\limits_{i=1}^{n} x_{i2}^2$	x → П 6	
		$\sum\limits_{i=1}^{n} y_i$	x → П 7	
		$\sum\limits_{i=1}^{n} y_i x_{i1}$	x → П 8	
		$\sum\limits_{i=1}^{n} y_i x_{i2}$	x → П 9	
		$\sum\limits_{i=1}^{n} y_i^2$	x → П d	
5	Запустить программу с шага 00	В/О С/П		
6	Списать оценки параметров корреляционного анализа и r_{y/x_1x_2}	Набрать		
		П → x 0		r_{y/x_1x_2}
		П → x 1		n
		П → x 2		$\bar{x}_1$
		П → x 3		$\bar{x}_2$
		П → x 4		s_1

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
		П → x 5 П → x 6 П → x 7 П → x 8 П → x 9 П → x d	r_{12} s_2 $\bar{y}$ r_{1y} r_{2y} s_y

Таблица 1.9. **Текст программы корреляционного анализа**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	П → x 2	62	35	:	13
01	П → x 1	61	36	П → x 6	66
02	:	13	37	:	13
03	x → П 2	42	38	x → П 5	45
04	П → x 3	63	39	F x^2	22
05	П → x 1	61	40	x → П a	4 –
06	:	13	41	П → x 8	68
07	x → П 3	43	42	П → x 1	61
08	П → x 7	67	43	:	13
09	П → x 1	61	44	П → x 2	62
10	:	13	45	П → x 7	67
11	x → П 7	47	46	x	12
12	П → x 2	62	47	–	11
13	П → x 4	64	48	П → x 4	64
14	ПП	53	49	:	13
15	87	87	50	П → x d	6*Г*
16	x → П 4	44	51	:	13
17	П → x 3	63	52	x → П 8	48
18	П → x 6	66	53	F x^2	22
19	ПП	53	54	x → П b	4*L*
20	87	87	55	П → x 9	69
21	x → П 6	46	56	П → x 1	61
22	П → x 7	67	57	:	13
23	П → x d	6*Г*	58	П → x 3	63
24	ПП	53	59	П → x 7	67
25	87	87	60	x	12
26	x → П d	4*Г*	61	–	11
27	П → x 5	65	62	П → x 6	66
28	П → x 1	61	63	:	13
29	:	13	64	П → x d	6*Г*
30	П → x 2	62	65	:	13
31	П → x 3	63	66	x → П 9	49
32	x	12	67	F x^2	22
33	–	11	68	x → П c	4*C*
34	П → x 4	64	69	П → x b	6*L*

Шаг	Клавиши	Код	Шаг	Клавиши	Код
70	П → x c	6C	82	−	11
71	+	10	83	:	13
72	2	02	84	F $\sqrt{\ }$	21
73	П → x 5	65	85	x → П 0	40
74	x	12	86	С/П	50
75	П → x 8	68	87	П → x 1	61
76	x	12	88	:	13
77	П → x 9	69	89	↔	14
78	x	12	90	F x^2	22
79	−	11	91	−	11
80	1	01	92	F $\sqrt{\ }$	21
81	П → x a	6 −	93	В/О	52

1.3.2. Программа вычисления оценок частных и множественных коэффициентов корреляции и детерминации

Эта программа (табл. 1.10, 1.11) позволяет получать оценки частных и множественных коэффициентов корреляции и детерминации.

Составим из частных коэффициентов корреляции матрицу. Так как матрица частных коэффициентов корреляции симметрическая с единицами на главной диагонали, то заполним в ней только элементы, лежащие над главной диагональю:

$$\begin{pmatrix} 1 & r_{12/y} & r_{1y/2} \\ & 1 & r_{2y/1} \\ & & 1 \end{pmatrix}.$$

Таблица 1.10. **Инструкция по работе с программой вычисления частных и множественных коэффициентов корреляции и детерминации**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 1.11)	Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ	
4	Подготовить счетчик цикла и адресную константу	В/О С/П	После останова на таблице число 13

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
5	Ввести в машину парные коэффициенты корреляции	Набрать r_{12} В↑ r_{1y} В↑ r_{2y}	
6	Вычисление частных коэффициентов корреляции, множественных коэффициентов корреляции и детерминации	С/П	На табло r_{2y}
7	Выписать из регистров памяти оценки параметров	П → x 4 П → x 5 П → x 6 П → x 7 П → x 8 П → x 9 П → x a П → x b П → x c П → x 2 П → x 3 П → x c	На табло: $r_{1/2y}$ $r^2_{1/2y}$ $r_{2y/1}$ $r_{2/1y}$ $r^2_{2/1y}$ $r_{1y/2}$ $r_{y/12}$ $r^2_{y/12}$ $r_{12/y}$ r_{2y} r_{12} r_{1y}
8	Для расчета с новыми значениями r_{12}, r_{1y} и r_{2y} перейти к п. 4		

Таблица 1.11. **Текст программы вычисления частных и множественных коэффициентов корреляции и детерминации**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	3	03	16	П → x 3	63
01	x → П 0	40	17	F x^2	22
02	1	01	18	−	11
03	3	03	19	1	01
04	x → П 1	41	20	П → x d	6Г
05	С/П	50	21	F x^2	22
06	x → П d	4Г	22	−	11
07	F ○	25	23	x	12
08	x → П 3	43	24	F $\sqrt{}$	21
09	F ○	25	25	:	13
10	x → П 2	42	26	К x → П 1	L1
11	П → x 3	63	27	П → x d	6Г
12	П → x d	6Г	28	F x^2	22
13	x	12	29	П → x 3	63
14	−	11	30	F x^2	22
15	1	01	31	+	10

Шаг	Клавиши	Код	Шаг	Клавиши	Код
32	2	02	43	−	11
33	П → x 2	62	44	:	13
34	x	12	45	К x → П 1	L1
35	П → x 3	63	46	F $\sqrt{\ }$	21
36	x	12	47	К x → П 1	L1
37	П → x d	6Г	48	П → x 3	63
38	x	12	49	П → x d	6Г
39	−	11	50	П → x 2	62
40	1	01	51	F L_0	5Г
41	П → x 2	62	52	0 6	06
42	F x^2	22	53	С/П	50

1.4. КОНТРОЛЬНЫЙ ПРИМЕР

По данным годовых отчетов десяти (n = 10) промышленных предприятий (табл. 1.12) провести регрессионный анализ зависимости себестоимости товарной продукции y (млн. руб.) от объема валовой продукции x_1 (млн. руб.) и производительности труда x_2 (тыс. руб. на чел.).

Таблица 1.12. **Исходная информация для анализа и результаты расчета**

№ п/п	x_{i1}	x_{i2}	y_i	$\hat{y}_i$	$e_i = y_i - \hat{y}_i$	$\delta_i = e_i/\hat{y}_i$
1	3	1,8	2,1	2,31472	−0,21472	−0,09276
2	4	1,5	2,8	3,48755	−0,68755	−0,19714
3	5	1,4	3,2	4,35777	−1,15777	−0,26568
4	5	1,3	4,5	4,50907	−0,00907	−0,00201
5	5	1,3	4,8	4,50907	0,29093	0,06452
6	5	1,5	4,9	4,20647	0,69353	0,16487
7	6	1,6	5,5	4,77408	0,72592	0,15205
8	7	1,2	6,5	6,09821	0,40179	0,06589
9	15	1,3	12,1	11,69825	0,40175	0,03434
10	20	1,2	15,0	15,44415	−0,44415	−0,02876

Решение. Определим вектор оценок коэффициентов регрессии уравнения

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2.$$

Согласно методу наименьших квадратов, вектор $\mathbf{b}$ получается из выражения $\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$, где

$$\mathbf{X}=\begin{pmatrix}1 & x_{11} & x_{12}\\ 1 & x_{21} & x_{22}\\ \cdot & \cdot\ \cdot & \cdot\ \cdot\\ 1 & x_{101} & x_{102}\end{pmatrix},\quad \mathbf{Y}=\begin{pmatrix}y_1\\ y_2\\ \cdots\\ y_{10}\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix}b_0\\ b_1\\ b_2\end{pmatrix};$$

$\mathbf{X}^T$ – транспонированная матрица $\mathbf{X}$; $(\mathbf{X}^T\mathbf{X})^{-1}$ – матрица, обратная матрице $\mathbf{X}^T\mathbf{X}$.

После реализации программы 1.2.2 получаем элементы матрицы $\mathbf{X}^T\mathbf{X}$ и вектора $\mathbf{X}^T\mathbf{Y}$. Так как матрица $\mathbf{X}^T\mathbf{X}$ симметрическая

$$\mathbf{X}^T\mathbf{X}=\begin{pmatrix} n & \sum\limits_{i=1}^{n} x_{i1} & \sum\limits_{i=1}^{n} x_{i2}\\ \sum\limits_{i=1}^{n} x_{i1} & \sum\limits_{i=1}^{n} x_{i1}^2 & \sum\limits_{i=1}^{n} x_{i1}x_{i2}\\ \sum\limits_{i=1}^{n} x_{i2} & \sum\limits_{i=1}^{n} x_{i1}x_{i2} & \sum\limits_{i=1}^{n} x_{i2}^2\end{pmatrix},\quad n=10,$$

в памяти хранятся только диагональные и наддиагональные ее элементы.

Для данного примера из регистров памяти извлекаем: $n = 10$, $\sum\limits_{i=1}^{10} x_{i1} = 75$, $\sum\limits_{i=1}^{10} x_{i2} = 14{,}1$, $\sum\limits_{i=1}^{10} x_{i1}^2 = 835$, $\sum\limits_{i=1}^{10} x_{i1}x_{i2} = 100{,}4$, $\sum\limits_{i=1}^{10} x_{i2}^2 = 20{,}21$. В результате формируем матрицу

$$\mathbf{X}^T\mathbf{X}=\begin{pmatrix}10 & 75 & 14{,}1\\ 75 & 835 & 100{,}4\\ 14{,}1 & 100{,}4 & 20{,}21\end{pmatrix}$$

и вектор

$$\mathbf{X}^T\mathbf{Y}=\begin{pmatrix}\sum\limits_{i=1}^{n} y_i\\ \sum\limits_{i=1}^{n} x_{i1}y_i\\ \sum\limits_{i=1}^{n} x_{i2}y_i\end{pmatrix}=\begin{pmatrix}61{,}4\\ 664{,}5\\ 82{,}23\end{pmatrix}.$$

По программе 1.2.3 рассчитываются элементы обратной матрицы $(\mathbf{X}^T\mathbf{X})^{-1}$ и вектор оценок коэффициентов регрессии $\mathbf{b}$. Обозначим

$$(\mathbf{X}^T\mathbf{X})^{-1}=\mathbf{A}=\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}.$$

Так как эта матрица симметрическая, то $a_{lj} = a_{jl}$ для $j, l = 1, 2, 3$.

После первого останова программы 1.2.3 из регистров памяти следует *обязательно* выписать элементы a_{11}, a_{12}, a_{13}.

Для данного примера имеем: $a_{11} = 11{,}13418$, $a_{12} = -0{,}16403$, $a_{13} = -6{,}95314$.

В результате реализации программы 1.2.3 получаем $a_{22} = 0{,}00539$, $a_{23} = 0{,}08766$, $a_{33} = 4{,}46502$, определитель матрицы $|\mathbf{X}^T\mathbf{X}| = 610{,}3$, а также элементы вектора

$$\mathbf{b} = \begin{pmatrix} 2{,}88142 \\ 0{,}71892 \\ -1{,}51303 \end{pmatrix}.$$

Формируем обратную матрицу

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 11{,}1341 & -0{,}16403 & -6{,}95314 \\ -0{,}16403 & 0{,}00539 & 0{,}08766 \\ -6{,}95314 & 0{,}08766 & 4{,}46502 \end{pmatrix}$$

и оценку уравнения регрессии

$$\hat{y} = 2{,}88142 + 0{,}71892x_1 - 1{,}51303x_2.$$

Программа 1.2.4 предназначена для статистического анализа полученного уравнения регрессии: проверки значимости уравнения и его коэффициентов, исследования абсолютных e_i и относительных δ_i ошибок аппроксимации. Хранящиеся в регистрах памяти значения $\hat{y}_i$, $e_i = y_i - \hat{y}_i$ и $\delta_i = e_i/\hat{y}_i$ заносятся в табл. 1.1. Эти значения получаются в регистрах памяти на основании реализации программы 1.2.4 после введения в микрокалькулятор очередных значений x_{i1}, x_{i2}, y_i, $i = 1, 2, \ldots, n$. После введения последней тройки (x_{n1}, x_{n2}, y_n) из регистров памяти выписываются значения

$$Q_{\text{ост}} = \sum_{i=1}^{n} e_i^2 = 3{,}47203, \quad Q_R = \sum_{i=1}^{n} \hat{y}_i^2 = 530{,}21976.$$

Тогда несмещенная оценка остаточной дисперсии

$$\hat{S}^2 = \frac{1}{n-3} \cdot Q_{\text{ост}} = \frac{1}{7} \cdot 3{,}47203 = 0{,}49601,$$

а оценка среднего квадратического отклонения

$$\hat{S} = \sqrt{\hat{S}^2} = 0{,}70427533.$$

Проверяем на уровне значимости $\alpha = 0{,}05$ значимость уравнения регрессии, т.е. гипотезу H_0: $\beta = 0$ $(\beta_0 = \beta_1 = \beta_2 = 0)$. Для этого вычисляем

$$F_{\text{набл}} = \frac{Q_R/(k+1)}{Q_{\text{ост}}/(n-k-1)} = \frac{1/3 \cdot 530{,}21976}{1/7 \cdot 3{,}47203} = 356{,}32776.$$

По таблице F-распределения для $\alpha = 0{,}05$, $\nu_1 = 3$ и $\nu_2 = 7$ находим $F_{кр} = 4{,}35$.

Так как $F_{набл} > F_{кр}$ $(356{,}328 > 4{,}35)$, гипотеза отвергается с вероятностью ошибки 0,05. Таким образом уравнение является значимым, т.е. хотя бы один из коэффициентов регрессии отличен от нуля.

Перед проверкой значимости отдельных коэффициентов регрессии найдем оценку ковариационной матрицы вектора $\mathbf{b}$:

$$\hat{\mathbf{S}}(\mathbf{b}) = \hat{\mathbf{S}}^2(\mathbf{X}^T\mathbf{X})^{-1} = \begin{pmatrix} 5{,}52259 & -0{,}08136 & -3{,}44878 \\ -0{,}08136 & 0{,}00267 & 0{,}04348 \\ -3{,}44878 & 0{,}04348 & 2{,}21466 \end{pmatrix}.$$

Учитывая, что на главной диагонали ковариационной матрицы находятся дисперсии коэффициентов регрессии, получаем следующие несмещенные оценки этих дисперсий:

$\hat{S}^2_{b_0} = 5{,}52259$ $(\hat{S}_{b_0} = 2{,}35002)$;
$\hat{S}^2_{b_1} = 0{,}00267$ $(\hat{S}_{b_1} = 0{,}05171)$;
$\hat{S}^2_{b_2} = 2{,}21467$ $(\hat{S}_{b_2} = 1{,}48818)$,

и имеем оценку корреляционной матрицы вектора $\mathbf{b}$ с элементами, определяемыми по формуле

$$r_{j-1\,l-1} = \frac{\hat{\mathrm{cov}}(b_{j-1}, b_{l-1})}{\hat{S}_{b_{j-1}} \hat{S}_{b_{l-1}}},$$

где $\hat{\mathrm{cov}}(b_{j-1}, b_{l-1})$ – элементы матрицы $\hat{\mathbf{S}}(\mathbf{b})$, стоящие на пересечении j-й строки и l-го столбца, $j, l = 1, 2, 3$.

Находим оценку корреляционной матрицы:

$$\hat{\mathbf{R}}(\mathbf{b}) = \begin{pmatrix} 1 & -0{,}66955 & -0{,}98614 \\ -0{,}66955 & 1 & 0{,}56504 \\ -0{,}98614 & 0{,}56504 & 1 \end{pmatrix}.$$

Для проверки значимости отдельных коэффициентов регрессии, т.е. гипотез H_0: $\beta_m = 0$, где $m = 0, 1, 2$, находим по таблицам F-распределения для $\alpha = 0{,}05$, $\nu_1 = 1$, $\nu_2 = 7$ критическое значение $F_{кр} = 5{,}59$.

Вычисляем $F_{набл}$ для каждого из коэффициентов регрессии по формуле

$$F_{набл}(b_m) = b_m^2 / \hat{S}^2_{b_m}, \quad m = 1, 2.$$

Подставляя данные, получаем:

$$F_{\text{набл}}(b_1) = \frac{0{,}51684338}{0{,}0026738} = 193{,}29919;$$

$$F_{\text{набл}}(b_2) = \frac{2{,}2892597}{2{,}2146648} = 1{,}0336822.$$

Так как $F_{\text{набл}}(b_1)$ больше $F_{\text{кр}} = 5{,}59$, то коэффициент регрессии β_1 значимо отличается от нуля.

Для коэффициента β_2 выполняется неравенство $F_{\text{набл}}(b_2) < F_{\text{кр}} = 5{,}59$, поэтому данный коэффициент можно считать равным нулю. Переходя к алгоритму пошагового регрессионного анализа, будем рассматривать оценку уравнения регрессии вида $\hat{y} = b'_0 + b'_1 x_1$. В этом случае вектор оценок $\mathbf{b}'$ определим по формуле

$$\mathbf{b}' = (\mathbf{X}'^T\mathbf{X}')^{-1}\mathbf{X}'^T\mathbf{Y},$$

где

$$\mathbf{X}' = \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix}; \quad (\mathbf{X}'^T\mathbf{X}') = \begin{pmatrix} n & \sum\limits_{i=1}^{10} x_{i1} \\ \sum\limits_{i=1}^{10} x_{i1} & \sum\limits_{i=1}^{10} x_{i1}^2 \end{pmatrix};$$

$$(\mathbf{X}'^T\mathbf{Y}) = \begin{pmatrix} \sum\limits_{i=1}^{n} y_i \\ \sum\limits_{i=1}^{n} x_i\, y_i \end{pmatrix}.$$

Используя ранее проведенные вычисления, получим

$$(\mathbf{X}'^T\mathbf{X}') = \begin{vmatrix} 10 & 75 \\ 75 & 835 \end{vmatrix}; \quad (\mathbf{X}'^T\mathbf{Y}) = \begin{vmatrix} 61{,}4 \\ 664{,}5 \end{vmatrix}.$$

Отсюда

$$(\mathbf{X}'^T\mathbf{X}')^{-1} = \frac{1}{10 \cdot 835 - (75)^2} \begin{vmatrix} 835 & -75 \\ -75 & 10 \end{vmatrix} =$$

$$= \begin{vmatrix} 0{,}306422 & -0{,}0275229 \\ -0{,}0275229 & 0{,}0036697 \end{vmatrix},$$

$$\mathbf{b}' = \begin{pmatrix} 0{,}306422 & -0{,}0275229 \\ -0{,}0275299 & 0{,}0036697 \end{pmatrix} \cdot \begin{pmatrix} 61{,}4 \\ 664{,}5 \end{pmatrix} = \begin{pmatrix} 0{,}5253430 \\ 0{,}7486096 \end{pmatrix}.$$

Тогда оценка уравнения регрессии будет иметь вид

$$\hat{y} = 0{,}52534 + 0{,}74861 x_1.$$

Для статистического анализа полученного уравнения регрессии воспользуемся программой 1.2.4. Введя в микрокалькулятор последовательно тройки значений

x_{i1}	3	4	5	5	5	5	6	7	15	20
x_{i2}	0	0	0	0	0	0	0	0	0	0
y_i	2,1	2,8	3,2	4,5	4,8	4,9	5,5	6,5	12,1	15,0

и приняв $b_2 = 0$, получим после введения последней тройки наблюдений

$$Q'_{\text{ост}} = \sum_{i=1}^{n} e_i^2 = 3{,}9847341; \quad Q'_R = \sum_{i=1}^{n} \hat{y}_i^2 = 529{,}69917.$$

Тогда несмещенная оценка остаточной дисперсии равна

$$\hat{s}'^2 = \frac{1}{n-2} Q'_{\text{ост}} = \frac{1}{8} 3{,}9847314 = 0{,}49809176,$$

а оценка среднего квадратического отклонения $\hat{s}' = 0{,}70575616$.

Проверяем при $\alpha = 0{,}05$ значимость уравнения регрессии, т.е. гипотезу H_0: $\beta = 0$ $(\beta_0 = \beta_1 = 0)$. Для этого вычисляем

$$F'_{\text{набл}} = \frac{1/2\, Q'_R}{1/8\, Q'_{\text{ост}}} = \frac{264{,}84958}{0{,}49809176} = 531{,}72849.$$

По таблице F-распределения для $\alpha = 0{,}05$, $\nu'_1 = 2$ и $\nu'_2 = 8$ находим $F_{\text{кр}} = 4{,}46$. Так как $F'_{\text{набл}} > F'_{\text{кр}}$, то уравнение является значимым.

Найдем оценку ковариационной матрицы вектора $\mathbf{b}'$

$$\hat{S}(\mathbf{b}') = \hat{s}'^2 (\mathbf{X}'^T \mathbf{X}')^{-1} = \begin{matrix} 0{,}306422 & -0{,}0275299 \\ -0{,}0275229 & 0{,}0036697 \end{matrix} \cdot 0{,}49809176 =$$

$$= \begin{pmatrix} 0{,}15262627 & -0{,}013712416 \\ -0{,}013712416 & 0{,}0018278473 \end{pmatrix}.$$

Отсюда получаем несмещенные оценки дисперсий коэффициентов регрессии

$$\hat{s}^2_{b'_0} = 0{,}15262627 \qquad (\hat{s}_{b'_0} = 0{,}3906741),$$
$$\hat{s}^2_{b'_1} = 0{,}0018278473 \qquad (\hat{s}_{b'_1} = 0{,}0427527).$$

Для проверки значимости коэффициента регрессии, т.е. гипотезы H_0: $\beta_1 = 0$, находим по таблице F-распределения при $\alpha = 0{,}05$, $\nu_1 = 1$, $\nu_2 = 8$ значение $F_{кр} = 5{,}32$. Так как $F_{набл}(b'_1) = 306{,}599$ больше $F_{кр} = 5{,}32$, то коэффициент регрессии β_1 значимо отличается от нуля. Таким образом, окончательное уравнение регрессии имеет вид

$$\hat{y} = 0{,}52534 + 0{,}74861x_1.$$

Программа 1.3.1 предназначена для оценки параметров трехмерного нормального закона распределения, вектора средних $(\bar{x}_1, \bar{x}_2, \bar{y})^T$, вектора средних квадратических отклонений $(s_1, s_2, s_y)^T$ и корреляционной матрицы

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & r_{1y} \\ r_{21} & 1 & r_{2y} \\ r_{y1} & r_{y2} & 1 \end{pmatrix}, \text{ где } r_{12} = \frac{\overline{x_1 x_2} - \bar{x}_1 \bar{x}_2}{S_1 S_2},$$

а также оценки множественного коэффициента корреляции r_{y/x_1x_2}. Множественный коэффициент детерминации определяется по формуле:

$$r^2_{y/x_1x_2} = \frac{r^2_{1y} + r^2_{2y} - 2r_{12}r_{1y}r_{2y}}{1 - r^2_{12}}.$$

Так как матрица R симметрическая, то $r_{lj} = r_{jl}$ для $j, l = 1, 2, 3$.

Программа предусматривает *обязательное занесение* в соответствующие регистры памяти результатов расчетов по программе 1.2.2 (см. инструкцию по работе с программой 1.3.1).

Для данных примера имеем:

$\bar{x}_1 = 7{,}5$; $\quad s_1 = 5{,}22$;
$\bar{x}_2 = 1{,}41$; $\quad s_2 = 0{,}18$;
$\bar{y} = 6{,}14$; $\quad s_y = 3{,}91$;
$r_{12} = -0{,}565$; $r_{1y} = 0{,}997$; $r_{2y} = -0{,}612$; $r_{y/x_1x_2} = 0{,}998$.

Формируем корреляционную матрицу

$$\mathbf{R} = \begin{pmatrix} 1 & -0{,}565 & 0{,}997 \\ & 1 & -0{,}612 \\ & & 1 \end{pmatrix}.$$

По программе 1.3.2 находят оценки частных и множественных коэффициентов корреляции и детерминации.

Оценки частных коэффициентов корреляции:

$$r_{12/y} = \frac{r_{12} - r_{1y} \cdot r_{2y}}{\sqrt{(1 - r_{1y}^2)(1 - r_{2y}^2)}} = \frac{-0{,}565 - 0{,}997(-0{,}612)}{\sqrt{(1 - 0{,}997^2)(1 - 0{,}612^2)}} = 0{,}738.$$

$$r_{1y/2} = \frac{0{,}997 - 0{,}565 \cdot 0{,}612}{\sqrt{(1 - 0{,}565^2)(1 - 0{,}612^2)}} = 0{,}998; \quad r_{2y/1} = -0{,}762.$$

По полученным в регистрах памяти данным составим матрицу частных коэффициентов корреляции:

$$\begin{pmatrix} 1 & 0{,}738 & 0{,}998 \\ & 1 & -0{,}762 \\ & & 1 \end{pmatrix}.$$

Следует иметь в виду, что частный коэффициент корреляции может резко отличаться от соответствующего парного коэффициента и даже иметь противоположный знак. Любой из частных коэффициентов может быть равен нулю, в то время как парный – отличен от нуля. В примере $r_{12} = -0{,}565$, а $r_{12/y} = 0{,}738$. Такое различие вызвано тесной связью объема валовой продукции X_1 и себестоимостью Y товарной продукции ($r_{1y} = 0{,}997$). В случае независимости величин частный и парный коэффициенты корреляции равны нулю.

Значимость частных коэффициентов корреляции проверим с помощью таблицы Фишера–Иейтса, согласно которой $\alpha = 0{,}05$ и $\nu = n - 3 = 7$ соответствует критическое значение $r_{кр}(0{,}05; 7) = 0{,}666$. Так как все найденные значения частных коэффициентов корреляции по модулю больше $r_{кр}$, то гипотеза о равенстве нулю каждого из генеральных частных коэффициентов корреляции отвергается с вероятностью ошибки, равной 0,05. Следовательно, доказана "чистая", истинная зависимость между x_3 и y, x_2 и y, x_1 и x_2, не обусловленная влиянием третьей случайной величины.

Оценки множественных коэффициентов корреляции и детерминации соответственно равны:

$$r_{1/2y} = 0{,}998; \quad r^2_{1/2y} = 0{,}997;$$
$$r_{2/1y} = 0{,}846; \quad r^2_{2/1y} = 0{,}715;$$
$$r_{y/12} = 0{,}999; \quad r^2_{y/12} = 0{,}997.$$

В качестве примера проверим значимость множественного коэффициента детерминации $\rho^2_{y/12}$, т.е. гипотезы H_0: $\rho^2_{y/12} = 0$. С этой целью найдем

$$F_{\text{набл}} = \frac{\frac{1}{2} r^2_{y/12}}{\frac{1}{3}\cdot(1 - r^2_{y/12})} = \frac{\frac{1}{2}\, 0{,}997}{\frac{1}{3}\, 0{,}003} = 498{,}5.$$

По таблице F-распределения для $\alpha = 0{,}05$, $\nu_1 = 2$ и $\nu_3 = 3$ найдем $F_{\text{кр}} = 9{,}55$. Так как $F_{\text{набл}} > F_{\text{кр}}$, то гипотеза о равенстве $\rho^2_{y/12} = 0$ отвергается. Следовательно, доказано наличие зависимости y от x_1 и x_2, т.е. себестоимость продукции зависит от объема валовой продукции и производительности труда.

2. КОМПОНЕНТНЫЙ АНАЛИЗ

2.1. АЛГОРИТМ КОМПОНЕНТНОГО АНАЛИЗА

Компонентный анализ предназначен для преобразования системы k исходных признаков в систему k новых показателей (главных компонент). Главные компоненты не коррелированы между собой и упорядочены по величине их дисперсий, причем первая главная компонента имеет наибольшую дисперсию, а последняя, k-я, – наименьшую.

Компонентный анализ является одним из основных методов факторного анализа. В задачах снижения размерности и классификации обычно используются m первых компонент ($m \ll k$).

При наличии результативного показателя Y может быть построено уравнение регрессии на главных компонентах.

На основании матрицы исходных данных

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \dots & \dots & \dots \\ x_{i1} & x_{i2} & x_{i3} \\ \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} \end{pmatrix}, \quad i = 1, 2, \dots, n,$$

вычисляем оценки параметров распределения трехмерной генеральной совокупности $\bar{x}_1, \bar{x}_2, \bar{x}_3, s_1, s_2, s_3, r_{12}, r_{13}$ и r_{23}, где $\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$;

$$s_j^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 = \frac{1}{n}\sum_{i=1}^{n} x_{ij}^2 - (\bar{x}_j)^2;$$

$$r_{j\nu} = \frac{\overline{x_j x_\nu} - \bar{x}_j \bar{x}_\nu}{s_j s_\nu}; \quad \overline{x_j x_\nu} = \frac{1}{n}\sum_{i=1}^{n} x_{ij} x_{i\nu}; \; j, \nu = 1, 2, 3.$$

Таким образом получаем оценку матрицы парных коэффициентов корреляции

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{pmatrix}.$$

Преобразуем матрицу **R** в диагональную матрицу Λ собственных значений характеристического многочлена $|\lambda\mathbf{E} - \mathbf{R}|$.

Характеристический многочлен имеет вид

$$|\lambda\mathbf{E} - \mathbf{R}| = \begin{vmatrix} \lambda - 1 & -r_{12} & -r_{13} \\ -r_{21} & \lambda - 1 & -r_{23} \\ -r_{31} & -r_{32} & \lambda - 1 \end{vmatrix} =$$

$$= (\lambda - 1)^3 - 2r_{12}r_{13}r_{23} - (\lambda - 1)(r_{13}^2 + r_{23}^2 + r_{12}^2),$$

где **E** – единичная матрица.

Приняв $\lambda - 1 = \tau$, получим неполное кубическое уравнение

$$\tau^3 - \alpha\tau - 2\beta = 0,$$

где

$$\alpha = r_{12}^2 + r_{13}^2 + r_{23}^2, \quad \beta = r_{12} \cdot r_{13} \cdot r_{23}.$$

Решая это уравнение и учитывая выполнение неравенства $(\beta^2 - \alpha^3/27) < 0$, получим

$$\tau_1 = 2\sqrt{\frac{\alpha}{3}} \cdot \cos\frac{\varphi}{3}, \quad \tau_{2,3} = -2\sqrt{\frac{\alpha}{3}} \cdot \cos\left(\frac{\varphi}{3} \pm \frac{\pi}{3}\right),$$

где

$$\varphi = \arccos\left(\frac{\beta}{\sqrt{\frac{\alpha^3}{27}}}\right).$$

Отсюда получаем собственные значения ($\lambda_l = \tau_l + 1$), причем $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$, и матрицу собственных значений

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

Собственные значения характеризуют вклады соответствующих главных компонент в суммарную дисперсию исходных признаков $k = 3$. Таким образом первая главная компонента оказывает наибольшее влияние на общую вариацию, а третья – наименьшее. Заметим, что $\lambda_1 + \lambda_2 + \lambda_3 = k = 3$.

Вклад l-й главной компоненты в суммарную дисперсию определяется по формуле $\lambda_l/k \cdot 100\,\%$.

Найдем теперь матрицу преобразования **V** – ортогональную матрицу, составленную из собственных векторов матрицы **R**.

Собственный вектор $\mathbf{U}_j$, отвечающий собственному числу λ_j, находится как отличное от нуля решение уравнения $(\lambda_j\mathbf{E} - \mathbf{R})\mathbf{U}_j = \mathbf{0}$. Так как определитель $|\lambda_j\mathbf{E} - \mathbf{R}| = 0$, то можно считать, что третья строка есть линейная комбинация первых двух его строк. Составим два уравнения

$$(\lambda_j - 1)u_{1j} + (-r_{12})u_{2j} + (-r_{13})u_{3j} = 0,$$

$$-r_{i2}u_{1j} + (\lambda_j - 1)u_{2j} + (-r_{23})u_{3j} = 0.$$

Примем $u_{3j} = 1$ и получим решение системы двух уравнений с двумя неизвестными

$$u_{1j} = \frac{r_{13}(\lambda_j - 1) + r_{12}r_{23}}{(\lambda_j - 1)^2 - r_{12}^2},$$

$$u_{2j} = \frac{r_{23}(\lambda_j - 1) + r_{13}r_{12}}{(\lambda_j - 1)^2 - r_{12}^2}.$$

Тогда окончательно собственный вектор $\mathbf{u}_j$ имеет вид

$$\mathbf{u}_j = \begin{pmatrix} \dfrac{r_{13}(\lambda_j - 1) + r_{12}r_{23}}{(\lambda_j - 1)^2 - r_{12}^2} \\ \dfrac{r_{23}(\lambda_j - 1) + r_{13}r_{12}}{(\lambda_j - 1)^2 - r_{12}^2} \\ 1 \end{pmatrix} \qquad \text{для } j = 1, 2, 3.$$

Находим норму вектора $\mathbf{u}_j$: $|\mathbf{u}_j| = \sqrt{u_{1j}^2 + u_{2j}^2 + 1}$. Тогда матрица **V**, составленная из нормированных векторов

$$\mathbf{v}_j = \frac{\mathbf{u}_j}{|\mathbf{u}_j|} = \begin{pmatrix} v_{1j} \\ v_{2j} \\ v_{3j} \end{pmatrix},$$

имеет вид

$$\mathbf{V} = (\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3) = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}$$

и является ортогональной $\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{E}$.

Матрица факторных нагрузок получается по формуле

$$\mathbf{A} = \mathbf{V} \cdot \Lambda^{1/2} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

где $\Lambda^{1/2}$ – диагональная матрица:

$$\Lambda^{1/2} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \\ 0 & 0 & \sqrt{\lambda_3} \end{pmatrix}.$$

Таким образом, нагрузка l-й и главной компоненты f_l на j-ю переменную x_j вычисляется по формуле

$$a_{jl} = v_{jl}\sqrt{\lambda_l}\,; \quad j = 1, 2, 3; \quad l = 1, 2, 3.$$

Элемент матрицы факторных нагрузок a_{jl} есть коэффициент корреляции, который измеряет тесноту связи между l-й главной компонентой и x_j-м признаком $(-1 \leqslant a_{jl} \leqslant 1)$. При этом имеет место соотношение

$$\sum_{j=1}^{k} a_{jl}^2 = \lambda_l.$$

Матрица факторных нагрузок **A** используется для экономической интерпретации главных компонент, которые представляют собой линейные функции исходных признаков.

Значения главных компонент для каждого i-го объекта $(i = 1, 2,..., n)$ задаются матрицей **F**.

Матрицу значений главных компонент можно получить по формуле

$$\mathbf{F} = \mathbf{A}^{-1}\mathbf{Z} = \Lambda^{-1/2}\mathbf{V}^T\mathbf{Z} = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ f_{31} & f_{32} & \dots & f_{3n} \end{pmatrix},$$

где $\mathbf{Z}$ – матрица нормированных значений наблюдаемых переменных X_j размером $(3 \times n)$: $x_{j1}, x_{j2},..., x_{jn}$, $j = 1, 2, 3$; n – число наблюдений.

Таким образом значения главных компонент получаем из выражения

$$f_{li} = \frac{1}{\sqrt{\lambda_l}} \sum_{j=1}^{3} v_{jl} z_{ji},$$

где $z_{ji} = \dfrac{x_{ij} - \bar{x}_j}{s_j}$, $i = 1, 2, ..., n$; $l = 1, 2, 3$.

Полученные главные компоненты позволяют классифицировать множество исходных признаков на группы, обобщающими показателями которых и являются главные компоненты. В силу ортогональности (независимости) главные компоненты удобны для построения на них уравнения регрессии ввиду отсутствия мультиколлинеарности главных компонент. Для построения уравнения регрессии на главных компонентах в качестве исходных данных следует взять вектор наблюдаемых значений результативного признака **y** и вместо матрицы значений исходных показателей **X** – матрицу вычисленных значений главных компонент **F**.

2.2. СИСТЕМА ПРОГРАММ КОМПОНЕНТНОГО АНАЛИЗА

2.2.1. Назначение программ

Система программ включает в себя четыре программы, которые позволяют получить следующие результаты:

средние значения показателей – $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ и их средние квадратические отклонения s_1, s_2, s_3 – программа 2.2.2 (табл. 2.1 и 2.2);

матрицу парных коэффициентов корреляции **R** и матрицу собственных значений **Λ** – программа 2.2.3 (табл. 2.3 и 2.4);

матрицу нормированных собственных векторов **V** и матрицу факторных нагрузок **A** – программа 2.2.4 (табл. 2.5 и 2.6);

матрицу значений главных компонент **F** – программа 2.2.5 (табл. 2.7 и 2.8).

2.2.2. Программа вычисления средних показателей x_1, x_2, x_3 и их средних квадратических отклонений

Таблица 2.1. Инструкция по работе с программой

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память Расчеты проводить в радианах, для чего переключатель перевести в положение "Р"	По тексту программы (табл. 2.2)	Следить за кодами команд. При ошибке набрать ШГ а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
4	Занести в регистры памяти 0 и 1 число наблюдений	Набрать n x → П 0 x → П 1	
5	Занести нули в регистры памяти 2, 3, 4, 5, 6, 7, 8, 9	Cx x → П 2 x → П 3 ...... x → П 9	
6	Ввести в машину первое значение x_{11} матрицы исходных данных **X** и запустить программу с шага 00. После останова машины набрать второе наблюдение x_{12} 1-й строки матрицы **X** и запустить программу и т.д.	Набрать x_{11} В/О С/П x_{12} С/П x_{13} С/П ...... x_{i1} С/П ...... x_{ij} С/П ...... x_{n3} С/П	
7	Выписать из регистров памяти значения	П → x 2 П → x 3 П → x 4 П → x a П → x b П → x c П → x 8 П → x 9 П → x d	На табло $\bar{x}_1$ $\bar{x}_2$ $\bar{x}_3$ s_1 s_2 s_3 $\Sigma x_1 x_2$ $\Sigma x_1 x_3$ $\Sigma x_2 x_3$
8	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора программы 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 6 инструкции	F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ ... $\overrightarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл. 2.2) Набрать правильную команду

Таблица 2.2. **Текст программы вычисления средних $\bar{x}_1, \bar{x}_2, \bar{x}_3$ и средних квадратических отклонений**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	42	+	10
01	03	03	43	x → П 8	48
02	С/П	50	44	П → x a	6−
03	x → П a	4−	45	П → x c	6C
04	П → x 1	61	46	x	12
05	:	13	47	П → x 9	69
06	П → x 2	62	48	+	10
07	+	10	49	x → П 9	49
08	x → П 2	42	50	П → x b	6L
09	П → x a	6−	51	П → x c	6C
10	F x^2	22	52	x	12
11	П → x 5	65	53	П → x d	6Г
12	+	10	54	+	10
13	x → П 5	45	55	x → П d	4Г
14	С/П	50	56	F L_0	5Г
15	x → П b	4L	57	02	02
16	П → x 1	61	58	П → x 5	65
17	:	13	59	П → x 1	61
18	П → x 3	63	60	:	13
19	+	10	61	П → x 2	62
20	x → П 3	43	62	F x^2	22
21	П → x b	6L	63	−	11
22	F x^2	22	64	F $\sqrt{}$	21
23	П → x 6	66	65	x → П a	4−
24	+	10	66	П → x 6	66
25	x → П 6	46	67	П → x 1	61
26	С/П	50	68	:	13
27	x → П c	4C	69	П → x 3	63
28	П → x 1	61	70	F x^2	22
29	:	13	71	−	11
30	П → x 4	64	72	F $\sqrt{}$	21
31	+	10	73	x → П b	4L
32	x → П 4	44	74	П → x 7	67
33	П → x c	6C	75	П → x 1	61
34	F x^2	22	76	:	13
35	П → x 7	67	77	П → x 4	64
36	+	10	78	F x^2	22
37	x → П 7	47	79	−	11
38	П → x a	6−	80	F $\sqrt{}$	21
39	П → x b	6L	81	x → П c	4C
40	x	12	82	С/П	50
41	П → x 8	68	83	В/О	52

2.2.3. Программа вычисления матрицы парных коэффициентов корреляции R и матрицы собственных значений Λ

Таблица 2.3. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	БП 00 F ПРГ	
2	Записать программу в память	По тексту программы (табл. 2.4)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем — правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Ввести в машину следующие элементы*	Набрать n x → П 1 $\bar{x}_1$ x → П 2 $\bar{x}_2$ x → П 3 $\bar{x}_3$ x → П 4 s_1 x → П a s_2 x → П b s_3 x → П c Σx_1x_2 x → П 8 Σx_1x_3 x → П 9 Σx_2x_3 x → П d	
5	В регистр памяти 7 внести 3	3 x → П 7	
6	Занести в регистры памяти 0, 5, 6 нули	Cx x → П 0 x → П 5 x → П 6	
7	Запустить программу с шага 00	В/О С/П	При зацикливании нажать клавишу С/П и перейти к п. 4

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
8	Выписать из регистров памяти результаты счета	П → x 5 П → x 6 П → x 8 П → x 9 П → x d П → x 7 П → x 3 П → x 4	α β r_{12} r_{13} r_{23} $(\lambda_1 - 1)$ $(1 - \lambda_2)$ $(1 - \lambda_3)$
9	Если программа не работает:		
	1) перейти в режим программирования	F ПРГ	
	2) проверить правильность набора команды	$\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ .... $\overrightarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл. 2.4)
	3) при обнаружении несоответствия кода команды на табло тексту программы	$\overleftarrow{\text{ШГ}}$	Набрать правильную команду
	4) при ошибочном наборе номера подпрограммы, например 90, необходимо вернуться на два шага назад и набрать правильный номер подпрограммы	$\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ ПП 90	
	5) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции		

* П. 4 выполняется, если после проведения расчетов по программе 2.2.2 микрокалькулятор выключался или в регистрах памяти 1 – 4 и 8 – d не сохранены результаты расчетов по этой программе.

Таблица 2.4. **Текст программы вычисления матрицы парных коэффициентов корреляции R и матрицы собственных значений Λ**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	П → x 2	62	49	x → П 6	46
01	П → x 3	63	50	П → x 5	65
02	x	12	51	3	03
03	x → П 0	40	52	:	13
04	П → x a	6–	53	F $\sqrt{}$	21
05	П → x b	6*L*	54	x → П 1	41
06	x	12	55	2	02
07	x → П 5	45	56	x	12
08	П → x 8	68	57	x → П 0	40
09	ПП	53	58	П → x 7	67
10	90	90	59	П → x 1	61
11	x → П 8	48	60	F x^y	24
12	П → x 2	62	61	F 1/x	23
13	П → x 4	64	62	П → x 6	66
14	x	12	63	x	12
15	x → П 0	40	64	F $\cos^{-1}$	1–
16	П → x a	6–	65	3	03
17	П → x c	6*C*	66	:	13
18	x	12	67	x → П 2	42
19	x → П 5	45	68	F cos	1*Г*
20	П → x 9	69	69	П → x 0	60
21	ПП	53	70	x	12
22	90	90	71	x → П 7	47
23	x → П 9	49	72	F π	20
24	П → x 3	63	73	3	03
25	П → x 4	64	74	:	13
26	x	12	75	x → П a	4–
27	x → П 0	40	76	П → x 2	62
28	П → x b	6*L*	77	+	10
29	П → x c	6*C*	78	F cos	1*Г*
30	x	12	79	П → x 0	60
31	x → П 5	45	80	x	12
32	П → x d	6*Г*	81	x → П 3	43
33	ПП	53	82	П → x 2	62
34	90	90	83	П → x a	6–
35	x → П d	4*Г*	84	–	11
36	F x^2	22	85	F cos	1*Г*
37	П → x 8	68	86	П → x 0	60
38	F x^2	22	87	x	12
39	+	10	88	x → П 4	44
40	П → x 9	69	89	С/П	50
41	F x^2	22	90	П → x 1	61
42	+	10	91	:	13
43	x → П 5	45	92	П → x 0	60
44	П → x 8	68	93	–	11
45	П → x 9	69	94	П → x 5	65
46	x	12	95	:	13
47	П → x d	6*Г*	96	В/О	52
48	x	12			

2.2.4. Программа вычисления матрицы нормированных собственных векторов V и матрицы факторных нагрузок A

Таблица 2.5. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 2.6)	Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 1, 2, 4, 5, 6, 7, a, b, c нули	Cx x → П 1 x → П 2 x → П 4 .... x → П c	
5	Ввести в 0 регистр памяти 2	Набрать 2 x → П 0	
6	Занести в регистры памяти 8, 9, d парные коэффициенты корреляции*	r_{12} x → П 8 r_{13} x → П 9 r_{23} x → П d	
7	В регистр памяти 3 ввести очередное значение $(\lambda_l - 1)$	Набрать $(\lambda_l - 1)$ x → П 3	
8	Запустить программу с шага 00	В/О С/П	
9	После останова машины выписать из регистров памяти элементы матриц V и A	 П → x 7 П → x 5 П → x 4 П → x 1 П → x 2 П → x 6	На табло v_{1l} v_{2l} v_{3l} a_{1l} a_{2l} a_{3l}
10	Если введены не все $(\lambda_l - 1)$, то перейти к п. 4		

* П. 6 выполняется, если после проведения расчетов по программе 2.2.3 микрокалькулятор выключался или в регистрах памяти 8, 9, d не сохранены результаты расчетов по этой программе.

Таблица 2.6. **Текст программы вычисления матрицы нормированных собственных векторов V и факторных нагрузок A**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	36	+	10
01	05	05	37	F $\sqrt{}$	21
02	x → П c	4C	38	x → П 3	43
03	ПП	53	39	П → x c	6C
04	64	64	40	П → x b	6L
05	П → x 3	63	41	:	13
06	F x^2	22	42	x → П 7	47
07	П → x 8	68	43	П → x 3	63
08	F x^2	22	44	x	12
09	−	11	45	x → П 1	41
10	F 1/x	23	46	П → x a	6−
11	x → П a	4−	47	П → x b	6L
12	П → x 8	68	48	:	13
13	П → x d	6Г	49	x → П 5	45
14	x	12	50	П → x 3	63
15	x → П b	4L	51	x	12
16	П → x 3	63	52	x → П 2	42
17	П → x 9	69	53	П → x b	6L
18	x	12	54	F 1/x	23
19	П → x b	6L	55	x → П 4	44
20	+	10	56	П → x 3	63
21	П → x a	6−	57	x	12
22	x	12	58	x → П 6	46
23	F L_0	5Г	59	2	02
24	02	02	60	x → П 0	40
25	x → П a	4−	61	ПП	53
26	F x^2	22	62	64	64
27	П → x c	6C	63	С/П	50
28	F x^2	22	64	П → x 9	69
29	+	10	65	В↑	0E
30	1	01	66	П → x d	6Г
31	+	10	67	x → П 9	49
32	F $\sqrt{}$	21	68	F ○	25
33	x → П b	4L	69	x → П d	4Г
34	П → x 3	63	70	В/О	52
35	1	01			

2.2.5. Программа вычисления матрицы значений главных компонент F

Таблица 2.7. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 2.8)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем — правильную команду
3	Перевести машину в автоматический режим работы	*F* АВТ	
4	Ввести в регистры памяти 2, 3, 4 соответственно средние значения признаков x_1, x_2, x_3, а в регистры 5, 6, 7 — соответственно их среднеквадратические отклонения	Набрать $\bar{x}_1$ x → П 2 $\bar{x}_2$ x → П 3 $\bar{x}_3$ x → П 4 s_1 x → П 5 s_2 x → П 6 s_3 x → П 7	
5	Занести в регистры памяти a, b, c очередные значения *l*-го столбца матрицы **V**, а в регистр 9 – $1/\sqrt{\lambda_l}$ $(l = 1, 2, 3)$	Набрать $1/\sqrt{\lambda_l}$ x → П 9 V_{1l} x → П a V_{2l} x → П b V_{3l} x → П c	
6	В регистры памяти 0, 1, 8, d занести нули	Cx x → П 0 x → П 1 x → П 8 x → П d	
7	Набрать первое значение первой строки матрицы исходных значений **X** (x_{11}) и запустить программу с шага 00	Набрать x_{11} В/О С/П	
8	После останова машины набрать второе значение данной строки матрицы **X** (x_{12}) и запустить программу	Набрать x_{12} С/П	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
9	После запуска последнего значения (x_{13}) первой строки и останова машины: списать с табло значение первого элемента f_{l1} l-й строки матрицы факторных нагрузок; занести 0 в регистр памяти d	Набрать x_{13} С/П Сх x → П d	На табло f_{l1}
10	Набрать первое значение x_{21} второй строки матрицы **X** и запустить машину и т. д.	Набрать x_{21} С/П x_{22} С/П x_{23} С/П .... x_{i3} С/П .... x_{n1} С/П x_{n2} С/П	На табло f_{l2} ... f_{li} ... f_{ln}
11	После запуска последнего значения матрицы **X** (x_{n3}) и останова машины списать с табло последнее значение l-й строки матрицы **F** и перейти к п. 5	Набрать x_{n3} С/П	На табло f_{ln}

Таблица 2.8. **Текст программы вычисления матрицы главных компонент F**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	П → x 2	62	16	x → П d	4Г
01	−	11	17	С/П	50
02	П → x 5	65	18	П → x 4	64
03	:	13	19	−	11
04	П → x a	6−	20	П → x 7	67
05	x	12	21	:	13
06	x → П d	4Г	22	П → x c	6C
07	С/П	50	23	x	12
08	П → x 3	63	24	П → x d	6Г
09	−	11	25	+	10
10	П → x 6	66	26	П → x 9	69
11	:	13	27	x	12
12	П → x b	6L	28	С/П	50
13	x	12	29	БП	51
14	П → x d	6Г	30	00	00
15	+	10			

2.2.6. Работа с системой программ

Работа с системой программ заключается в следующем:

1. Выполнить инструкцию к программе 2.2.2 (табл. 2.1). Входными данными программы являются значения x_{ij}, содержащиеся, например, в табл. 2.9, ввод которых осуществляется по строкам x_{i1}, x_{i2}, x_{i3} ($i = \overline{1, n}$).

После окончания вычислений по программе 2.2.2 (табл. 2.2) в адресуемых регистрах памяти сохраняются значения средних величин признаков, средних квадратических отклонений и сумм произведений $\Sigma x_1 x_2$, $\Sigma x_1 x_3$, $\Sigma x_2 x_3$. Значения $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ и s_1, s_2, s_3 необходимо списать. Если между программами 2.2.2 и 2.2.3 не будет проводиться дополнительных расчетов и микрокалькулятор не выключается, то значения $\Sigma x_1 x_2$, $\Sigma x_1 x_3$, $\Sigma x_2 x_3$ выписывать не обязательно.

Переход от программы 2.2.2 к программе 2.2.3 осуществляется с помощью команд БП 00 F ПРГ, после чего производится непосредственный набор текста второй программы (табл. 2.4).

2. Выполнить инструкцию по работе с программой 2.2.3 (табл. 2.3).

Входными данными этой программы являются результаты счета, полученные в программе 2.2.2.

После окончания вычислений по программе 2.2.3 в адресуемых регистрах памяти сохраняются значения матрицы парных коэффициентов корреляции R и величин α, β, $\lambda_1 - 1$, $1 - \lambda_2$, $1 - \lambda_3$.

3. Выполнить инструкцию по работе с программой 2.2.4 (табл. 2.5).

Входными данными этой программы (табл. 2.6) являются парные коэффициенты корреляции r_{12}, r_{13}, r_{23}, а также $\lambda_l - 1 = \tau_l$.

После каждого введения значения $(\lambda_l - 1)$ осуществляется запуск программы. После ее выполнения в адресуемых регистрах памяти сохраняются соответствующие столбцы матриц нормированных собственных векторов **V** и факторных нагрузок **A**, которые необходимо выписывать после прогона каждой строки матрицы исходных данных (см. табл. 2.9).

Запуск программы 2.2.4 осуществляется три раза.

После окончания расчетов по программе получим матрицы **V** и **A**.

4. Вычислить значения $1/\sqrt{\lambda_l}$ ($l = 1, 2, 3$).

5. Выполнить инструкцию по работе с программой 2.2.5 (табл. 2.7).

Входными данными этой программы (табл. 2.8) являются средние значения $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$ и средние квадратические отклонения s_1, s_2, s_3.

Для расчета каждого элемента f_{li} матрицы **F** значений главных компонент осуществляется ввод соответствующей величины $1/\sqrt{\lambda_l}$ и столбца матрицы **V** и, кроме того, построчный ввод матрицы исходных значений **X**. Причем запуск программы производится после набора каждого аргумента x_{ij}, а значение f_{li} высвечивается на табло лишь после пропуска целой строки матрицы **X**, например первой: x_{11}, x_{12}, x_{13}. В результате после введения в машину всех значений x_{ij} матрицы **X** получаем первую строку матрицы значений главных компонент $\mathbf{F} - f_{1i}$.

Затем в соответствующие регистры памяти вводятся следующие значения $1/\sqrt{\lambda_l}$, столбцы матрицы **V**, элементы матрицы **X** и расчеты повторяются.

После окончания вычислений по программе получим матрицу факторных нагрузок **F**.

2.3. КОНТРОЛЬНЫЙ ПРИМЕР

По данным годовых отчетов n машиностроительных предприятий (табл. 2.9) провести компонентный анализ на основе показателей: x_1 – трудоемкость единицы продукции; x_2 – удельный вес рабочих в составе промышленно-производственного персонала (ППП); x_3 – удельный вес покупных изделий. Так как пример носит в основном контрольный характер, то можно ограничиться пятью наблюдениями (n = 5), между тем как для получения содержательных результатов обычно требуется, чтобы число наблюдений не менее чем в три раза превышало число признаков.

Таблица 2.9. **Исходная информация для анализа**

№ п/п	x_1	x_2	x_3
1	0,23	0,78	0,40
2	0,24	0,75	0,26
3	0,19	0,68	0,40
4	0,17	0,70	0,50
5	0,23	0,62	0,40

Решение. Определим средние значения показателей x_1, x_2, x_3 и их средние квадратические отклонения. В результате расчетов по программе 2.2.2 получим:

$\bar{x}_1 = 0{,}212$; $s_1 = 0{,}027129319$;
$\bar{x}_2 = 0{,}706$; $s_2 = 0{,}055713553$;
$\bar{x}_3 = 0{,}392$; $s_3 = 0{,}076524505$,

а также суммы произведений $\Sigma x_1 x_2 = 0{,}7502$, $\Sigma x_1 x_3 = 0{,}4074$ и $\Sigma x_2 x_3 = 1{,}377$.

Программа 2.2.3 позволяет получить матрицу парных коэффициентов корреляции **R**. Так как матрица **R** симметрическая с единицами на главной диагонали, то в регистрах памяти сохраняются значения, лежащие над главной диагональю

$$\begin{pmatrix} 1 & r_{12} & r_{13} \\ & 1 & r_{23} \\ & & 1 \end{pmatrix}.$$

Для данного примера имеем

$$\begin{pmatrix} 1 & 0{,}24347145 & -0{,}78225186 \\ & 1 & -0{,}31711391 \\ & & 1 \end{pmatrix}.$$

По программе 2.2.3 рассчитываются также значения α, β и $\tau_l = 1 - \lambda_l$: $\tau_1 = 0{,}94823232$; $-\tau_2 = 0{,}16202787$; $-\tau_3 = 0{,}78620452$. Отсюда определяем собственные значения ($\lambda_l = \tau_l + 1$): $\lambda_1 = 1{,}94823232$; $\lambda_2 = 0{,}83797213$; $\lambda_3 = 0{,}21379548$. Вклад в суммарную дисперсию первой главной компоненты равен 64,94 %, второй – 27,93 и третьей – 7,13 %. Для экономической интерпретации следует оставить две первые главные компоненты, суммарный вклад которых составляет 92,87 %.

По программе 2.2.4 рассчитываются элементы матрицы нормированных собственных векторов **V** и матрицы факторных нагрузок **A**. По данным примера имеем

$$\mathbf{V} = \begin{pmatrix} -0{,}64397431 & -0{,}32745547 & 0{,}69142617 \\ -0{,}38620753 & 0{,}91930254 & 0{,}07567394 \\ 0{,}66040968 & 0{,}21830183 & 0{,}71847297 \end{pmatrix},$$

$$\mathbf{A} = \begin{pmatrix} -0{,}89885339 & -0{,}29975554 & 0{,}31970181 \\ -0{,}53906491 & 0{,}84153705 & 0{,}03499014 \\ 0{,}92179387 & 0{,}19983528 & 0{,}33220771 \end{pmatrix}.$$

Из матрицы факторных нагрузок **A** следует, что первая главная компонента наиболее тесно связана с третьим признаком – удельный вес покупных изделий ($a_{31} = 0{,}92179382$) и первым – трудоемкость единицы продукции ($a_{11} = -0{,}89885346$), причем с первым признаком зависимость обратная. В этой связи первую главную компоненту можно интерпретировать как фактор, характеризующий уровень организации производства в отрасли.

Вторая главная компонента практически полностью определяется вторым признаком ($a_{22} = 0{,}84153705$) и поэтому может быть интерпретирована как удельный вес рабочих в составе ППП, т.е. так же, как и сам признак x_2.

Программа 2.2.5 рассчитывает элементы матрицы **F** значений главных компонент.

Для данного примера получаем

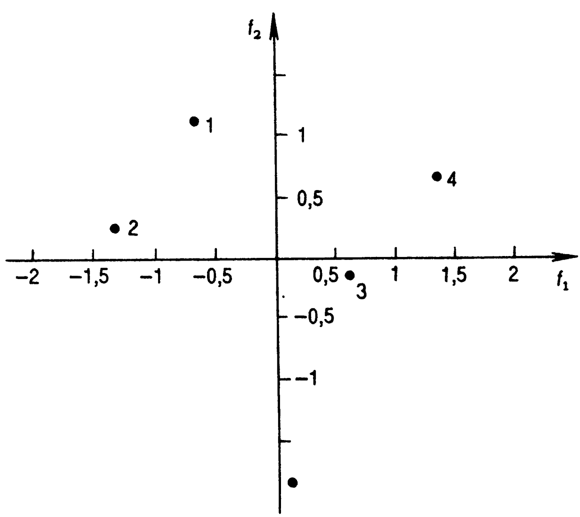
$$\mathbf{F} = \begin{pmatrix} -0{,}6241615 & -1{,}5108396 & 0{,}5527271 & 1{,}4118156 & 0{,}1704585 \\ 1{,}1214635 & 0{,}0125640 & -0{,}1536453 & 0{,}7822041 & -1{,}7625864 \\ 1{,}3719785 & -1{,}0077003 & -1{,}1265692 & -0{,}1396786 & 0{,}9019694 \end{pmatrix}.$$

Элементы этой матрицы f_{li} характеризуют значение l-й главной компоненты для i-го предприятия, $l = 1, 2, 3$, а $i = 1, 2, \ldots, 5$.

Таким образом, элементы каждого столбца матрицы **F** характеризуют соответствующее предприятие в пространстве главных компонент.

Положение машиностроительных предприятий по данным этого примера в пространстве двух первых главных компонент показано на рис. 2.1.

По рисунку видно, что пятое предприятие значительно отличается от остальных по уровню организации производства (f_1) и удельному весу рабочих в составе ППП (f_2). Остальные предприятия можно разбить на две группы (1,2 и 3,4).



Матрицу значений главных компонент **F** можно использовать в качестве исходной информации для многомерной классификации объектов с применением методов дискриминантного и кластерного анализа.

3. ДИСКРИМИНАНТНЫЙ АНАЛИЗ

3.1. АЛГОРИТМ ДИСКРИМИНАНТНОГО АНАЛИЗА

Классификацией называют разделение рассматриваемой совокупности объектов или явлений на однородные в определенном смысле группы.

Различают классификацию *при наличии обучающих выборок* (дискриминантный анализ) и классификацию *без обучения*. К классификации без обучения относят методы автоматической классификации (кластерный анализ).

Решение задач дискриминации (дискриминантный анализ) состоит в разбиении всего выборочного пространства (множества реализации всех рассматриваемых многомерных случайных величин) на некоторое число областей.

Пусть имеются две генеральные совокупности X и Y, имеющие многомерный (трехмерный) нормальный закон распределения с неизвестными, но равными ковариационными матрицами.

Из этих совокупностей взяты обучающие выборки объемами n_1 и n_2 соответственно:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \cdots & \cdots & \cdots \\ x_{n_1 1} & x_{n_1 2} & x_{n_1 3} \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ \cdots & \cdots & \cdots \\ y_{n_2 1} & y_{n_2 2} & y_{n_2 3} \end{pmatrix}.$$

Целью дискриминантного анализа в этом случае является отнесение нового наблюдения (строки) из матрицы

$$\mathbf{Z} = \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ \cdots & \cdots & \cdots \\ z_{l1} & z_{l2} & z_{l3} \end{pmatrix}$$

либо к X, либо к Y.

Для решения задачи по обучающим выборкам определяем оценки векторов средних

$$\bar{\mathbf{X}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} \text{ и } \bar{\mathbf{Y}} = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \end{pmatrix}$$

и ковариационных матриц $\mathbf{S}_X = [\mathbf{S}_{kj}]_x$ и $\mathbf{S}_Y = [S_{kj}]_y$, где

$$\bar{x}_j = \frac{1}{n_1} \sum_{i=1}^{n_1} x_{ij}, \quad \mathbf{S}_{kj(x)} = \frac{1}{n_1} \sum_{i=1}^{n_1} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) = \overline{x_j x_k} - \bar{x}_j \bar{x}_k,$$

$j, k = 1, 2, 3.$

Определим несмещенную оценку суммарной ковариационной матрицы

$$\hat{\mathbf{S}} = \frac{1}{n_1 + n_2 - 2} [n_1 \mathbf{S}_X + n_2 \mathbf{S}_Y].$$

Находим обратную матрицу $\hat{\mathbf{S}}^{-1}$ и вектор оценок коэффициентов дискриминантной функции $\mathbf{a} = \hat{\mathbf{S}}^{-1}(\bar{\mathbf{X}} - \bar{\mathbf{Y}})$.

Вычислим оценки дискриминантной функции для матриц исходных данных $\mathbf{X}$ и $\mathbf{Y}$:

$$\hat{\mathrm{U}}_X = \mathbf{X} \cdot \mathrm{a}, \quad \hat{\mathrm{U}}_Y = \mathbf{Y} \cdot \mathrm{a}.$$

Тогда их средние значения будут определяться как

$$\bar{\hat{u}}_X = \frac{1}{n_1} \sum_{i=1}^{n_1} \hat{u}_{X_i}, \quad \bar{\hat{u}}_Y = \frac{1}{n_2} \sum_{i=1}^{n_2} \hat{u}_{Y_i}.$$

Определим границу дискриминации – константу $\hat{c}$ по формуле

$$\hat{c} = \frac{1}{2} (\bar{\hat{u}}_X + \bar{\hat{u}}_Y).$$

Оценку дискриминантной функции для ν-й строки матрицы $\mathbf{Z}$, которая характеризует ν-е наблюдение, подлежащее дискриминации, получим, решив уравнение

$$\hat{u}_\nu = z_{\nu 1} a_1 + z_{\nu 2} a_2 + z_{\nu 3} a_3.$$

Если $\hat{u}_\nu \geqslant \hat{c}$, то ν-е наблюдение следует отнести к совокупности $\mathbf{X}$, если же $\hat{u}_\nu < \hat{c}$, то ν-е наблюдение будем относить к совокупности $\mathbf{Y}$.

Дискриминантный анализ допускает наличие более двух обучающих выборок, однако в этом случае задача существенно усложняется и не всегда приводит к однозначной дискриминации, т.е. не все объекты удается отнести к какому-либо классу.

3.2.1. Назначение программ

Система программ включает в себя три программы, которые позволяют получить следующие результаты:

векторы средних $\bar{X}$ и $\bar{Y}$ и оценки ковариационных матриц S_x и S_y – программа 3.2.2 (табл. 3.1 и 3.2);

несмещенную оценку суммарной ковариационной матрицы $\hat{S}$ и обратной матрицы $\hat{S}^{-1}$, вектор оценок коэффициентов дискриминантной функции **a**, оценки дискриминантной функции для матриц исходных данных **X** и **Y** – $\hat{u}_x$ и $\hat{u}_y$ – программа 3.2.3 (табл. 3.3 и 3.4);

средние значения $\bar{u}_x$ и $\bar{u}_y$ и константу $\hat{c}$ – программа 3.2.4 (табл. 3.5 и 3.6).

3.2.2. Программа вычисления векторов средних $\bar{X}$, $\bar{Y}$ и оценок ковариационных матриц S_x, S_y

Таблица 3.1. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 3.2)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести количество строк матрицы **X** n_1 в регистры памяти 0 и d, а во все остальные регистры – нули При работе с матрицей **Y** в регистры 0 и d заносится n_2	Набрать n_1 x → П 0 x → П d Cx x → П 1 x → П 2 .... x → П 9 x → П a ...... x → П c	
5	Набрать элемент x_{i1} матрицы **X** и запустить программу с шага 00	Набрать x_{i1} В/О С/П	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
6	После останова машины набрать следующий элемент матрицы $\mathbf{X} - x_{i2}$ и запустить программу	Набрать x_{i2} С/П	
7	Набрать последний элемент i-й строки матрицы $\mathbf{X} - x_{i3}$ и запустить программу	Набрать x_{i3} С/П	
8	Если введены не все строки матрицы **X**, то перейти к п. 5, иначе – к п. 9.		
9	После останова машины списать с табло значение $n_1 s_{23}$, а из регистров памяти 1, 2, 3, 7, 8 – элементы матрицы $n_1 s_x$; из регистров памяти 4, 5, 6 – элементы вектора $n_1 \overline{x}$	 П → x 1 П → x 2 П → x 3 П → x 7 П → x 8 П → x 4 П → x 5 П → x 6	На табло $n_1 s_{23}$ $n_1 s_{11}$ $n_1 s_{12}$ $n_1 s_{13}$ $n_1 s_{22}$ $n_1 s_{33}$ $n_1 \overline{x}_1$ $n_1 \overline{x}_2$ $n_1 \overline{x}_3$
10	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора команд 3) при обнаружении несоответствия кода команды на табло тексту программы 4) при ошибочном наборе номера подпрограммы, например 90, необходимо вернуться на два шага назад и набрать правильный номер подпрограммы 5) при ошибочном наборе адреса команды перехода, например 97, необходимо вернуться на два шага назад и набрать правильный адрес	 F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ .... $\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ ПП 90 $\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ F L_0 97	 Сличать высвечиваемые коды с текстом программы (табл. 3.2) Набрать правильную команду

Примечания: 1. Расчеты по матрице **Y** в п. 5 – 9 выполняются аналогично указанному для матрицы **X**.

2. В п. 9 для получения также матрицы $\mathbf{S}_x$ и вектора $\overline{\mathbf{X}}$ необходимо после выполнения п. 9 при выписывании из регистров памяти элементов матрицы $n_1 s_x$ и вектора $n_1 \overline{\mathbf{X}}$ делить их на n_1, например:

Нажимаемые клавиши	На табло
П → x 1	$n_1 s_{11}$
П → x d	n_1
:	s_{11}

Таблица 3.2. **Текст программы вычисления векторов средних $\overline{\mathbf{X}}$ и $\overline{\mathbf{Y}}$ и оценок ковариационных матриц $\mathbf{S}_x$ и $\mathbf{S}_y$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x → П a	4–	49	F L_0	5Г
01	С/П	50	50	97	97
02	x → П b	4L	51	П → x 1	61
03	С/П	50	52	x → П a	4–
04	x → П c	4C	53	П → x 4	64
05	F x^2	22	54	П → x 4	64
06	П → x 8	68	55	ПП	53
07	+	10	56	90	90
08	x → П 8	48	57	x → П 1	41
09	П → x a	6–	58	П → x 7	67
10	П → x b	6L	59	x → П a	4–
11	x	12	60	П → x 5	65
12	П → x 2	62	61	П → x 5	65
13	+	10	62	ПП	53
14	x → П 2	42	63	90	90
15	П → x a	6–	64	x → П 7	47
16	П → x c	6C	65	П → x 8	68
17	x	12	66	x → П a	4–
18	П → x 3	63	67	П → x 6	66
19	+	10	68	П → x 6	66
20	x → П 3	43	69	ПП	53
21	П → x a	6–	70	90	90
22	П → x 4	64	71	x → П 8	48
23	+	10	72	П → x 2	62
24	x → П 4	44	73	x → П a	4–
25	П → x b	6L	74	П → x 4	64
26	П → x 5	65	75	П → x 5	65
27	+	10	76	ПП	53
28	x → П 5	45	77	90	90
29	П → x c	6C	78	x → П 2	42
30	П → x 6	66	79	П → x 3	63
31	+	10	80	x → П a	4–
32	x → П 6	46	81	П → x 4	64
33	П → x b	6L	82	П → x 6	66
34	F x^2	22	83	ПП	53
35	П → x 7	67	84	90	90
36	+	10	85	x → П 3	43
37	x → П 7	47	86	П → x 9	69
38	П → x a	6–	87	x → П a	4–
39	F x^2	22	88	П → x 5	65
40	П → x 1	61	89	П → x 6	66
41	+	10	90	x	12
42	x → П 1	41	91	П → x d	6Г
43	П → x b	6L	92	:	13
44	П → x c	6C	93	П → x a	6–
45	x	12	94	↔	14
46	П → x 9	69	95	–	11
47	+	10	96	В/О	52
48	x → П 9	49	97	С/П	50

3.2.3. Программа вычисления матриц $\hat{S}$ и $\hat{S}^{-1}$, вектора $\mathbf{a}$ и оценок значений дискриминантной функции $\hat{u}_X$ и $\hat{u}_Y$

Таблица 3.3. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 3.4)	Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистр памяти 0 число 6	6 x → П 0	
5	Перевести программу на шаг 00	В/О	
6	Занести в регистры памяти: c – число $(n_1 + n_2 - 2)$, b – $n_1 s_{11(x)}$, d – $n_2 s_{11(y)}$ и запустить программу	Набрать $(n_1 + n_2 - 2)$ x → П c $n_1 s_{11(x)}$ x → П b $n_2 s_{11(y)}$ x → П d С/П	
7	После останова машины списать с табло число $\hat{s}_{11}$ и занести его в регистр памяти 7	x → П 7	На табло $\hat{s}_{11}$
8	Занести в регистры памяти b и d соответственно $n_1 s_{12(x)}$ и $n_2 s_{12(y)}$ и запустить программу	Набрать $n_1 s_{12(x)}$ x → П b $n_2 s_{12(y)}$ x → П d С/П	
9	После останова машины списать с табло число $\hat{s}_{12}$ и занести его в регистры памяти 4 и 8	x → П 4 x → П 8	На табло $\hat{s}_{12}$
10	Занести в регистры памяти b и d $n_1 s_{13(x)}$ и $n_2 s_{13(y)}$ и запустить программу	Набрать $n_1 s_{13(x)}$ x → П b $n_2 s_{13(y)}$ x → П d С/П	
11	После останова машины списать с табло число $\hat{s}_{13}$ и запомнить его в регистрах 1 и 9	x → П 1 x → П 9	На табло $\hat{s}_{13}$
12	Занести в регистры памяти b и d $n_1 s_{22(x)}$ и $n_2 s_{22(y)}$ и запустить программу	Набрать $n_1 s_{22(x)}$ x → П b $n_2 s_{22(y)}$ x → П d С/П	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
13	После останова машины списать с табло $\hat{s}_{22}$ и занести его в регистр памяти 5	x → П 5	На табло $\hat{s}_{22}$
14	Занести в регистры памяти b и d $n_1 s_{23(x)}$ и $n_2 s_{23(y)}$ и запустить программу	Набрать $n_1 s_{23(x)}$ x → П b $n_2 s_{23(y)}$ x → П d С/П	
15	После останова машины списать с табло $\hat{s}_{23}$ и занести его в регистры памяти 2 и 6	x → П 2 x → П 6	На табло $\hat{s}_{23}$
16	Занести в регистры памяти b и d $n_1 s_{33(x)}$ и $n_2 s_{33(y)}$ и запустить программу	Набрать $n_1 s_{33(x)}$ x → П b $n_2 s_{33(y)}$ x → П d С/П	
17	После останова машины списать с табло $\hat{s}_{33}$ и снова запустить программу	С/П	На табло $\hat{s}_{33}$
18	После останова машины выписать из регистров памяти элементы обратной матрицы $\hat{\mathbf{S}}^{-1}$:	П → x 7 П → x 8 П → x 9 П → x 5 П → x 6 П → x 3	$\hat{s}^{-1}_{11}$ $\hat{s}^{-1}_{12}$ $\hat{s}^{-1}_{13}$ $\hat{s}^{-1}_{22}$ $\hat{s}^{-1}_{23}$ $\hat{s}^{-1}_{33}$
19	Занести в регистры памяти 0, a, b соответственно $\bar{x}_1 - \bar{y}_1$, $\bar{x}_2 - \bar{y}_2$, $\bar{x}_3 - \bar{y}_3$ элементы вектора $(\bar{\mathbf{X}} - \bar{\mathbf{Y}})^3$ и запустить программу	Набрать $(\bar{x}_1 - \bar{y}_1)$ x → П 0 $(\bar{x}_2 - \bar{y}_2)$ x → П a $(\bar{x}_3 - \bar{y}_3)$ x → П b С/П	
20	После останова машины списать с табло элемент a_1. Занести в регистры памяти 7, 8, 9 соответственно вторую строку матрицы $\hat{\mathbf{S}}^{-1}$ и запустить программу. Учитывается, что $\hat{s}^{-1}_{21} = \hat{s}^{-1}_{12}$	Набрать $\hat{s}^{-1}_{12}$ x → П 7 $\hat{s}^{-1}_{22}$ x → П 8 $\hat{s}^{-1}_{23}$ x → П 9 С/П	На табло a_1

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
21	Списать с табло элемент a_2. Занести в регистры памяти 7, 8, 9 элементы третьей строки матрицы $\hat{\mathbf{S}}^{-1}$ и запустить программу ($\hat{s}^{-1}_{31} = \hat{s}^{-1}_{13}$; $\hat{s}^{-1}_{32} = \hat{s}^{-1}_{23}$)	Набрать $\hat{s}^{-1}_{13}$ x → П 7 $\hat{s}^{-1}_{23}$ x → П 8 $\hat{s}^{-1}_{33}$ x → П 9 С/П	На табло a_2
22	Списать с табло элемент a_3. Занести в регистры памяти 0, a, b элементы вектора **a**	Набрать a_1 x → П 0 a_2 x → П a a_3 x → П b	На табло a_3
23	Занести в регистры памяти 7, 8, 9 элементы i-й строки матрицы **X** и запустить программу	Набрать x_{i1} x → П 7 x_{i2} x → П 8 x_{i3} x → П 9 С/П	
24	Списать с табло i-й элемент вектора $\hat{\mathbf{u}}_X$		На табло $\hat{u}_{Xi}$
25	Если введены не все строки матрицы **X**, то перейти к п. 23, иначе – к п. 26		
26	Занести в регистры памяти 7, 8, 9 элементы i-й строки матрицы **Y** и запустить программу	Набрать y_{i1} x → П 7 y_{i2} x → П 8 y_{i3} x → П 9 С/П	
27	Списать с табло i-й элемент вектора $\hat{\mathbf{u}}_Y$		На табло $\hat{u}_{Yi}$
28	Если введены не все строки матрицы **Y**, то перейти к п. 26		
29	При неправильном наборе адреса безусловного перехода необходимо вернуться на два шага назад и набрать правильный адрес	$\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ БП 03	

Таблица 3.4. **Текст программы вычисления матриц $\hat{S}$ и $\hat{S}^{-1}$, вектора $\hat{a}$ и оценок дискриминантной функции $\hat{u}_X$ и $\hat{u}_Y$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	Б П	51	45	x → П 4	44
01	03	03	46	П → x 3	63
02	С/П	50	47	П → x d	6Г
03	П → x b	6L	48	П → x 9	69
04	П → x d	6Г	49	x	12
05	+	10	50	−	11
06	П → x c	6C	51	x → П 5	45
07	:	13	52	П → x 8	68
08	F L_0	5Г	53	П → x 7	67
09	02	02	54	:	13
10	x → П 3	43	55	x → П 1	41
11	С/П	50	56	П → x 9	69
12	3	03	57	П → x 7	67
13	x → П 0	40	58	:	13
14	П → x 4	64	59	x → П 2	42
15	П → x 7	67	60	П → x 7	67
16	:	13	61	F 1/x	23
17	x → П d	4Г	62	x → П 3	43
18	П → x d	6Г	63	П → x a	6−
19	/−/	0L	64	x → П 7	47
20	x → П c	4C	65	П → x b	6L
21	П → x 5	65	66	x → П 8	48
22	П → x d	6Г	67	П → x c	6C
23	П → x 8	68	68	x → П 9	49
24	x	12	69	F L_0	5Г
25	−	11	70	14	14
26	x → П a	4−	71	С/П	50
27	П → x 6	66	72	П → x 7	67
28	П → x d	6Г	73	П → x 0	60
29	П → x 9	69	74	x	12
30	x	12	75	x → П 7	47
31	−	11	76	П → x 8	68
32	x → П b	4L	77	П → x a	6−
33	П → x 1	61	78	x	12
34	П → x 7	67	79	П → x 7	67
35	:	13	80	+	10
36	x → П d	4Г	81	x → П 7	47
37	П → x d	6Г	82	П → x 9	69
38	/−/	0L	83	П → x b	6L
39	x → П 6	46	84	x	12
40	П → x 2	62	85	П → x 7	67
41	П → x d	6Г	86	+	10
42	П → x 8	68	87	x → П 7	47
43	x	12	88	БП	51
44	−	11	89	71	71

3.2.4. Программа вычисления средних значений $\bar{\hat{u}}_X$ и $\bar{\hat{u}}_Y$ и константы $\hat{c}$

Таблица 3.5. Инструкция по работе с программой

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 3.6)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ В/О	
4	Занести в регистры памяти 0 и а значение n_1, а в регистры 1 и b – n_2. Во все остальные регистры памяти занести нули	Набрать n_1 x → П 0 x → П a n_2 x → П 1 x → П b Cx x → П 2 x → П 3 .... x → П 9 x → П c x → П d	
5	Набрать $\hat{u}_{Xi}$ и запустить программу	Набрать $\hat{u}_{Xi}$ С/П	
6	Если введены не все элементы вектора $\hat{u}_X$, то перейти к п.5, иначе – к п. 7		
7	Набрать i-й элемент вектора $\hat{u}_Y$ и запустить программу	Набрать $\hat{u}_{Yi}$ С/П	
8	Если введены не все элементы вектора $\hat{u}_Y$, то перейти к п. 7, иначе – к п. 9		
9	После останова машины списать с табло значение константы $\hat{c}$, а из регистров памяти 4 и 3 соответственно средние значения $\bar{\hat{u}}_X$ и $\bar{\hat{u}}_Y$	П → x 4 П → x 3	На табло $\hat{c}$ $\bar{\hat{u}}_X$ $\bar{\hat{u}}_Y$

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
10	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора команд 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции	F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ ... $\overleftarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл. 3.6) Набрать правильную команду

Таблица 3.6. Текст программы вычисления средних значений $\hat{\bar{u}}_X$ и $\hat{\bar{u}}_Y$ и константы $\hat{c}$

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	14	П → x 4	64
01	03	03	15	П → x a	6–
02	С/П	50	16	:	13
03	П → x 4	64	17	x → П 4	44
04	+	10	18	П → x 2	62
05	x → П 4	44	19	П → x b	6*L*
06	F L_0	5*Г*	20	:	13
07	02	02	21	x → П 3	43
08	С/П	50	22	П → x 4	64
09	П → x 2	62	23	+	10
10	+	10	24	2	02
11	x → П 2	42	25	:	13
12	F L_1	5*L*	26	С/П	50
13	08	08			

3.2.5. Работа с системой программ

Работа с системой программ заключается в следующем:

1. Выполнить инструкцию к программе 3.2.2 (табл. 3.1) сначала для матрицы исходных данных **X**, а затем – для матрицы **Y**, начиная с п.4 инструкции. Входными данными программы являются значения $x_{ij}(i = \overline{1, n_1})$ и $y_{ij}(i = \overline{1, n_2})$, ввод которых осуществляется по строкам.

После окончания вычислений по программе 3.2.2 (табл. 3.2) для исходных данных матрицы $\mathbf{X}$ в адресуемых регистрах памяти сохраняются значения вектора $n_1\bar{\mathbf{X}}$ и матрицы $n_1\mathbf{S}_x$, которые необходимо списать. Затем по программе рассчитываются вектор $n_2\bar{Y}$ и матрица $n_2\mathbf{S}_y$ для исходных данных $\mathbf{Y}$, которые также необходимо выписать из адресуемых регистров памяти.

Определение векторов $n_1\bar{\mathbf{X}}$ и $n_2\bar{Y}$ и матриц $n_1\mathbf{S}_x$ и $n_2\mathbf{S}_y$ более удобно для проведения дальнейших расчетов по алгоритму дискриминантного анализа. Чтобы определить векторы средних $\bar{\mathbf{X}}$ и $\bar{\mathbf{Y}}$ и матрицы $\mathbf{S}_x$ и $\mathbf{S}_y$, необходимо выполнить примечание к п. 9 инструкции по работе с этой программой.

2. Вычислить вектор $(\bar{\mathbf{X}} - \bar{\mathbf{Y}})$.

3. Выполнить инструкцию по работе с программой 3.2.3 (табл. 3.3).

Вычисления по программе (табл. 3.4) проводятся в несколько этапов. На первом этапе (п. 1 – 18 инструкции) определяются оценка суммарной ковариационной матрицы $\hat{\mathbf{S}}$ и обратная матрица $\hat{\mathbf{S}}^{-1}$, входными данными являются результаты счета по этой программе. После вычисления каждого элемента матрицы $\hat{\mathbf{S}}$ необходимо списывать его с табло микрокалькулятора. Элементы матрицы $\hat{\mathbf{S}}^{-1}$ сохраняются в адресуемых регистрах памяти и выписываются из них после окончания первого этапа расчетов по программе.

На втором этапе (п. 19 – 22 табл. 3.3) вычисляется вектор оценок ($\mathbf{a}$) коэффициентов дискриминантной функции. Исходными данными служат элементы обратной матрицы $\hat{\mathbf{S}}^{-1}$, полученной на первом этапе, и вектора $(\bar{\mathbf{X}} - \bar{Y})$. После вычисления каждого элемента вектора $\mathbf{a}$ его необходимо списать с табло машины.

На третьем этапе получаем оценки дискриминантной функции для матриц исходных данных $\mathbf{X}$ (п. 22 – 25 инструкции) и $\mathbf{Y}$ (п. 26 – 28), которые являются входными данными для данного этапа программы. Все элементы векторов $\hat{\mathbf{u}}_X$ и $\hat{\mathbf{u}}_Y$ необходимо выписывать по ходу выполнения расчетов.

4. Выполнить инструкцию по работе с программой 3.2.4 (табл. 3.5).

Перевод машины в автоматический режим работы осуществляется с помощью команд F АВТ В/О.

Входными данными этой программы (табл. 3.6) являются векторы $\hat{\mathbf{u}}_X$ и $\hat{\mathbf{u}}_Y$. Запуск программы осуществляется после ввода каждого элемента сначала вектора $\hat{\mathbf{u}}_X$, а затем вектора $\hat{\mathbf{u}}_Y$.

После окончания расчетов по этой программе получим средние значения $\bar{\hat{u}}_X$ и $\bar{\hat{u}}_Y$ и константу $\hat{c}$.

3.3. КОНТРОЛЬНЫЙ ПРИМЕР

На основании данных годовых отчетов деятельность производственных объединений электротехнической промышленности оценивается по трем основным показателям: среднегодовой стоимости основных производственных фондов (млн.руб.), среднесписочной численности промышленно-производственного персонала (тыс. чел.) и балансовой прибыли (млн. руб.). Требуется выделить группу передовых объединений отрасли.

С этой целью экспертами были определены обучающие выборки. Первая выборка ($\mathbf{X}$, $n_1 = 4$) представляла группу передовых объединений, а вторая ($\mathbf{Y}$, $n_2 = 5$) – остальных (табл. 3.7).

Таблица 3.7. **Таблица исходных данных**

Группы объединений	Стоимость основных фондов	Численность ППП	Прибыль
Передовые (X)	224,228 151,827 147,313 152,253	17,115 14,904 13,627 10,545	22,981 21,481 28,669 10,199
Остальные (Y)	46,757 29,033 52,134 37,050 63,979	4,428 5,510 4,214 5,527 4,211	11,124 6,091 11,842 11,873 12,860

Можно ли к группе передовых отнести объединение z, имеющее следующие характеристики: стоимость основных фондов – 55,451 млн. руб., численность ППП – 9,592 тыс. чел., прибыль – 12,840 млн. руб.?

Решение. Определим векторы средних $\bar{\mathbf{X}}$ и $\bar{\mathbf{Y}}$ и оценки ковариационных матриц $\mathbf{S}_x$ и $\mathbf{S}_y$ соответственно для групп передовых и остальных объединений. В результате расчетов по программе 3.2.2 получим:

$$n_1\hat{X}=\begin{pmatrix}675,681\\ 56,191\\ 83,330\end{pmatrix},\quad n_1S_x=\begin{pmatrix}4102,44 & 222,663 & 115,779\\ & 22,58745 & 41,0946\\ & & 179,5187\end{pmatrix},$$

$$n_2\hat{Y}=\begin{pmatrix}228,963\\ 23,890\\ 53,790\end{pmatrix},\quad n_2S_y=\begin{pmatrix}729,333 & -33,0476 & 113,9347\\ & 1,85891 & -4,51242\\ & & 28,75151\end{pmatrix}.$$

Выполнив примечание к пункту 9 табл. 3.1, имеем

$$\bar{\mathbf{X}}=\begin{pmatrix}168,92025\\ 14,04775\\ 20,8325\end{pmatrix},\quad \mathbf{S}_x=\begin{pmatrix}1025,61 & 55,66575 & 28,94475\\ & 5,6468625 & 10,27365\\ & & 44,879675\end{pmatrix},$$

$$\bar{\mathbf{Y}} = \begin{pmatrix} 45,7926 \\ 4,778 \\ 10,758 \end{pmatrix}, \quad \mathbf{S}_y = \begin{pmatrix} 145,8666 & -6,60952 & 22,78694 \\ & 0,371782 & -0,902484 \\ & & 5,750302 \end{pmatrix}.$$

Вычисляем вектор $(\bar{\mathbf{X}} - \bar{\mathbf{Y}})$:

$$(\bar{\mathbf{X}} - \bar{\mathbf{Y}}) = \begin{pmatrix} 123,12765 \\ 9,26975 \\ 10,0745 \end{pmatrix}.$$

Программа 3.2.3 позволяет получить следующие результаты: несмещенную оценку суммарной ковариационной матрицы

$$\hat{\mathbf{S}} = \begin{pmatrix} 690,25328 & 27,087914 & 32,816242 \\ & 3,4923371 & 5,2260257 \\ & & 29,752887 \end{pmatrix};$$

обратную матрицу $\hat{\mathbf{S}}^{-1}$ и вектор оценок коэффициентов дискриминантной функции **a**:

$$\hat{\mathbf{S}}^{-1} = \begin{pmatrix} 0,0020945371 & -0,017349116 & 0,00073714 \\ & 0,53214363 & -0,07433441 \\ & & 0,04585381 \end{pmatrix}; \quad \mathbf{a} = \begin{pmatrix} 0,10449979 \\ 2,0478006 \\ -0,13634981 \end{pmatrix}.$$

Матрицы $n_1\mathbf{S}_x$, $n_2\mathbf{S}_y$, $\mathbf{S}_x$, $\mathbf{S}_y$, $\hat{\mathbf{S}}$ и $\hat{\mathbf{S}}^{-1}$ симметрические и поэтому выписываются только элементы, лежащие на главной диагонали и над ней;

оценки дискриминантной функции для матриц **X** и **Y**:

$$\hat{\mathbf{U}}_X = \begin{pmatrix} 55,346433 \\ 43,457381 \\ 39,390544 \\ 36,113833 \end{pmatrix}, \quad \hat{\mathbf{U}}_Y = \begin{pmatrix} 12,437003 \\ 13,486817 \\ 12,46277 \\ 13,571031 \\ 13,555623 \end{pmatrix}.$$

По программе 3.2.4 рассчитываем средние значения $\bar{\hat{u}}_X = 43,577047$ и $\bar{\hat{u}}_Y = 13,102648$, по которым определяется константа $\hat{c} = 28,339847$.

Чтобы определить, к какой группе относится указанное выше объединение, необходимо значения его показателей $z_1 = 55,451$, $z_2 = 9,592$ и $z_3 = 12,84$ подставить в уравнение

$$u_z = a_1 \cdot z_1 + a_2 \cdot z_2 + a_3 \cdot z_3.$$

Отсюда находим $\hat{u}_z = 23,68639$. Сравнивая $\hat{u}_z$ и $\hat{c}$, получаем, что $\hat{u}_z < \hat{c}$ $(23,7 < 28,3)$, поэтому данное объединение нельзя отнести к группе передовых.

Расчет $\hat{u}_z$ в автоматическом режиме:

a_1 В↑ z_1 × a_2 В↑ z_2 × a_3 В↑ z_3 × + +

на табло $\hat{u}_z = 23,68639$.

В данном примере $a_1 = 0,10449979$, $a_2 = 2,048006$, $a_3 = -0,13634981$.

4. КЛАСТЕРНЫЙ АНАЛИЗ

4.1. АЛГОРИТМ КЛАСТЕРНОГО АНАЛИЗА

Кластерный анализ – это совокупность методов классификации многомерных наблюдений или объектов, основанных на определении понятия расстояния между объектами с последующим выделением из них групп, "сгустков" наблюдений (кластеров, таксонов). При этом не требуется априорной информации о распределении генеральной совокупности.

Выбор конкретного метода кластерного анализа зависит от цели классификации.

Кластерный анализ используется при исследовании структуры совокупностей социально-экономических показателей или объектов: предприятий, регионов, социологических анкет, коллективов и т.д.

От матрицы исходных данных

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ \cdots & \cdots & \cdots & \cdots \\ x_{i1} & x_{i2} & x_{i3} & x_{i4} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & x_{n3} & x_{n4} \end{pmatrix}$$

переходим к матрице нормированных значений $\mathbf{Z}$ с элементами

$$z_{ij} = \frac{x_{ij} - \overline{x}_j}{s_j},$$

где $j = 1, 2, 3, 4$ – номер показателя, $i = 1, 2, \ldots, n$ – номер наблюдения;

$$\overline{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}; \quad s_j = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \overline{x}_j)^2} = \sqrt{\frac{1}{n} \sum x_{ij}^2 - (\overline{x}_j)^2}.$$

В качестве расстояния между двумя наблюдениями z_i и z_ν используют "взвешенное" евклидово расстояние, определяемое по формуле

$\rho_{BE}(z_i, z_\nu) = \sqrt{\sum_{l=1}^{4} w_l (z_{il} - z_{\nu l})^2}$, где w_l – "вес" показателя; $0 < w_l \leq 1$.

Если $w_l = 1$ для всех $l = 1, 2, 3, 4$, то получаем обычное евклидово расстояние

$$\rho_E(z_i, z_\nu) = \sqrt{\sum_{l=1}^{4} (z_{il} - z_{\nu l})^2}.$$

Полученные значения удобно представить в виде матрицы расстояний

$$\mathbf{R} = \begin{pmatrix} 0 & \rho_{12} & \rho_{13} & \dots & \rho_{1n} \\ \rho_{21} & 0 & \rho_{23} & \dots & \rho_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{i1} & \rho_{i2} & 0 & \dots & \rho_{in} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{n1} & \rho_{n2} & \rho_{n3} & \dots & 0 \end{pmatrix}; \quad \rho_{i\nu} = \rho_{\nu i}.$$

Так как матрица **R** симметрическая, т.е. $\rho_{i\nu} = \rho_{\nu i}$, то достаточно ограничиться записью наддиагональных элементов матрицы.

Используя матрицу расстояний, можно реализовать агломеративную иерархическую процедуру кластерного анализа. Расстояния между кластерами определяют по *принципу "ближайшего соседа"* или *"дальнего соседа"*. В первом случае за расстояние между кластерами принимают расстояние между ближайшими элементами этих кластеров, а во втором – между наиболее удаленными друг от друга.

Принцип работы иерархических агломеративных процедур состоит в последовательном объединении групп элементов сначала самых близких, а затем все более отдаленных друг от друга.

На первом шаге алгоритма каждое наблюдение $z_i (i = 1, 2, \dots, n)$ рассматривается как отдельный кластер. В дальнейшем на каждом шаге работы алгоритма происходит объединение двух самых близких кластеров, и вновь строится матрица расстояний, размерность которой снижается на единицу. Работа алгоритма заканчивается, когда все наблюдения объединены в один класс.

4.2. СИСТЕМА ПРОГРАММ КЛАСТЕРНОГО АНАЛИЗА

Система включает две программы. Программа 4.2.1 (табл. 4.1 и 4.2) предназначена для получения средних значений, средних квадратических отклонений и матрицы, состоящей из нормированных значений исходных показателей. Программа 4.2.2 (табл. 4.3 и 4.4) позволяет рассчитать матрицу расстояний между всеми парами наблюдений.

4.2.1. Программа преобразования исходной матрицы наблюдений X в нормированную матрицу Z

Таблица 4.1. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 4.2)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 0 и a число строк матрицы наблюдений X – n	Набрать n x → П 0 x → П a	
5	Занести в регистры памяти 1, 2,..., 9, b, c, d нули	Сх x → П 1 x → П 2 .... x → П 9 x → П b x → П c x → П d	
6	Ввести в машину первое значение j-го столбца матрицы X и запустить программу с шага 00. После останова машины набрать второе значение j-го столбца матрицы X и запустить программу и т. д.	Набрать x_{1j} В/О С/П x_{2j} С/П .... x_{ij} С/П .... x_{nj} С/П	
7	Если введены все значения j-го столбца матрицы X, то повторно набрать первое наблюдение j-го столбца матрицы X и запустить программу	Набрать x_{1j} С/П	
8	После останова машины списать с табло первый элемент нормированной матрицы Z, а из регистров памяти 2 и 4 – соответственно s_1 и $\bar{x}_1$	П → x 2 П → x 4	На табло z_{1j} s_1 $\bar{x}_1$

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
9	Ввести в машину очередное значение j-го столбца матрицы **X** и запустить программу	Набрать x_{ij} С/П	
10	После останова машины списать с табло элемент z_{ij} нормированной матрицы **Z**		На табло z_{ij}
11	Если введены не все значения j-го столбца матрицы наблюдений **X**, то перейти к п. 9, иначе – к п. 5		
12	При ошибочном наборе адреса безусловного перехода, например 03, необходимо вернуться на два шага назад и набрать правильный адрес	$\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ БП 03	

Таблица 4.2. **Текст программы вычисления нормированной матрицы Z**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	17	02	02
01	03	03	18	П → x a	6–
02	С/П	50	19	x → П 0	40
03	x → П 2	42	20	П → x 3	63
04	F x^2	22	21	П → x 4	64
05	П → x a	6–	22	F x^2	22
06	:	13	23	–	11
07	П → x 3	63	24	F $\sqrt{\ }$	21
08	+	10	25	x → П 2	42
09	x → П 3	43	26	С/П	50
10	П → x 2	62	27	П → x 4	64
11	П → x a	6–	28	–	11
12	:	13	29	П → x 2	62
13	П → x 4	64	30	:	13
14	+	10	31	БП	51
15	x → П 4	44	32	26	26
16	F L_0	5Г			

4.2.2. Программа вычисления евклидова расстояния между объектами

Таблица 4.3. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 4.4)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 0, a, b, c "весовые" коэффициенты w_l*	Набрать w_1 x → П 0 w_2 x → П a w_3 x → П b w_4 x → П c	
5	Занести в регистры памяти 1, 2, 3, 4 значения $\bar{i}$-й строки матрицы исходных нормированных данных **Z**	Набрать z_{i1} x → П 1 z_{i2} x → П 2 z_{i3} x → П 3 z_{i4} x → П 4	
6.	Занести в регистры памяти 5, 6,..., 9, d нули	Cx x → П 5 x → П 6 .... x → П 9 x → П d	
7	Ввести в стековую память k-ю строку матрицы **Z** и запустить программу с шага 00, где $k > i$, $k \leqslant n$	Набрать z_{k1} В↑ z_{k2} В↑ z_{k3} В↑ z_{k4} В/О С/П	
8	После останова машины списать с табло элемент ρ_{ik} матрицы расстояний **R**		На табло ρ_{ik}.
9	Если рассчитаны не все наблюдения i-й строки матрицы расстояний **R**, то перейти к п. 6, иначе – к п. 5		

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
10	Если программа не работает: 1) перейти в режим программирования	F ПРГ	
	2) проверить правильность набора команды	$\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ ... $\overrightarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл 4.4)
	3) при обнаружении несоответствия кода команды на табло тексту программы	$\overleftarrow{\text{ШГ}}$	Набрать правильную команду
	4) при ошибочном наборе номера подпрограммы, например 35, необходимо вернуться на два шага назад и набрать правильный номер подпрограммы	$\overleftarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$ ПП 35	
	5) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции		

* Если число показателей $l < 4$, то вместо отсутствующих показателей заносятся нули. К нулю приравниваются и соответствующие "веса".

Таблица 4.4. **Текст программы вычисления расстояния между двумя объектами**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x→П 5	45	20	П→x 3	63
01	F ○	25	21	−	11
02	x→П 6	46	22	F x^2	22
03	F ○	25	23	П→x b	6*L*
04	x→П 7	47	24	ПП	53
05	F ○	25	25	35	35
06	П→x 1	61	26	П→x 5	65
07	−	11	27	П→x 4	64
08	F x^2	22	28	−	11
09	П→x 0	60	29	F x^2	22
10	ПП	53	30	П→x c	6*C*
11	35	35	31	ПП	53
12	П→x 7	67	32	35	35
13	П→x 2	62	33	F $\sqrt{\ }$	21
14	−	11	34	С/П	50
15	F x^2	22	35	x	12
16	П→x 0	60	36	П→x 8	68
17	ПП	53	37	+	10
18	35	35	38	x→П 8	48
19	П→x 6	66	39	В/О	52

4.3. КОНТРОЛЬНЫЙ ПРИМЕР

Провести классификацию пяти предприятий электротехнической промышленности (табл. 4.5), каждое из которых характеризуется следующими экономическими показателями: x_1 – прибыль от реализации (млн. руб.); x_2 – удельный вес продукции высшей категории качества (%); x_3 – выработка товарной продукции на одного работника ППП (тыс. руб.); x_4 – среднегодовая стоимость основных производственных фондов (млн. руб.).

Таблица 4.5. **Значения основных экономических показателей предприятий электротехнической промышленности**

Номер предприятия	x_1	x_2	x_3	x_4
1	3,338	78,46	5,013	7,312
2	1,909	50,83	3,423	17,785
3	6,653	26,12	3,314	21,544
4	2,105	72,11	2,534	8,125
5	6,178	13,70	1,863	1,780

Решение. Для устранения различия в единицах измерения показателей нормируем их. В результате расчетов по программе 4.2.1 получаем матрицу нормированных исходных данных

$$\mathbf{Z} = \begin{pmatrix} -0{,}34776501 & 1{,}1996448 & 1{,}688891 & -0{,}55050379 \\ -1{,}0591251 & 0{,}1026702 & 0{,}1833199 & 0{,}89186241 \\ 1{,}3024511 & -0{,}8783738 & 0{,}0801078 & 1{,}4095607 \\ -0{,}96155583 & 0{,}9475352 & -0{,}6584743 & -0{,}43853551 \\ 1{,}0699948 & -1{,}3714763 & -1{,}2938443 & -1{,}3123838 \end{pmatrix},$$

а также средние значения показателей x_1, x_2, x_3, x_4 и их средние квадратические отклонения:

$$\begin{aligned} \bar{x}_1 &= 4{,}0366; & s_1 &= 2{,}0088277; \\ \bar{x}_2 &= 48{,}244; & s_2 &= 25{,}187455; \\ \bar{x}_3 &= 3{,}2294; & s_3 &= 1{,}0560776; \\ \bar{x}_4 &= 11{,}3092; & s_4 &= 7{,}2609854. \end{aligned}$$

В качестве расстояния между объектами возьмем взвешенное евклидово расстояние, причем "веса" w_l зададим пропорционально степени важности экономического показателя: $w_1 = 0{,}4$; $w_2 = 0{,}3$; $w_3 = 0{,}2$; $w_4 = 0{,}1$. Программа 4.2.2 позволяет получить матрицу расстояний между всеми пятью предприятиями:

$$\mathbf{R}_1 = \begin{pmatrix} 0 & 1,159804 & 1,9283079 & 1,1311047 & 2,2980731 \\ & 0 & 1,6262618 & 0,77977305 & 1,8968315 \\ & & 0 & 1,9581917 & 1,1126867 \\ & & & 0 & 1,9881173 \\ & & & & 0 \end{pmatrix}$$

Из матрицы $\mathbf{R}_1$ следует, что объекты 2 и 4 наиболее близки ($\rho_{2,4} = 0,78$) и поэтому объединяются в один кластер. После объединения имеем четыре кластера:

Номер кластера	1	2	3	4
Состав кластера	(1)	(2,4)	(3)	(5)

Расстояние между кластерами будем находить по принципу "ближайшего соседа". За расстояние между кластерами S_1 и $S_{(2,4)}$ берем минимальное из расстояний $\rho_{12} = 1,159804$ и $\rho_{14} = 1,1311047$. Аналогично находим расстояния между S_3, S_5 и $S_{(2,4)}$, которые соответственно равны: $\rho_{3(2,4)} = 1,6262618$ и $\rho_{5(2,4)} = 1,8968315$. Расстояние между остальными кластерами остается без изменения. Таким образом, получаем матрицу расстояний

$$\mathbf{R}_2 = \begin{pmatrix} 0 & 1,1311047 & 1,9283079 & 2,2980731 \\ & 0 & 1,6262618 & 1,8968315 \\ & & 0 & 1,1126867 \\ & & & 0 \end{pmatrix}.$$

Из матрицы $\mathbf{R}_2$ следует, что кластеры S_3 и S_5 наиболее близки ($\rho_{35} = 1,1126867$) и поэтому объединяются в новый кластер $S_{(3,5)}$. После объединения будем иметь три кластера S_1, $S_{(2,4)}$ и $S_{(3,5)}$. Расстояния между новым кластером $S_{(3,5)}$ и кластерами S_1, $S_{(2,4)}$ соответственно равны: $\rho_{1,(3,5)} = 1,9283079$ ($\rho_{13} = 1,9283079$ меньше $\rho_{15} = 2,298073$) и $\rho_{(2,4)(3,5)} = 1,6262618$. Матрица расстояний имеет следующий вид:

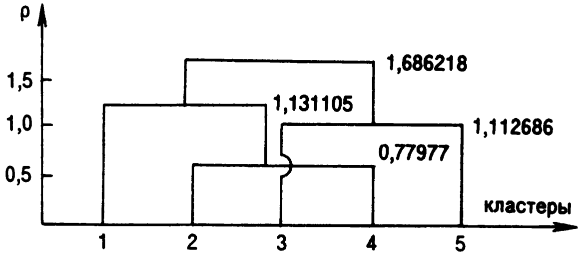
$$\mathbf{R}_3 = \begin{pmatrix} 0 & 1,1311047 & 1,9283079 \\ & 0 & 1,6262618 \\ & & 0 \end{pmatrix}.$$

Из этой матрицы следует, что кластеры S_1 и $S_{(2,4)}$ объединяются в новый кластер $S_{(1,2,4)}$, так как расстояние между ними минимально – $\rho_{1(2,4)} = 1,1311047$. Тогда получим матрицу расстояний

$$\mathbf{R}_4 = \begin{pmatrix} 0 & 1,6262618 \\ & 0 \end{pmatrix}.$$

Таким образом на расстоянии $\rho_{(1,2,4)(3,5)}$ = 1,6262618 два кластера $S_{(1,2,4)}$ и $S_{(3,5)}$ объединяются в один.

Результаты иерархической классификации наблюдений представлены на рис. 4.1 в виде дендрограммы, где по оси ординат приводятся расстояния между объединяемыми на данном этапе кластерами.



В задаче предпочтение следует отдать предпоследнему этапу классификации, когда все объекты объединены в два кластера $S_{(1,2,4)}$ и $S_{(3,5)}$, что наглядно видно на рис. 4.1.

5. ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

5.1. АЛГОРИТМ ОДНОФАКТОРНОГО ДИСПЕРСИОННОГО АНАЛИЗА

Дисперсионный анализ предназначен для выявления влияния на результативный показатель Y отдельных факторов. При этом обычно предполагается, что Y имеет закон нормального распределения. Факторные величины могут измеряться как в количественных, так и в неколичественных шкалах. Они могут иметь фиксированные и случайные уровни.

Модели дисперсионного анализа в зависимости от числа факторов классифицируются на однофакторные, двухфакторные и т. д.

Решается задача проверки влияния на результативный признак Y фактора A, имеющего m уровней A_j, где $j = 1, 2, \ldots, m$.

Предполагается, что Y подчиняется закону нормального распределения с условным математическим ожиданием μ_j, зависящим от уровней фактора A_j и постоянной дисперсии σ^2.

Задача сводится к проверке на уровне значимости α гипотезы H_0: $\mu_1 = \mu_2 = \ldots = \mu_m$.

Обозначим:

y_{ij} – результат i-го наблюдения ($i = 1, 2, \ldots, n_j$), выполненного при A_j-м уровне фактора;

$y_{*j} = \frac{1}{n_j} \sum_{i=1}^{n_j} y_{ij}$ – средняя арифметическая, соответствующая A_j-му уровню;

$y_{**} = \frac{1}{N} \sum_{j=1}^{m} n_j y_{*j}$ – общая средняя комплекса;

$N = \sum_{j=1}^{m} n_j$ – общее число наблюдений;

$Q_A = \sum_{j=1}^{m} n_j (y_{*j} - y_{**})^2$ – сумма квадратов, обусловленная влиянием фактора A;

$Q_{ост} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (y_{ij} - y_{*j})^2$ – остаточная сумма квадратов;

$Q_{общ} = \sum_{j=1}^{m} \sum_{i=1}^{n_j} (y_{ij} - y_{**})^2$ – общая сумма квадратов. Причем $Q_{общ} = Q_A + Q_{ост}$.

Для проверки гипотезы H_0: $\mu_1 = \mu_2 = \ldots = \mu_m$ вычисляется

$$F_{\text{набл}} = \frac{\frac{1}{m-1} Q_A}{\frac{1}{N-m} Q_{\text{ост}}} .$$

Если $F_{\text{набл}} \leqslant F_{\text{кр}}(\alpha, m-1, N-m)$, где $F_{\text{кр}}(\alpha, m-1, N-m)$ находится по таблице F-распределения для уровня значимости α и числа степеней свободы числителя $\nu_1 = m - 1$ и знаменателя $\nu_2 = n - m$, то гипотеза не отвергается. Из этого следует, что влияние фактора A на результативный признак Y не доказано.

Если $F_{\text{набл}} > F_{\text{кр}}(\alpha, m-1, N-m)$, то гипотеза отвергается с вероятностью ошибки α, из чего следует, что фактор A влияет на результативный признак Y.

5.2. ОДНОФАКТОРНЫЙ КОМПЛЕКС С РАВНЫМ ЧИСЛОМ НАБЛЮДЕНИЙ

5.2.1. Назначение программы

Программа однофакторного дисперсионного анализа с равным числом наблюдений на каждом уровне фактора A (табл. 5.1, 5.2), т.е. когда $n_j = n$ для всех $j = 1, 2, \ldots, m$, позволяет вести расчеты без ограничения по числу уровней m и числу наблюдений n.

Программа рассчитывает Q_A, $Q_{\text{ост}}$ и $Q_{\text{общ}}$, преобразованные к виду:

$$Q_A = \frac{1}{n} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} y_{ij} \right)^2 - \frac{1}{mn} \left(\sum_{j=1}^{m} \sum_{i=1}^{n} y_{ij} \right)^2;$$

$$Q_{\text{ост}} = \sum_{j=1}^{m} \sum_{i=1}^{n} y_{ij}^2 - \frac{1}{n} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} y_{ij} \right)^2;$$

$$Q_{\text{общ}} = \sum_{j=1}^{m} \sum_{i=1}^{n} y_{ij}^2 - \frac{1}{mn} \left(\sum_{j=1}^{m} \sum_{i=1}^{n} y_{ij} \right)^2.$$

5.2.2. Программа однофакторного дисперсионного анализа с равным числом наблюдений

Таблица 5.1. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 5.2)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести нули в регистры памяти 4, 5, 6, 7, 8	Сх x→П 4 x→П 5 .... x→П 8	
5	Занести в регистры памяти 0, 1 число наблюдений на уровне фактора – n; в регистры памяти 2, 3 – число уровней фактора – m	Набрать n x→П 0 n x→П 1 m x→П 2 m x→П 3	
6	Ввести в машину очередные числа y_{ij}	Набрать y_{ij}	
7	Запустить программу с шага 00	В/О С/П	Если машина зациклилась, нажать клавишу С/П и перейти к п. 10
8	Если введены не все данные, т.е. y_{ij}, то вновь возвращаемся к п. 6, иначе переходим к п. 9		Сначала вводят все данные первого уровня A_1, затем второго A_2 и т.д.
9	Списать из памяти значения Q_A, $Q_{общ}$, $Q_{ост}$	П→x 4 П→x 5 П→x 7	На табло Q_A $Q_{общ}$ $Q_{ост}$

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
10	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора команды 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции	F ПРГ $\overrightarrow{ШГ}$ $\overrightarrow{ШГ}$ ... $\overrightarrow{ШГ}$ $\overleftarrow{ШГ}$	Сличать высвечиваемые коды с текстом программы (табл 5.2) Набрать правильную команду

Таблица 5.2. **Текст программы однофакторного дисперсионного анализа с равным числом наблюдений**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	27	F L_2	58
01	0 3	03	28	0 2	02
02	С/П	50	29	П → x 5	65
03	x → П 4	44	30	F x^2	22
04	П → x 5	65	31	x → П 5	45
05	+	10	32	П → x 1	61
06	x → П 5	45	33	П → x 3	63
07	П → x 4	64	34	x	12
08	F x^2	22	35	F 1/x	23
09	П → x 6	66	36	П → x 5	65
10	+	10	37	x	12
11	x → П 6	46	38	x → П 7	47
12	П → x 4	64	39	П → x 8	68
13	П → x 7	67	40	П → x 1	61
14	+	10	41	:	13
15	x → П 7	47	42	П → x 7	67
16	F L_0	5Г	43	–	11
17	0 2	02	44	x → П 4	44
18	П → x 7	67	45	П → x 6	66
19	F x^2	22	46	П → x 7	67
20	П → x 8	68	47	–	11
21	+	10	48	x → П 5	45
22	x → П 8	48	49	П → x 4	64
23	П → x 1	61	50	–	11
24	x → П 0	40	51	x → П 7	47
25	0	00	52	С/П	50
26	x → П 7	47			

5.2.3. Контрольный пример

По данным годовых отчетов (табл. 5.3) исследовать влияние размера предприятия на производительность труда.

Таблица 5.3. **Производительность труда на предприятии (в рублях на одного работника промышленно-производственного персонала)**

Номер наблюдения	Вид предприятия			
	мелкое	среднее	крупное	крупнейшее
1	88	87	93	92
2	89	86	93	94
3	87	86	91	95

Решение. Объем наблюдений $N = 12$, исследуемый фактор A – размер предприятия имеет четыре уровня ($m = 4$). Рассматривалось по три предприятия ($n = 3$) каждого вида.

В результате расчетов по программе 5.2.2 получаем следующее:

$Q_A = 108{,}917$; $Q_{\text{ост}} = 10{,}0$; $Q_{\text{общ}} = 118{,}917$.

Для проверки гипотезы о влиянии фактора A вычислим

$$F_{\text{набл}} = \frac{\frac{1}{m-1} Q_A}{\frac{1}{N-m} Q_{\text{ост}}} = \frac{\frac{1}{3} \cdot 108{,}917}{\frac{1}{8} \cdot 10{,}0} = 29{,}045.$$

Сравним рассчитанное значение с критическим, полученным по таблице F-распределения; для $\alpha = 0{,}05$, $\nu_1 = m - 1 = 3$ и $\nu_2 = N - m = 8$ $F_{\text{кр}}(0{,}05; 3; 8;) = 4{,}07$. Так как наблюдаемое значение превосходит критическое, нулевая гипотеза отвергается с вероятностью ошибки 0,05. Следовательно, доказано влияние размера предприятия на производительность труда.

5.3. ОДНОФАКТОРНЫЙ КОМПЛЕКС С РАЗЛИЧНЫМ ЧИСЛОМ НАБЛЮДЕНИЙ

5.3.1. Назначение программы

Программа однофакторного дисперсионного анализа с различным числом наблюдений на каждом уровне фактора A, т.е. когда $n_j \neq n$ для некоторых $j = 1, 2, \ldots, m$, предполагает не более четырех уровней фактора $A (m \leqslant 4)$ при любом числе наблюдений n_j. Программа рассчитывает Q_A, $Q_{\text{ост}}$ и $Q_{\text{общ}}$ по формулам, приведенным в 5.1.

5.3.2. Программа однофакторного дисперсионного анализа с различным числом наблюдений

Таблица 5.4. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 5.5)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Ввести в регистр памяти 4 число 9, в регистр памяти 1 – число уровней фактора *A*, в регистр памяти 5 – число 5	9 x → П 4 m x → П 1 5 x → П 5	
5	Ввести в регистры памяти a, b, c, d соответственно количество наблюдений на 1, 2, 3 и 4 уровнях фактора A*	n_1 x → П a n_2 x → П b n_3 x → П c n_4 x → П d	
6	В регистры памяти 0, 2, 3, 6, 7, 8, 9 занести нули	Cx x → П 0 x → П 2 x → П 3 x → П 6 .... x → П 9	
7	Запустить программу с шага 00	В/О С/П	
8	Набрать первое наблюдение первого уровня (y_{11}). После останова машины набрать второе наблюдение первого уровня и т.д. После запуска последнего наблюдения первого уровня набрать первое наблюдение второго уровня, затем после останова машины набрать y_{22} и т.д.	Набрать y_{11} С/П y_{21} С/П .... $y_{n_1 1}$ С/П y_{12} С/П y_{22} С/П .... $y_{n_m m}$ С/П	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
9	Выписать из регистров памяти средние $y_{*1}, y_{*2}, y_{*3}, y_{*4}, y_{**}$**	П → x 6 П → x 7 П → x 8 П → x 9 П → x 3	На табло: y_{*1} y_{*2} y_{*3} y_{*4} y_{**}
10	Повторить шаг 8		
11	Выписать из регистров памяти Q_A, $Q_{ост}$, $Q_{общ}$***	П → x 4 П → x 2 П → x 3	На табло: $Q_{общ}$ $Q_{ост}$ Q_A

* В регистры памяти, для которых нет данных (т.е. число уровней фактора A $m < 4$), заносятся нули вместо n_j.

** Если число уровней фактора A $m < 4$, то из регистров памяти извлекаются средние (y_{*j}) только для имеющихся уровней фактора A.

*** Если число уровней фактора A $m < 4$, то из памяти извлекаются только $Q_{общ}$ и $Q_{ост}$, а Q_A вычисляется в автоматическом режиме как разность $Q_{общ}$ и $Q_{ост}$.

Таблица 5.5. **Текст программы однофакторного дисперсионного анализа с различным числом наблюдений**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	К П → x 4	*Г*4	20	К П → x 4	*Г*4
01	x → П 0	40	21	:	13
02	0	00	22	К x → П 5	*L*5
03	x → П 9	49	23	F L_1	5*L*
04	С/П	50	24	0 0	00
05	x → П 2	42	25	П → x a	6–
06	П → x 9	69	26	П → x b	6*L*
07	+	10	27	+	10
08	x → П 9	49	28	П → x c	6*C*
09	П → x 2	62	29	+	10
10	П → x 3	63	30	П → x d	6*Г*
11	+	10	31	+	10
12	x → П 3	43	32	F 1/x	23
13	F L_0	5*Г*	33	П → x 3	63
14	0 4	04	34	x	12
15	П → x 4	64	35	x → П 3	43
16	1	01	36	4	04
17	–	11	37	x → П 1	41
18	x → П 4	44	38	0	00
19	П → x 9	69	39	x → П 2	42

Шаг	Клавиши	Код	Шаг	Клавиши	Код
40	x → П 4	44	69	БП	51
41	П → x a	6–	70	9 1	91
42	x → П 0	40	71	П → x 1	61
43	С/П	50	72	3	03
44	x → П 5	45	73	–	11
45	П → x 6	66	74	F x=0	5E
46	–	11	75	8 2	82
47	F x^2	22	76	П → x c	6C
48	П → x 2	62	77	x → П 0	40
49	+	10	78	П → x 8	68
50	x → П 2	42	79	x → П 6	46
51	П → x 5	65	80	БП	51
52	П → x 3	63	81	9 1	91
53	–	11	82	П → x 1	61
54	F x^2	22	83	2	02
55	П → x 4	64	84	–	11
56	+	10	85	F x=0	5E
57	x → П 4	44	86	9 1	91
58	F L_0	5Г	87	П → x d	6Г
59	4 3	43	88	x → П 0	40
60	П → x 1	61	89	П → x 9	69
61	4	04	90	x → П 6	46
62	–	11	91	F L_1	5L
63	F x=0	5E	92	4 3	43
64	7 1	71	93	П → x 4	64
65	П → x b	6L	94	П → x 2	62
66	x → П 0	40	95	–	11
67	П → x 7	67	96	x → П 3	43
68	x → П 6	46	97	С/П	50

5.3.3. Контрольный пример

Исследовать влияние количества осадков за год на урожайность пшеницы (ц/га) по данным табл. 5.6.

Таблица 5.6. **Урожайность пшеницы, ц/га**

Номер наблюдения	Количество осадков, мм			
	250–260	260–270	270–280	280–290
1	31	27	35	33
2	30	29	34	37
3		28	34	36
4		30	36	
5			35	

Решение. Объем наблюдений $N = 14$, исследуемый фактор A (количество осадков за год) имеет четыре уровня ($m = 4$), каждому из которых соответствует различное число наблюдений: $n_1 = 2$; $n_2 = 4$; $n_3 = 5$; $n_4 = 3$.

Предварительно проверим гипотезу об однородности остаточных дисперсий, т.е. H_0: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$.

Так как $n_j = n$ для всех $j = 1, 2, 3, 4$, то для проверки гипотезы воспользуемся критерием Кохрана и определим

$$G_{\text{набл}} = \frac{\max\limits_j \hat{s}_j^2}{\sum\limits_{j=1}^{m} \hat{s}_j^2},$$

где $\max\limits_j \hat{s}_j^2$ – наибольшая из исправленных оценок дисперсий

$$\hat{s}_j^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_{ij} - y_{*j})^2.$$

По таблице G-распределения (табл. П.1.7) для уровня значимости $\alpha = 0{,}05$, числа сравниваемых совокупностей $l = 4$ и степеней свободы $\nu = n - 1 = 2$ найдем $G_{\text{кр}} = 0{,}768$.

Для нашего примера

$$G_{\text{набл}} = \frac{2{,}333}{5{,}000} = 0{,}4667.$$

Так как $G_{\text{набл}} < G_{\text{кр}}$, то гипотеза о равенстве дисперсий не отвергается. Примем, что $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$, и перейдем к проверке гипотезы о равенстве генеральных средних, т.е. проверим гипотезу H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

В результате расчетов по программе 5.3.2 после первого ввода исходных данных извлекаем из регистров памяти величины:

$y_{*1} = 30{,}5$; $y_{*2} = 28{,}5$; $y_{*3} = 34{,}8$; $y_{*4} = 35{,}3$; $y_{**} = 32{,}5$.

После окончания работы всей программы, т.е. после второго ввода исходных данных, получаем результаты:

$Q_{\text{общ}} = 139{,}5$; $Q_A = 122{,}533$; $Q_{\text{ост}} = 16{,}967$.

Для проверки гипотезы о влиянии фактора A вычислим

$$F_{\text{набл}} = \frac{\frac{1}{3}\,122{,}533}{\frac{1}{10}\,16{,}967} = 24{,}073,$$

которое сравним с критическим значением, полученным по таблице F-распределения; при $\alpha = 0{,}05$; $\nu_1 = 3$ и $\nu_2 = 10$ $F_{\text{кр}}(0{,}05;\ 3;\ 10) = 3{,}71$. Так как наблюдаемое значение превосходит критическое, нулевая гипотеза отвергается с вероятностью ошибки 0,05, что доказывает влияние количества осадков на урожайность пшеницы.

6. ДВУХФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

6.1. АЛГОРИТМ ДВУХФАКТОРНОГО ДИСПЕРСИОННОГО АНАЛИЗА

По природе факторов модели дисперсионного анализа подразделяются на фиксированные ($M1$), если все факторы, входящие в модель, имеют фиксированные уровни, на случайные ($M2$), если все факторы имеют случайные уровни, и смешанные, у которых часть факторов имеют случайные, а часть – фиксированные уровни.

Проверяются влияние на результативный признак Y фактора A, имеющего m уровней A_i ($i = 1, 2,..., m$), фактора B, имеющего r уровней B_j ($j = 1, 2,..., r$), и их взаимодействия, причем каждому сочетанию уровней A_iB_j соответствует n наблюдений ($k = 1, 2,..., n$). Таким образом, общее число наблюдений в двухфакторном комплексе $N = mrn$.

В двухфакторном дисперсионном анализе проверяются гипотезы H_0 об отсутствии влияния на результативный признак Y: фактора A; фактора B; взаимодействия двух факторов (A и B). Для проверки гипотез используют следующее разложение общей вариации результативного признака:

$$Q_{общ} = Q_A + Q_B + Q_{AB} + Q_{ост},$$

где $Q_{общ} = \sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - y_{***})^2$ – общая сумма квадратов;

$Q_A = rn \sum_{i=1}^{m}(y_{i**} - y_{***})^2$ – сумма квадратов, измеряющая главное влияние фактора A;

$Q_B = mn \sum_{j=1}^{r}(y_{*j*} - y_{***})^2$ – сумма квадратов, измеряющая главное влияние фактора B;

$Q_{AB} = n \sum_{i=1}^{m}\sum_{j=1}^{r}[y_{ij*} - (y_{i**} + y_{*j*}) + y_{***}]^2$ – сумма квадратов, измеряющая взаимодействие факторов A и B;

$Q_{ост} = \sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n}(y_{ijk} - y_{ij*})^2$ – остаточная сумма квадратов, обусловленная влиянием всех неконтролируемых факторов.

Здесь y_{ijk} – результат k-го наблюдения, выполненного при A_i-м уровне фактора A и B_j-м уровне фактора B;

$$y_{***} = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \text{ – общая средняя;}$$

$$y_{i**} = \frac{1}{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \text{ – средняя, соответствующая } A_i\text{-му уровню;}$$

$$y_{ij*} = \frac{1}{n} \sum_{k=1}^{n} y_{ijk} \text{ – средняя внутри ячейки } (A_i\ B_j).$$

Для проверки гипотез об отсутствии влияния факторов A, B и их взаимодействия используют F-критерий, статистику которого выбирают по табл. 6.1 в зависимости от вида модели.

Полученные значения $F_{набл}$ сравниваются с табличными. Если $F_{набл} \leq F_{кр}(\alpha, \nu_1, \nu_2)$, где $F_{кр}(\alpha, \nu_1, \nu_2)$ находится по таблице F-распределения по уровню значимости α, числу степеней свободы числителя ν_1 и знаменателя ν_2, то гипотеза не отвергается. Из этого следует, что влияние фактора (или взаимодействия факторов) на результативный признак Y не доказано.

Если $F_{набл} \geq F_{кр}(\alpha, \nu_1, \nu_2)$, то гипотеза H_0 отвергается с вероятностью ошибки, равной α. Из этого следует, что соответствующий фактор (взаимодействие факторов) влияет на результативный признак Y.

Трехфакторный дисперсионный анализ отличается от двухфакторного наличием большего разнообразия взаимодействий факторов, что ведет к значительному увеличению объема вычислений и усложнению содержательной интерпретации результатов. Начиная с трех факторов, в дисперсионном анализе без дополнительных условий не удается проверить влияние всего разнообразия факторов и их взаимодействия.

Таблица 6.1. **Двухфакторный комплекс дисперсионного анализа**

Сумма квадратов	Число степеней свободы	Средняя сумма квадратов		$M1$	$M2$	Смешанная модель	
						A – случайные уровни	B – случайные уровни
				$F_{набл}$	$F_{набл}$	$F_{набл}$	$F_{набл}$
Q_A	$m-1$	$Q_A/(m-1)$	(1)	(1) : (4)	(1) : (3)	(1) : (4)	(1) : (3)
Q_B	$r-1$	$Q_B/(r-1)$	(2)	(2) : (4)	(2) : (3)	(2) : (3)	(2) : (4)
Q_{AB}	$(m-1)(r-1)$	$Q_{AB}/(m-1)(r-1)$	(3)	(3) : (4)	(3) : (4)	(3) : (4)	(3) : (4)
$Q_{ост}$	$mr(n-1)$	$Q_{ост}/mr(n-1)$	(4)				

6.2. СИСТЕМА ПРОГРАММ ДВУХФАКТОРНОГО ДИСПЕРСИОННОГО АНАЛИЗА

6.2.1. Назначение программ

При разработке программного обеспечения двухфакторного дисперсионного анализа использовались следующие преобразованные формулы:

$$Q_{\text{общ}} = \sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{mrn} \left(\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2;$$

$$Q_A = \frac{1}{rn} \sum_{i=1}^{m} \left(\sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2 - \frac{1}{mrn} \left(\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2 ;$$

$$Q_B = \frac{1}{mn} \sum_{j=1}^{r} \left(\sum_{i=1}^{m} \sum_{k=1}^{n} y_{ijk} \right)^2 - \frac{1}{mrn} \left(\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2;$$

$$Q_{AB} = \frac{1}{n} \sum_{i=1}^{m} \sum_{j=1}^{r} \left(\sum_{k=1}^{n} y_{ijk} \right)^2 - \frac{1}{mn} \sum_{j=1}^{r} \left(\sum_{i=1}^{m} \sum_{k=1}^{n} y_{ijk} \right)^2 -$$

$$- \frac{1}{nr} \sum_{i=1}^{m} \left(\sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2 + \frac{1}{mrn} \left(\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2;$$

$$Q_{\text{ост}} = \sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk}^2 - \frac{1}{n} \sum_{i=1}^{m} \sum_{j=1}^{r} \left(\sum_{k=1}^{n} y_{ijk} \right)^2.$$

Предлагаемые программы для проведения расчетов на программируемом микрокалькуляторе предусматривают не более четырех уровней фактора $A (m \leqslant 4)$ при любом числе уровней фактора B и любом числе наблюдений n.

Программа 6.2.2 (табл. 6.2 и 6.3) предназначена для вычисления сумм: $\left(\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2$, $\sum_{i=1}^{m} \sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk}^2$, $\sum_{i=1}^{m} \sum_{j=1}^{r} \left(\sum_{k=1}^{n} y_{ijk} \right)^2$, $\sum_{i=1}^{m} \left(\sum_{j=1}^{r} \sum_{k=1}^{n} y_{ijk} \right)^2$, $\sum_{j=1}^{r} \left(\sum_{i=1}^{m} \sum_{k=1}^{n} y_{ijk} \right)^2$, а программа 6.2.3 (табл. 6.4. и 6.5) предназначена для вычисления квадратичных форм Q_A, Q_B, Q_{AB}, $Q_{\text{ост}}$ и $Q_{\text{общ}}$.

6.2.2. Программа вычисления сумм

Таблица 6.2. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 6.3)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем — правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести нули в регистры памяти 3, 4, 5,..., 9, a, b, c, d	Cx x → П 3 x → П 4 .... x → П 9 x → П a .... x → П d	
5	Занести в регистр памяти 0 число наблюдений в ячейке n; в регистр памяти 1 — число уровней фактора A — m; в регистр памяти 2 — число уровней фактора B — r	Набрать n x → П 0 m x → П 1 r x → П 2	
6	Ввести в машину очередные числа y_{ijk}	Набрать y_{ijk}	
7	Запустить машину с шага 00	В/О С/П	Если машина зациклилась, нажать клавишу С/П и перейти к п. 10
8	Если введены не все данные, т.е. y_{ijk}, то вновь возвращаемся к п. 6, иначе переходим к п. 9		Данные вводят по столбцам: сначала все элементы ячейки 1.1, затем 2.1 и т.д.

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
9	Списать из памяти значения	П → x 5 П → x 6 П → x 8 П → x 3 П → x 9	На табло: $(\sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk})^2$ $\sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk}^2$ $\sum_{i=1}^{m}\sum_{j=1}^{r}(\sum_{k=1}^{n} y_{ijk})^2$ $\sum_{j=1}^{r}(\sum_{i=1}^{m}\sum_{k=1}^{n} y_{ijk})^2$ $\sum_{i=1}^{m}(\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk})^2$
10	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора команды 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции	F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ ... $\overrightarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$	Сличать высвечиваемые коды с текстом программы (табл 6.3) Набрать правильную команду

Таблица 6.3. **Текст программы вычислений**

$$(\sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk})^2,\ \sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk}^2,\ \sum_{i=1}^{m}\sum_{j=1}^{r}(\sum_{k=1}^{n} y_{ijk})^2,\ \sum_{i=1}^{m}(\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk})^2,$$

$$\sum_{j=1}^{r}(\sum_{i=1}^{m}\sum_{k=1}^{n} y_{ijk})^2$$

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	БП	51	06	x → П 5	45
01	0 3	03	07	П → x 4	64
02	С/П	50	08	F x^2	22
03	x → П 4	44	09	П → x 6	66
04	П → x 5	65	10	+	10
05	+	10	11	x → П 6	46

Шаг	Клавиши	Код
12	П → x 4	64
13	П → x 7	67
14	+	10
15	x → П 7	47
16	П → x 4	64
17	П → x 9	69
18	+	10
19	x → П 9	49
20	F L_0	5Г
21	0 2	02
22	П → x 7	67
23	F x^2	22
24	П → x 8	68
25	+	10
26	x → П 8	48
27	П → x 1	61
28	4	04
29	−	11
30	F x = 0	5E
31	3 8	38
32	П → x 7	67
33	П → x a	6−
34	+	10
35	x → П a	4−
36	БП	51
37	6 4	64
38	П → x 1	61
39	3	03
40	−	11
41	F x = 0	5E
42	4 9	49
43	П → x 7	67
44	П → x b	6L
45	+	10
46	x → П b	4L
47	БП	51
48	6 4	64
49	П → x 1	61
50	2	02
51	−	11
52	F x = 0	5E
53	6 0	60
54	П → x 7	67
55	П → x c	6C
56	+	10
57	x → П c	4C
58	БП	51
59	6 4	64
60	П → x 7	67
61	П → x d	6Г
62	+	10
63	x → П d	4Г
64	0	00
65	x → П 7	47
66	Набрать *n*	*n**
67	x → П 0	40
68	F L_1	5L
69	0 2	02
70	П → x 9	69
71	F x^2	22
72	П → x 3	63
73	+	10
74	x → П 3	43
75	0	00
76	x → П 9	49
77	Набрать *m*	0*m***
78	x → П 1	41
79	F L_2	58
80	0 2	02
81	П → x a	6−
82	F x^2	22
83	П → x b	6L
84	F x^2	22
85	+	10
86	П → x c	6C
87	F x^2	22
88	+	10
89	П → x d	6Г
90	F x^2	22
91	+	10
92	x → П 9	49
93	П → x 5	65
94	F x^2	22
95	x → П 5	45
96	С/П	50

* *n* — число наблюдений.

** *m* — число уровней фактора *A*.

6.2.3. Программа вычисления квадратичных форм Q_A, Q_B, Q_{AB}, $Q_{ост}$, $Q_{общ}$

Таблица 6.4. **Инструкция по работе с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 6.5)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистр памяти a количество наблюдений в ячейке – n; в регистр памяти b – число уровней фактора A – m; в регистр памяти c – число уровней фактора B – r; в регистры 0, 1, 2, 4, 7, d заносятся нули	Набрать n x → П a m x → П b r x → П c Cx x → П 0 x → П 1 x → П 2 x → П 4 x → П 7 x → П d	
5	Ввести в машину значения	Набрать $r\ m\ n$ $\Sigma(\Sigma\ \Sigma y_{ijk})^2$ x → П 3 $m\ r\ n$ $(\Sigma\ \Sigma\Sigma y_{ijk})^2$ x → П 5 $m\ r\ n$ $\Sigma\ \Sigma\Sigma\ y^2_{ijk}$ x → П 6 $m\ r\ n$ $\Sigma\ \Sigma(\Sigma y_{ijk})^2$ x → П 8 $m\ \ r\ n$ $\Sigma\ (\Sigma\Sigma y_{ijk})^2$ x → П 9	
6	Запустить программу с шага 00	В/О С/П	При зацикливании нажать клавишу С/П и перейти к п. 8
7	Списать из памяти результаты	П → x 9 П → x 3 П → x 6 П → x 8 П → x 7	На табло: Q_A Q_B Q_{AB} $Q_{ост}$ $Q_{общ}$

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
8	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора команды 3) при обнаружении несоответствия кода команды на табло тексту программы 4) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 инструкции	F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ . . . $\overrightarrow{\text{ШГ}}$ $\overleftarrow{\text{ШГ}}$	 Сличать высвечиваемые коды с текстом программы (табл 6.5) Набрать правильную команду

Примечание. П. 5 выполняется, если после проведения расчетов по программе 6.2.2 микрокалькулятор выключался или в регистрах памяти 3, 5, 6, 8, 9 не сохранены результаты расчетов по этой программе.

Таблица 6.5. **Текст программы вычисления квадратичных форм Q_A, Q_B, Q_{AB}, $Q_{ост}$ и $Q_{общ}$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	П→x a	6–	22	П→x c	6*C*
01	П→x b	6*L*	23	x	12
02	x	12	24	F 1/x	23
03	П→x c	6*C*	25	П→x 9	69
04	x	12	26	x	12
05	F 1/x	23	27	П→x 5	65
06	П→x 5	65	28	–	11
07	x	12	29	x→П 9	49
08	x→П 5	45	30	П→x 6	66
09	П→x a	6–	31	П→x 5	65
10	П→x b	6*L*	32	–	11
11	x	12	33	x→П 7	47
12	F 1/x	23	34	П→x 3	63
13	П→x 3	63	35	П→x 5	65
14	x	12	36	–	11
15	x→П 3	43	37	x→П 3	43
16	П→x a	6–	38	П→x 6	66
17	F 1/x	23	39	П→x 8	68
18	П→x 8	68	40	–	11
19	x	12	41	x→П 8	48
20	x→П 8	48	42	П→x 7	67
21	П→x a	6–	43	П→x 3	63

Шаг	Клавиши	Код
44	–	11
45	П → x 9	69
46	–	11
47	П → x 8	68

Шаг	Клавиши	Код
48	–	11
49	x → П 6	46
50	С/П	50

6.3. КОНТРОЛЬНЫЙ ПРИМЕР

В течение трех рабочих дней недели ($n = 3$) исследовалась зависимость времени, проведенного покупателем в очереди для оплаты покупок в крупных универсамах, от времени посещения магазина (фактор A) и числа работающих кассиров-контролеров (фактор B). Результаты наблюдений представлены в табл. 6.6.

Таблица 6.6. **Время, проведенное покупателем в очереди, мин**

Время суток (фактор A)	Число работающих касс (фактор B)		
	4	6	8
8 – 11	15, 17, 16	7, 6, 9	3, 4, 2
11 – 14	11, 9, 12	5, 3, 3	1, 2, 2
14 – 17	14, 13, 11	5, 7, 6	3, 2, 3
17 – 20	20, 22, 24	10, 9, 7	4, 6, 5

Из табл. 6.6 следует, что фактор A имеет четыре уровня ($m = 4$), фактор B – три ($r = 3$), а в каждой ячейке $n = 3$ наблюдениям.

Можно предположить, что фактор A имеет случайные уровни, а фактор B – фиксированные, т.е. имеется смешанная модель двухфакторного дисперсионного анализа.

Решение. В результате расчетов по программе 6.2.2 (табл. 6.2, 6.3) получим значения следующих величин:

$$\left(\sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk}\right)^2 = 89401; \quad \sum_{i=1}^{m}\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk}^2 = 3761;$$

$$\sum_{i=1}^{m}\sum_{j=1}^{r}\left(\sum_{k=1}^{n} y_{ijk}\right)^2 = 11165; \quad \sum_{i=1}^{m}\left(\sum_{j=1}^{r}\sum_{k=1}^{n} y_{ijk}\right)^2 = 24219;$$

$$\sum_{j=1}^{r}\left(\sum_{i=1}^{m}\sum_{k=1}^{n} y_{ijk}\right)^2 = 41309.$$

По программе 6.2.3 (табл. 6.4, 6.5) рассчитываются суммы квадратов Q_A, Q_B, Q_{AB}, $Q_{ост}$ и $Q_{общ}$. Для данных примеров имеем:

$Q_A = 207{,}6389$; $Q_B = 959{,}0557$; $Q_{AB} = 71{,}6110$;

$Q_{ост} = 39{,}3334$; $Q_{общ} = 1277{,}639$.

Для проверки гипотезы об отсутствии влияния фактора A, т.е. H_0: $\sigma_\alpha^2 = 0$, находим (см. табл. 6.1, гр. 6)

$$F_A = \frac{\frac{1}{m-1} Q_A}{\frac{1}{mnr - mr} Q_{ост}} = 68{,}5779,$$

которое сравниваем с табличным значением $F_{кр}(\alpha = 0{,}05;\ \nu_1 = 3;\ \nu_2 = 24) = 3{,}01$. Так как $F_A > F_{кр}$, то делаем вывод о влиянии фактора A, т.е. время, затраченное покупателем для оплаты покупок, зависит от момента посещения универсама.

Для проверки гипотезы об отсутствии влияния фактора B, т.е. H_0: $\sum_{j=1}^{r} \beta_j^2 = 0$, находим (см. табл. 6.1, гр. 6)

$$F_B = \frac{\frac{1}{r-1} Q_B}{\frac{1}{(m-1)(r-1)} Q_{AB}} = 73{,}1482,$$

которое сравниваем с табличным значением $F_{кр}$ $(\alpha = 0{,}05;\ \nu_1 = 2;\ \nu_2 = 6) = 5{,}14$. Так как $F_B > F_{кр}$, то делаем вывод о влиянии фактора B, т.е. число работающих кассиров-контролеров влияет на время, затраченное покупателем в очереди.

Для проверки гипотезы об отсутствии совместного влияния факторов A и B, т.е. H_0: $\sigma_{\alpha\beta}^2 = 0$, находим (см. табл. 6.1, гр. 6)

$$F_{AB} \approx \frac{\frac{1}{(m-1)(r-1)} Q_{AB}}{\frac{1}{mr(n-1)} Q_{ост}} = 7{,}8838,$$

которое сравниваем с $F_{кр}$ $(\alpha = 0{,}05;\ \nu_1 = 6;\ \nu_2 = 24) = 2{,}51$. Так как $F_{AB} > F_{кр}$, то делаем вывод о влиянии взаимодействия факторов A и B на результативный показатель.

Таким образом время, затраченное покупателем в очереди для оплаты покупок, зависит от времени суток (фактор A), числа работающих касс (фактор B) и взаимного влияния этих факторов.

7. ОЦЕНИВАНИЕ ПАРАМЕТРОВ И ПРОВЕРКА ГИПОТЕЗЫ О НОРМАЛЬНОМ ЗАКОНЕ РАСПРЕДЕЛЕНИЯ

7.1. ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Методику оценивания параметров и проверки гипотезы о нормальном законе распределения покажем на сквозном примере.

По результатам выборочного обследования 100 однотипных предприятий получены следующие данные объема основных фондов (табл. 7.1).

Таблица 7.1. **Объем основных фондов 100 предприятий легкой промышленности, млн. руб.**

5,56	5,43	5,47	5,47	5,33	5,37	5,43	5,54	5,61
5,33	5,43	5,61	5,11	5,43	5,33	5,54	5,33	5,11
5,54	5,43	5,33	5,54	5,43	5,43	5,43	5,33	5,11
5,43	5,43	5,43	5,33	5,40	5,43	5,47	5,68	5,47
5,43	5,68	5,21	5,33	5,58	5,47	5,47	5,21	5,54
5,64	5,47	5,27	5,27	5,37	5,33	5,47	5,47	5,54
5,40	5,58	5,47	5,27	5,05	5,79	5,79	5,64	5,64
5,71	5,85	5,47	5,47	5,43	5,47	5,54	5,64	5,64
5,79	5,03	5,33	5,68	5,43	5,61	5,54	5,64	5,54
5,39	5,33	5,21	5,68	5,54	5,33	5,21	5,21	5,81
5,27	5,64	5,27	5,27	5,33	5,37	5,27	5,54	5,54
5,47								

Рассматривается задача проверки гипотезы о нормальном законе распределения генеральной совокупности. Для ее решения из генеральной совокупности берется выборка объемом $n = 100$. По этой выборке строится интервальный вариационный ряд (п. 7.2), вычисляются выборочные характеристики распределения генеральной совокупности (п. 7.3, 7.4). По этим характеристикам находятся теоретические частоты (п. 7.6).

В результате сравнения теоретических и эмпирических частот с помощью критерия согласия χ^2 (Пирсона) делается вывод о том, подчиняется или нет рассматриваемая генеральная совокупность нормальному закону распределения (п. 7.7).

7.2. ПОСТРОЕНИЕ ИНТЕРВАЛЬНОГО ВАРИАЦИОННОГО РЯДА РАСПРЕДЕЛЕНИЯ

Построение интервального вариационного ряда распределения включает следующие этапы.

1. Определение среди имеющихся наблюдений (табл. 7.1) минимального $x_{\min}$ и максимального $x_{\max}$ значений признака. В данном примере это будут $x_{\min} = 5{,}03$ и $x_{\max} = 5{,}85$.

2. Определение размаха варьирования признака $R = x_{\max} - x_{\min} = 5{,}85 - 5{,}03 = 0{,}82$.

3. Определение длины интервала по формуле

$h = \dfrac{R}{1 + 3{,}32 \lg n}$, где n – объем выборки.

В данном примере $h = \dfrac{0{,}82}{1 + 3{,}32 \lg 100} = 0{,}11$ (млн. руб.).

4. Определение граничных значений интервалов $(a_i - b_i)$. Так как $x_{\min}$ и x_{max} являются случайными величинами, рекомендуется отступить влево от нижнего предела варьирования $(x_{\min})$.

За нижнюю границу первого интервала предлагается принимать величину, равную $a_1 = x_{\min} - h/2$.

Если оказывается, что $a_1 < 0$, хотя по смыслу это величина неотрицательная, то можно принять $a_1 = 0$.

Верхняя граница первого интервала $b_1 = a_1 + h$. Тогда, если b_i – верхняя граница i-го интервала (причем $a_{i+1} = b_i$), то $b_2 = a_2 + h$, $b_3 = a_3 + h$ и т.д. Построение интервалов продолжается до тех пор, пока начало следующего по порядку интервала не будет равно или больше $x_{\max}$.

В примере граничные значения составят:

$a_1 = 5{,}03 - \dfrac{0{,}11}{2} = 4{,}97$; $b_1 = 4{,}97 + 0{,}11 = 5{,}08$;

$a_2 = 5{,}08$; $b_2 = 5{,}08 + 0{,}11 = 5{,}19$ и т.д.

Границы последовательных интервалов запишем в первой графе табл. 7.2.

5. Сгруппируем результаты наблюдений.

Просматриваем статистические данные в том порядке, в каком они записаны в табл. 7.1, и значения признака разносим по соответствующим интервалам, обозначая их черточками: I, II, III, IIII, IIIII (по одной для каждого наблюдения). Так как граничные значения признака могут совпадать с границами интервалов, то условимся в каждый интервал включать варианты, *большие*, чем нижняя граница интервала

($x_i > a_i$), и *меньшие* или *равные* верхней границе ($x_i \leqslant b_i$). Общее количество штрихов, отмеченных в интервале (табл. 7.2, гр. 2) даст его частоту (табл. 7.2, гр. 3). В результате получим интервальный статистический ряд распределения частот (табл. 7.2, гр. 1 и 3).

Число интервалов обычно берут равным от 7 до 11 в зависимости от числа наблюдений и точности измерений с таким расчетом, чтобы интервалы были достаточно наполнены частотами. Если получают интервалы с нулевыми частотами, то нужно увеличить ширину интервала (особенно в середине интервального ряда).

Таблица 7.2. **Интервальный ряд распределения объемов основных фондов 100 предприятий**

Интервалы $a_i - b_i$	Подсчет частот	Частота m_i	Накопленная частота $m_{нi}$
4,97–5,08	II	2	2
5,08–5,19	III	3	5
5,19–5,30	IIIII IIIII II	12	17
5,30–5,41	IIIII IIIII IIIII IIII	19	36
5,41–5,52	IIIII IIIII IIIII IIIII IIIII IIII	29	65
5,52–5,63	IIIII IIIII IIIII III	18	83
5,63–5,74	IIIII IIIII III.	13	96
5,74–5,85	IIII	4	100

7.3. ВЫЧИСЛЕНИЕ ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК РАСПРЕДЕЛЕНИЯ

Для вычисления средней арифметической, дисперсии, коэффициентов асимметрии и эксцесса рекомендуется следующий порядок вычислений.

Заменяем интервальный ряд дискретным, для чего все значения признака в пределах интервала приравниваем к его серединному значению, и считаем, что частота относится к середине интервала. Значения середин интервалов равны $x_i = \frac{a_i + b_i}{2}$.

Для удобства вычислений целесообразно составить табл. 7.3. Значения середин интервалов заносят в первую графу, соответствующие частоты – во вторую графу и т.д.

В таблице $\Delta_i = (x_i - \bar{x})$, где $i = 1, 2,..., l$.

Таблица 7.3. **Вспомогательная таблица для вычисления выборочных характеристик**

x_i	m_i	$\Delta_i m_i$	$\Delta_i^2 m_i$	$\Delta_i^3 m_i$	$\Delta_i^4 m_i$
x_1	m_1	$\Delta_1 m_1$	$\Delta_1^2 m_1$	$\Delta_1^3 m_1$	$\Delta_1^4 m_1$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
x_i	m_i	$\Delta_i m_i$	$\Delta_i^2 m_i$	$\Delta_i^3 m_i$	$\Delta_i^4 m_i$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
x_l	m_l	$\Delta_l m_l$	$\Delta_l^2 m_l$	$\Delta_l^3 m_l$	$\Delta_l^4 m_l$
	$\sum_i m_i$	0	$\sum_i \Delta_i^2 m_i$	$\sum_i \Delta_i^3 m_i$	$\sum_i \Delta_i^4 m_i$

Пользуясь табл. 7.3, можно вычислить среднюю арифметическую

$$\bar{x} = \frac{\Sigma x_i m_i}{\Sigma m_i},$$

которая характеризует среднее положение наблюдаемых значений, и выборочный центральный момент k-го порядка

$$\hat{\mu}_k = \frac{\Sigma (x_i - \bar{x})^k m_i}{\Sigma m_i}, \quad k = 1, 2, 3, 4.$$

Выборочные коэффициенты асимметрии $\hat{\mathrm{Ac}}$ и эксцесса $\hat{\mathrm{Ek}}$ определяются по формулам

$$\hat{\mathrm{Ac}} = \frac{\hat{\mu}_3}{s^3}; \quad \hat{\mathrm{Ek}} = \frac{\hat{\mu}_4}{s^4} - 3.$$

Наиболее трудоемкие вычисления предлагается выполнять по программе на программируемом микрокалькуляторе. Описание программы и порядок работы с ней приведены ниже.

7.4. СИСТЕМА ПРОГРАММ ВЫЧИСЛЕНИЯ ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК РАСПРЕДЕЛЕНИЯ

7.4.1. Назначение программ

Входными данными программ являются столбцы x_i и m_i табл. 7.3. Предварительно по программе 7.4.2 вычисляется средняя $\bar{x}$.

Остальные характеристики вычисляются по программе 7.4.3 (табл. 7.6), которая позволяет получать:

а) в каждой строке табл. 7.3 значения величин $(x_i - \bar{x})m_i$, $(x_i - \bar{x})^2 m_i$, $(x_i - \bar{x})^3 m_i$, $(x_i - \bar{x})^4 m_i$;

б) значения сумм соответствующих столбцов $\sum_i (x_i - \bar{x})m_i$, $\sum_i m_i$ (для контроля правильности ввода исходных данных), $\sum_i (x_i - \bar{x})^2 m_i$, $\sum_i (x_i - \bar{x})^3 m_i$, $\sum_i (x_i - \bar{x})^4 m_i = \sum_i \Delta_i^4 m_i$.

В программе для списывания промежуточных результатов и ввода новых данных *предусмотрены остановы машины.* Для этого программа разбивается на две логически законченные части.

Первая часть располагается в программной памяти машины с адреса 00. Эта часть вычисляет величины, указанные в п. *а* и значения сумм, указанных в п. *б*. Входными данными являются x_i и m_i.

Вторая часть располагается в программной памяти машины с адреса 34. Результатом работы этой части являются значения величин s, s^2, $\hat{\mu}_3$, $\hat{\mu}_4$, $\hat{Ac}$, $\hat{Ek}$.

В случае необходимости повторного вычисления этих величин достаточно выполнить инструкцию работы с программой начиная с п. 10.

7.4.2. Программа вычисления средней арифметической ($\bar{x}$)

Таблица 7.4. **Инструкция работы с программой вычисления $\bar{x}$**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы вычисления $\bar{x}$	Следить за кодами команд
3	Перевести машину в автоматический режим	F АВТ	
4	Занести нули в регистры 1, 2	Сх х→П 1 х→П 2	
5	Ввести в программу очередную пару чисел: x_i, m_i	Набрать x_i В↑ Набрать m_i	
6	Запустить программу с шага 00	В/О С/П	
7	Если заполнены не все строки табл. 7.3, то перейти к п. 5, иначе – к п. 8		
8	Запустить программу с шага 11	С/П	На табло $\bar{x}$
9	После окончания счета получены результаты	П→х 0 П→х 2	На табло $\bar{x}$ На табло $\sum_i m_i$

Таблица 7.5. **Текст программы вычисления $\bar{x}$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
1	x → П 3	43	9	x → П 2	42
2	x	12	10	С/П	50
3	П → x 1	61	11	F 1/x	23
4	+	10	12	П → x 1	61
5	x → П 1	41	13	x	12
6	П → x 2	62	14	x → П 0	40
7	П → x 3	63	15	С/П	50
8	+	10			

7.4.3. Программа вычисления выборочных характеристик s^2, s, $\widehat{Ac}$, $\widehat{Ek}$, $\hat{\mu}_3$, $\hat{\mu}_4$

Таблица 7.6. **Инструкция работы с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память	По тексту программы (табл. 7.7)	Следить за кодами команд
3	Перевести машину в автоматический режим	F АВТ	
4	Занести нули в регистры 7, 8, 9, a, b, c	Сх x → П 7 x → П 8 ...... x → П c	
5	Ввести в машину среднюю арифметическую $\bar{x}$ в регистр 0	Набрать $\bar{x}$ x → П 0	
6	Ввести в программу очередную пару чисел: m_i и x_i	Набрать m_i x → П 1 Набрать x_i	
7	Запустить программу с шага 00	В/О С/П	
8	Списать значение строки в табл. 7.3	П → x 2 П → x 3 П → x 4 П → x 5	На табло: $(x_i-\bar{x})m_i$ На табло: $(x_i-\bar{x})^2m_i$ На табло: $(x_i-\bar{x})^3m_i$ На табло: $(x_i-\bar{x})^4m_i$
9	Если заполнены не все строки табл. 7.3, то перейти к п. 5, иначе – к п. 10		

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
10	Списать строку сумм	П → x 8 П → x 9 П → x a П → x b П → x c	На табло: Σm_i На табло: $\Sigma(x_i-\bar{x})m_i=0$ На табло: $\Sigma(x_i-\bar{x})^2 m_i$ На табло: $\Sigma(x_i-\bar{x})^3 m_i$ На табло: $\Sigma(x_i-\bar{x})^4 m_i$
11	Запустить программу с шага 34	С/П	
12	После окончания счета списать результаты	П → x 1 П → x 2 П → x 3 П → x 4 П → x 5 П → x 6	На табло — s На табло — s^2 На табло — $\hat{\mu}_3$ На табло — $\hat{\mu}_4$ На табло — $\hat{A}c$ На табло — $\hat{E}k$
13	Если содержимое регистра 9 не равно 0, т.е. $\Sigma(x_i-\bar{x})m_i \neq 0$, то имеет место ошибка: а) при вычислении $\bar{x}$; б) при наборе программы; в) при наборе чисел x_i, m_i	б) F ПРГ $\overrightarrow{\text{ШГ}}$ $\overrightarrow{\text{ШГ}}$ ... $\overrightarrow{\text{ШГ}}$	а) проверить правильность вычисления $\bar{x}$ б) сличать высвечиваемые коды в тексте программы (табл. 7.7) в) повторно ввести данные x_i, m_i, провести расчеты по программе
14	При обнаружении несоответствия записанного кода команды коду графы "Код" в тексте программы (табл. 7.7) исправить этот код	$\overleftarrow{\text{ШГ}}$ Клавиши соответствующего шага табл. 7.7	

Таблица 7.7. **Текст программы вычисления выборочных характеристик s, s^2, $\hat{\mu}_3$, $\hat{\mu}_4$, $\hat{Ac}$, $\hat{Ek}$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	П→x 0	60	30	П→x 8	68
01	−	11	31	+	10
02	B↑	0E	32	x→П 8	48
03	B↑	0E	33	С/П	50
04	B↑	0E	34	П→x a	6−
05	П→x 1	61	35	П→x 8	68
06	x	12	36	:	13
07	x→П 2	42	37	x→П 2	42
08	x	12	38	F √	21
09	x→П 3	43	39	x→П 1	41
10	x	12	40	П→x b	6L
11	x→П 4	44	41	П→x 8	68
12	x	12	42	:	13
13	x→П 5	45	43	x→П 3	43
14	П→x c	6C	44	П→x 2	62
15	+	10	45	:	13
16	x→П c	4C	46	П→x 1	61
17	П→x 4	64	47	:	13
18	П→x b	6L	48	x→П 5	45
19	+	10	49	П→x c	6C
20	x→П b	4L	50	П→x 8	68
21	П→x 3	63	51	:	13
22	П→x a	6−	52	x→П 4	44
23	+	10	53	П→x 2	62
24	x→П a	4−	54	F x^2	22
25	П→x 2	62	55	:	13
26	П→x 9	69	56	3	03
27	+	10	57	−	11
28	x→П 9	49	58	x→П 6	46
29	П→x 1	61	59	С/П	50

7.4.4. Работа с системой программ

Работа с программами заключается в следующем:

1. Заготавливается табл. 7.3 и в ней заполняются столбцы x_i и m_i.
2. По программе 7.4.2 вычисляется средняя арифметическая $\bar{x}$.
3. Выполняется инструкция работы с программой 7.4.3 (табл. 7.5).
4. После окончания вычислений по программе (табл. 7.6) в адресуемых регистрах памяти *сохраняются значения величин,* которые можно использовать для дальнейшего вычисления "вручную". На этом работа с микроЭВМ *по программе заканчивается.*

Содержимое адресуемых регистров памяти после останова программы на шаге 33 и шаге 59 приведено в табл. 7.8.

Таблица 7.8. **Содержимое регистров памяти машины в моменты останова программы вычисления выборочных характеристик**

Регистр	Шаг	
	33	59
X	$\sum_i m_i$	$\hat{Ek}$
0	$\bar{x}$	$\bar{x}$
1	m_i	s
2	$(x_i-\bar{x})m_i$	s^2
3	$(x_i-\bar{x})^2 m_i$	$\hat{\mu}_3$
4	$(x_i-\bar{x})^3 m_i$	$\hat{\mu}_4$
5	$(x_i-\bar{x})^4 m_i$	$\hat{Ac}$
6		$\hat{Ek}$
7		
8	Σm_i	Σm_i
9	$\Sigma(x_i-\bar{x})m_i=0$	$\Sigma(x_i-\bar{x})m_i=0$
a	$\Sigma(x_i-\bar{x})^2 m_i$	$\Sigma(x_i-\bar{x})^2 m_i$
b	$\Sigma(x_i-\bar{x})^3 m_i$	$\Sigma(x_i-\bar{x})^3 m_i$
c	$\Sigma(x_i-\bar{x})^4 m_i$	$\Sigma(x_i-\bar{x})^4 m_i$
d		

При "ручных" вычислениях содержимое адресуемых регистров памяти *не затрагивается*. Поэтому в случае ошибки (нажатия не на те клавиши) имеется возможность повторного счета.

7.4.5. Пример вычисления выборочных характеристик

С учетом данных табл. 7.2 найдем по данным примера значения x_i, m_i и занесем их во вспомогательную таблицу (табл. 7.9) для вычисления выборочных характеристик. По программе 7.4.2 найдем среднюю арифметическую $\bar{x}$ = 5,4656 млн. руб.

Таблица 7.9. **Вспомогательная таблица для вычисления выборочных характеристик**

x_i	m_i	$x_i m_i$	$\Delta_i m_i$	$\Delta_i^2 m_i$	$\Delta_i^3 m_i$	$\Delta_i^4 m_i$
1	2	3	4	5	6	7
5,03	2	10,06	−0,8712	0,37949	−0,1653	0,07201
5,14	3	15,42	−0,9768	0,31805	−0,10356	0,03372

1	2	3	4	5	6	7
5,25	12	63,00	−2,5872	0,55780	−0,12026	0,025928
5,36	19	101,84	−2,0064	0,21188	−0,02237	0,00236
5,47	29	158,63	· 0,1276	0,00056	0,00000	0,00000
5,58	18	100,44	2,0592	0,23557	0,02695	0,00308
5,69	13	73,97	2,9172	0,65462	0,14690	0,03296
5,80	4	23,20	1,3376	0,44729	0,14957	0,05002
Итого	100	546,56	0	2,80526	−0,08808	0,22008

Для проверки правильности вычисления $\bar{x}$ и ввода в микрокалькулятор значений x_i, m_i рассчитывают:

$$\sum_{i=1}^{l} \Delta_i m_i = \sum_{i=1}^{l} (x_i - \bar{x}) m_i = 0.$$

В примере тождество выполняется. В итоговой строке столбца имеем 0.

По данным примера $\hat{\mu}_2 = 0{,}028$, $\hat{\mu}_3 = -0{,}00088$, $\hat{\mu}_4 = 0{,}0022$.

Выборочная дисперсия s^2 равна центральному моменту второго порядка:

$$s^2 = \frac{\Sigma(x_i - \bar{x})^2 m_i}{\Sigma m_i} = \hat{\mu}_2.$$

По данным примера $s^2 = 0{,}028$, а выборочное среднее квадратическое отклонение $s = \sqrt{s^2} = 0{,}167$ млн. руб.; $\hat{Ac} = -0{,}188$; $\hat{Ek} = -0{,}203$.

Медиана $\hat{Me}$ – значение признака $x_{(l)}$, приходящееся на середину ранжированного ряда наблюдений ($n = 2$, $l - 1$). При четном числе наблюдений ($n = 2l$) медианой $\hat{Me}$ является средняя арифметическая двух значений, расположенных в середине ранжированного ряда:

$$\hat{Me} = \frac{x_{(l)} + x_{(l+1)}}{2}.$$

Если ранжировать значения, попавшие в медианный интервал [5,41; 5,52], – интервал, в котором накопленная частота $m_н$ впервые превышает половину объема выборки $n/2 = 50$, – до значений $x_{(50)}$ и $x_{(51)}$, получим

$$x_{(37)} = x_{(38)} = \ldots = x_{(50)} = x_{(51)} = 5{,}43.$$

Следовательно, $\hat{Me} = \frac{5{,}43 + 5{,}43}{2}$ (млн. руб.).

Если исходить из интервального ряда, то медиану следует вычислять по формуле

$$\hat{Me} = a_{Me} + h\frac{\frac{n}{2} - m_{H(Me-1)}}{m_{Me}},$$

где $_{Me}$ означает номер медианного интервала, $_{(Me-1)}$ – интервала, предшествующего медианному.

В нашем примере $\hat{Me} = 5{,}41 + \frac{50-36}{29}0{,}11 = 5{,}41 + 0{,}0531 = 5{,}46$ млн. руб.

Мода $\hat{Mo}$ для совокупности наблюдений равна тому значению признака (табл. 7.1), которому соответствует наибольшая частота.

Поскольку вариант 5,43 имеет наибольшую частоту (m = 15), это означает, что $\hat{Mo} = 5{,}43$ млн. руб.

Для одномодального интервального ряда моду можно вычислять по формуле

$$\hat{Mo} = a_{Mo} + h\frac{m_{Mo} - m_{(Mo-1)}}{2m_{Mo} - m_{(Mo-1)} - m_{(Mo+1)}},$$

где $_{Mo}$ означает номер модального интервала (интервала с наибольшей частотой), $_{Mo-1}$ и $_{Mo+1}$ – номера предшествующего модальному и следующего за ним интервалов.

В примере

$$\hat{Mo} = 5{,}41 + \frac{29-19}{2\cdot 29-19-18}0{,}11 = 5{,}41 + 0{,}0524 = 5{,}46 \text{ млн. руб.}$$

Так как по величине $\bar{x}$, $\hat{Mo}$ и $\hat{Me}$ почти не отличаются друг от друга, есть основания предполагать теоретическое распределение нормальным.

При расчете $\hat{Mo}$ и $\hat{Me}$ по интервальному ряду целесообразно воспользоваться микрокалькулятором (в автоматическом режиме, без программы). Коэффициент вариации $V_s = \frac{s}{\bar{x}} \cdot 100\,\% = 3{,}06\,\%$.

7.5. ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ ВАРИАЦИОННЫХ РЯДОВ

Для визуального подбора теоретического распределения, а также выявления положения среднего значения ($\bar{x}$) и характера рассеивания (s^2 и s) вариационные ряды изображают графически.

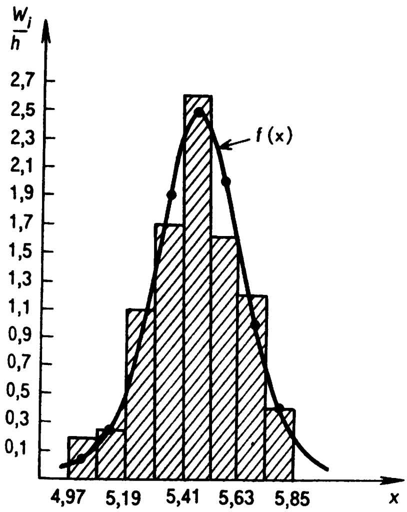
Для изображения как дискретных, так и интервальных рядов применяются полигон и кумулята, для изображения только интервальных рядов – гистограмма. Для построения этих графиков запишем вариационные ряды распределения (интервальный и дискретный) относительных частот (частостей) $W_i = m_i/n$, накопленных относительных частот $W_{нi}$ и найдем отношение W_i/h, заполнив табл. 7.10.

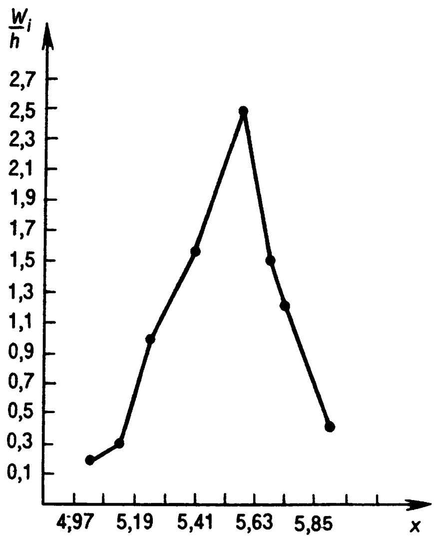
Таблица 7.10. **Статистический ряд распределения объемов основных фондов 100 предприятий**

Интервалы $a_i - b_i$	x_i	W_i	$W_{нi}$	W_i/h
4,97–5,08	5,03	0,02	0,02	0,18
5,08–5,19	5,14	0,03	0,05	0,27
5,19–5,30	5,25	0,12	0,17	1,09
5,30–5,41	5,36	0,19	0,36	1,73
5,41–5,52	5,47	0,29	0,65	2,64
5,52–5,63	5,58	0,18	0,83	1,64
5,63–5,74	5,69	0,13	0,96	1,18
5,74–5,85	5,80	0,04	1,00	0,36
	–	1,00	–	–

Для построения гистограммы относительных частот (частостей) по оси абсцисс откладываем частичные интервалы, на каждом из которых строим прямоугольник, площадь которого равна относительной частоте W_i данного i-го интервала. Тогда высота элементарного прямоугольника должна быть равна W_i/h; в нашем примере $h = 0{,}11$ (рис. 7.1). Следовательно, площадь под гистограммой равна сумме всех относительных частот, т.е. единице.

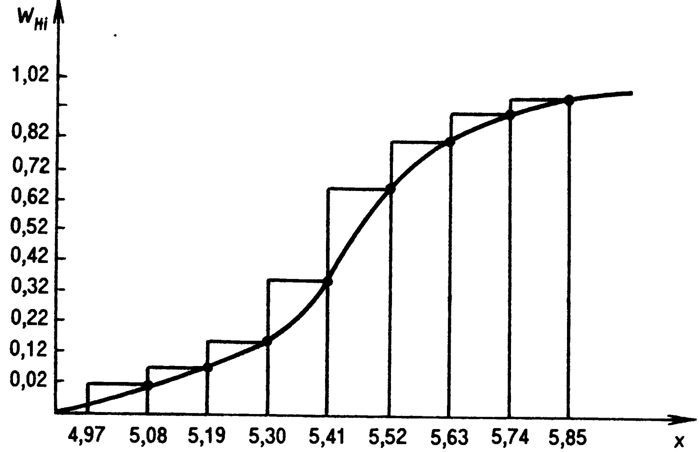
Из гистограммы можно получить полигон того же распределения, если середины верхних оснований прямоугольников соединить отрезками прямой (рис. 7.2).





Гистограмма и полигон являются аппроксимациями кривой плотности (дифференциальной функции) теоретического распределения (генеральной совокупности). Поэтому по их виду можно судить о гипотетическом законе распределения.

Для построения кумуляты дискретного ряда по оси абсцисс откладывают значения признака x_i, а по оси ординат – накопленные относительные частоты $W_{нi}$. Для интервального ряда по оси абсцисс откладывают интервалы (рис. 7.3).



С кумулятой сопоставляется график интегральной функции распределения $F(x)$.

В нашем примере коэффициенты асимметрии и эксцесса не намного отличаются от нуля. Коэффициент асимметрии оказался отрицательным ($\widehat{Ac}$ = – 0,188), что свидетельствует о небольшой левосторонней асимметрии данного распределения. Эксцесс оказался также отрицательным ($\widehat{Ek}$ = – 0,203). Это говорит о том, что кривая, изображающая ряд распределения, по сравнению с нормальной имеет несколько более плоскую вершину. Гистограмма и полигон напоминают кривую нормального распределения (рис. 7.1 и 7.2). Все это дает возможность выдвинуть гипотезу о том, что распределение объемов фондов является нормальным.

7.6. РАСЧЕТ ТЕОРЕТИЧЕСКОЙ КРИВОЙ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Приведем один из способов расчета теоретического нормального распределения по двум найденным выборочным характеристикам x и s эмпирического ряда.

При расчете теоретических частот m_i^T за оценку математического ожидания μ и среднего квадратического отклонения σ нормального закона распределения принимают значения соответствующих выборочных характеристик $\bar{x}$ и s, т.е. $\mu = \bar{x} = 5{,}466$; $\sigma = s = 0{,}167$.

Теоретические частоты находят по формуле

$$m_i^{\text{T}} = np_i,$$

где n – объем; p_i – вероятность попадания значения нормально распределенной случайной величины в i-интервал.

Вероятность p_i определяется по формуле

$$p_i = p(a_i < x \leqslant b_i) = \frac{1}{2}\left[\Phi(t_{2i}) - \Phi(t_{1i})\right],$$

где $\Phi(t) = \frac{2}{\sqrt{2\pi}} \int\limits_0^t e^{-x^2/2}\,dx$ – интегральная функция Лапласа, находится по таблице для

$$t_{2i} = \frac{b_i - \bar{x}}{s}, \quad t_{1i} = \frac{a_i - \bar{x}}{s}.$$

Для вычисления вероятности p_i и теоретических частот m_i^{T} составим табл. 7.11.

Таблица 7.11. **Расчет теоретической кривой нормального распределения**

Интервалы $a_i - b_i$	m_i	t_{1i}	t_{2i}	$1/2\Phi(t_{1i})$	$1/2\Phi(t_{2i})$	p_i	np_i	m_i^{T}	W_i^{T}/h
1	2	3	4	5	6	7	8	9	10
4,97–5,08	2	$-\infty$	–2,39	–0,5	–0,4913	0,0087	0,87	1	0,09
5,08–5,19	3	–2,39	–1,71	–0,4913	–0,4564	0,0349	3,49	3	0,27
5,19–5,30	12	–1,71	–1,03	–0,4564	–0,3485	0,1079	10,9	11	1,00
5,30–5,41	19	–1,03	–1,35	–0,3485	–0,1368	0,2117	21,17	21	1,91
5,41–5,52	29	–0,35	0,33	–0,1368	0,1293	0,2661	26,61	27	3,45
5,52–5,63	18	0,33	1,02	0,1233	0,3461	0,2228	22,28	22	2,00
5,63–5,74	13	1,02	1,70	0,3461	0,4554	0,1093	10,98	11	1,00
5,74–5,85	4	1,70	$+\infty$	0,4554	0,5000	0,0446	4,46	4	0,36
	100	–	–	–	–	1,000	–	100	–

Примечание. При использовании критерия согласия Пирсона общее число наблюдений должно быть достаточно большим ($n \geqslant 50$) и интервалы должны быть достаточно заполнены частотами. Если отдельные теоретические частоты на концах распределения окажутся малыми ($m_i^{\text{T}} < 5$), то при вычислении $\chi^2_{\text{набл}}$ необходимо объединить такие интервалы, сложив соответствующие частоты.

Построим теоретическую нормальную кривую $f(x)$ на рис. 7.1. Для этого из середины частных интервалов восстановим перпендикуляры высотой W_i^T/h (табл. 7.11, гр. 10), где $W_i^T = m_i^T/n$. На рис. 7.1 концы этих перпендикуляров отмечены точками. Полученные точки соединены плавной кривой.

Сравнение гистограммы и нормальной кривой наглядно показывает согласованность между теоретическим и эмпирическим распределениями.

7.7. ПРОВЕРКА ГИПОТЕЗЫ О НОРМАЛЬНОМ ЗАКОНЕ РАСПРЕДЕЛЕНИЯ

Часто для проверки соответствия эмпирического ряда распределения нормальному закону используют критерий χ^2, основанный на сравнении эмпирических частот m_i с теоретическими m_i^T, которые можно ожидать при принятии определенной нулевой гипотезы.

Значение $\chi^2_{набл}$ – наблюдаемое значение критерия, полученное по результатам наблюдений, равно

$$\chi^2_{набл} = \sum_{i=1}^{k} \frac{(m_i - m_i^T)^2}{m_i^T},$$

где k – число интервалов (после объединения), m_i^T – теоретические частоты. Все вспомогательные расчеты, необходимые для вычисления χ^2, сведем в табл. 7.12.

Таблица 7.12. **Вычисление критерия χ^2 при проверке нормальности распределения объемов основных фондов**

Интервалы $a_i - b_i$	m_i	m_i^T	$(m_i - m_i^T)^2$	$\frac{(m_i - m_i^T)^2}{m_i^T}$
4,97 – 5,08	2	1		
5,08 – 5,19	3 } 17	3 } 15	4	0,267
5,19 – 5,30	12	11		
5,30 – 5,41	19	21	4	0,190
5,41 – 5,52	29	27	4	0,148
5,52 – 5,63	18	22	16	0,727
5,63 – 5,74	13	11		
5,74 – 5,85	4	4	4	0,267
	100	100	–	$\chi^2_{набл} = 1,599$

Правило проверки гипотезы заключается в следующем. Определяем по таблице распределения χ^2 (Хи-квадрат) критическое значение $\chi^2_{кр}(\alpha, \nu)$ для числа степеней свободы $\nu = k - 3$ и заданного уровня значимости α. Затем сравниваем $\chi^2_{набл}$ и $\chi^2_{кр}$.

Если $\chi^2_{набл} \leqslant \chi^2_{кр}$, то выдвинутая гипотеза о законе распределения *не отвергается* (не противоречит опытным данным).

Если $\chi^2_{набл} > \chi^2_{кр}$, то выдвинутая гипотеза о нормальном законе распределения *отвергается* с вероятностью ошибки α.

Для нашего примера $\chi^2_{набл} = 1{,}599$, $\alpha = 0{,}05$, $\nu = 5 - 3 = 2$ (число интервалов после объединения стало равным 5) и $\chi^2_{кр}\ (0{,}05;\ 2) = 6{,}000$.

Так как $\chi^2_{набл} < \chi^2_{кр}$, то, согласно критерию Пирсона, гипотеза о нормальном законе *не отвергается*. Можно сделать вывод, что распределение объемов основных фондов 100 предприятий является нормальным.

8. МОДЕЛИРОВАНИЕ РЯДОВ ДИНАМИКИ С ИСПОЛЬЗОВАНИЕМ ГАРМОНИЧЕСКОГО АНАЛИЗА

8.1. АЛГОРИТМ МОДЕЛИРОВАНИЯ ВРЕМЕННОГО РЯДА С УЧЕТОМ ТРЕНДА И СЕЗОННЫХ КОЛЕБАНИЙ

Пусть рассматриваемый временной ряд $x_1, x_2, \ldots, x_t, \ldots, x_T$ содержит три компоненты: тренд U_t, сезонные колебания V_t и случайную составляющую ε_t, т.е.

$$X_t = U_t + V_t + \varepsilon_t, \quad t = 1, 2, \ldots, T.$$

Оценку уравнения тренда получим с помощью метода наименьших квадратов из функций вида: $\hat{U}_t = b_0 + b_1 t$; $\hat{U}_t = b_0 + b_1 t^2$; $\hat{U}_t = b_0 e^{b_1 t}$ и $\hat{U}_t = b_0 + b_1 1/t$.

Вид уравнения тренда выбирается прежде всего на основании качественного экономического анализа рассматриваемого процесса. При выборе тренда руководствуются также соображениями адекватности получаемой модели, оцениваемой по величине остаточной дисперсии s_u^2 и средней относительной ошибки аппроксимации $\overline{\delta u}$, где

$$s_u^2 = \frac{1}{T} \sum_{t=1}^{T} (x_t - \hat{u}_t^2), \quad \delta_{u_t} = \frac{|x_t - \hat{u}_t|}{x_t},$$

$$\overline{\delta}_u = \frac{1}{T} \sum_{t=1}^{T} \frac{|x_t - \hat{u}_t|}{x_t} = \frac{1}{T} \sum_{t=1}^{T} |\delta_{u_t}| \quad .$$

Считается, что лучшими аппроксимирующими свойствами обладает функция, которой соответствуют наименьшие значения s_u^2 или $\overline{\delta}_u$.

Рассмотрим задачу сглаживания сезонных колебаний исходя из ряда $v_t = x_t - \hat{u}_t$, где x_t – значение исходного временного ряда в момент времени t, а $\hat{u}_t$ – оценка соответствующего значения тренда ($t = 1, 2, \ldots, T$).

Так как сезонные колебания представляют собой циклический, повторяющийся во времени процесс, то в качестве сглаживающих функций используется гармонический ряд (ряд Фурье) вида

$$v_t = \sum_{i=1}^{n} \alpha_i \cos w_i t + \sum_{i=1}^{n} \beta_i \sin w_i t.$$

Оценки параметров α_i и β_i модели определяют из выражений:

$$\hat{\alpha}_i = \begin{cases} (2/T) \sum_{t=1}^{T} x_t \cos w_i t & \text{для } i = 1, 2, \dots, n-1, \\ (1/T) \sum_{t=1}^{T} x_t \cos w_i t & \text{для } i = 0, n; \end{cases}$$

$$\hat{\beta}_i = \begin{cases} (2/T) \sum_{t=1}^{T} x_t \sin w_i t & \text{для } i = 1, 2, \dots, n-1, \\ (1/T) \sum_{t=1}^{T} x_t \sin w_i t & \text{для } i = 0, n, \end{cases}$$

где $n = T/2$ – максимально допустимое число гармоник; $w_i = \frac{2\pi i}{T}$ – угловая частота i-й гармоники ($i = 1, 2, \dots, m$).

Пусть m – число гармоник, используемых для сглаживания сезонных колебаний ($m < n$). Тогда оценка гармонического ряда имеет вид

$$\hat{v}_t = \sum_{i=0}^{m} \hat{\alpha}_i \cos w_i t + \sum_{i=0}^{m} \hat{\beta}_i \sin w_i t,$$

а расчетные значения временного ряда исходного показателя определяются по формуле

$$\hat{x}_t = \hat{u}_t + \hat{v}_t.$$

Качество полученной аппроксимации исходного временного ряда следует оценивать по величине остаточной дисперсии

$$s_x^2 = \frac{1}{T} \sum_{t=1}^{T} (x_t - \hat{x}_t)^2$$

и средней относительной ошибки

$$\hat{\delta}_x = \frac{1}{T} \sum_{t=1}^{T} \frac{|x_t - \hat{x}_t|}{x_t} = \frac{1}{T} \sum_{t=1}^{T} |\delta_{x_t}|.$$

О качестве сглаживания исходного временного ряда в отдельных точках можно судить по величине индивидуальных относительных отклонений $\delta_{x_t} = |x_t - \hat{x}_t| / x_t$.

8.2.1. Назначение программ

Система программ включает в себя четыре программы, которые позволяют получить следующие результаты:

оценку уравнения тренда $\hat{u}_t$ из класса линейных, квадратичных, гиперболических и экспоненциальных функций (программа 8.2.2);

расчетные значения уравнения тренда, гармонический ряд V_t, индивидуальные (δ_t) и среднюю ($\bar{\delta}$) относительные ошибки аппроксимации, оценку остаточной дисперсии (программа 8.2.3);

оценки параметров α_i и β_i ($i = 1, 2, 3, 4$) гармонического ряда V_t (программа 8.2.4);

расчетные значения гармонического ряда $\hat{V}_t$ (программа 8.2.5).

8.2.2. Программа вычисления оценок параметров тренда b_0 и b_1

Таблица 8.1. **Инструкция работы с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Набрать программу с шага 00 по шаг 70, с шага 71 для уравнения типа $U_t = b_0 + b_1 t$ набрать подпрограммы 1 и 3, для уравнения типа $U_t = b_0 + b_1 t^2$ — подпрограммы 1 и 4, для уравнения типа $U_t = b_0 + b_1/t$ — подпрограммы 1 и 5, для уравнения типа $U_t = b_0 e^{b_1 t}$ — подпрограммы 2 и 3	По тексту программы (табл. 8.2)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем — правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести 1 в регистр памяти a, число наблюдений T — в регистр памяти 7, а остальные регистры обнулить	1 x→П a T x→П 7 Cx x→П 1 x→П 2 ...... x→П 6 x→П 8 x→П 9 x→П b x→П c x→П d	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
5	Набрать первый уровень исходного ряда x_t и запустить программу с шага 00	Набрать x_1 В/О С/П	
6	После останова машины набрать следующий уровень ряда x_t и запустить программу и т.д.	Набрать x_2 С/П x_3 С/П ...... x_t С/П ...... x_T С/П	
7	Запустить программу с шага 30	БП 30 С/П	
8	После останова машины выписать из регистров памяти 1 и 2 соответственно b_0 и b_1	П → x 1 П → x 2	На табло b_0 b_1

Примечание. В п. 8 для уравнения регрессии типа $U_t = b_0 e^{b_1 t}$ в регистре памяти 1 находится значение $\ln b_0$. Чтобы найти b_0, необходимо набрать следующие команды:

Нажимаемые клавиши	На табло
П → x 1	$\ln b_0$
F e^x	b_0

Таблица 8.2. **Текст программы вычисления оценок параметров тренда b_0 и b_1**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x → П 1	41	11	x → П c	4*C*
01	ПП	53	12	П → x b	6*L*
02	71	71	13	F x^2	22
03	П → x 2	62	14	П → x d	6*Г*
04	+	10	15	+	10
05	x → П 2	42	16	x → П d	4*Г*
06	П → x a	6–	17	П → x 1	61
07	ПП	53	18	П → x b	6*L*
08	75	75	19	x	12
09	П → x c	6*C*	20	П → x 0	60
10	+	10	21	+	10

Шаг	Клавиши	Код
22	x → П 0	40
23	П → x a	6–
24	1	01
25	+	10
26	x → П a	4–
27	С/П	50
28	Б/П	51
29	00	00
30	П → x 7	67
31	x → П 3	43
32	П → x c	6С
33	x → П 4	44
34	x → П 5	45
35	П → x d	6Г
36	x → П 6	46
37	ПП	53
38	63	63
39	x → П 9	49
40	П → x 2	62
41	x → П 3	43
42	П → x 0	60
43	x → П 4	44
44	ПП	53
45	63	63
46	П → x 9	69
47	:	13
48	x → П 1	41
49	П → x 7	67
50	x → П 3	43
51	П → x c	6С
52	x → П 4	44
53	П → x 2	62
54	x → П 5	45
55	П → x 0	60
56	x → П 6	46
57	ПП	53
58	63	63

Шаг	Клавиши	Код
59	П → x 9	69
60	:	13
61	x → П 2	42
62	С/П	50
63	П → x 3	63
64	П → x 6	66
65	x	12
66	П → x 4	64
67	П → x 5	65
68	x	12
69	–	11
70	В/О	52
	Подпрограмма 1	
71	1	01
72	x	12
73	x → П 1	41
74	В/О	52
	Подпрограмма 2	
71	F ln	18
72	x → П 1	41
73	x → П 1	41
74	В/О	52
	Подпрограмма 3	
75	1	01
76	x	12
77	x → П b	4L
78	В/О	52
	Подпрограмма 4	
75	F x^2	22
76	x → П b	4L
77	x → П b	4L
78	В/О	52
	Подпрограмма 5	
75	F x^{-1}	23
76	x → П b	4L
77	x → П b	4L
78	В/О	52

8.2.3. Программа вычисления выборочных характеристик $\hat{u}_t$, v_t, s_u^2, δ_{u_t} и $\bar{\delta}_u$

Таблица 8.3. **Инструкция работы с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
2	Набрать программу с шага 00 по 43 шаг, затем для уравнения типа $U_t = b_0 + b_1 t$ набрать подпрограмму 1, для уравнения типа $U_t = b_0 + b_1 t^2$ – подпрограмму 2, для уравнения типа $U_t = b_0 + b_1/t$ – подпрограмму 3, для уравнения типа $U_t = b_0 e^{b_1 t}$ – подпрограмму 4	По тексту программы (табл. 8.4)	Следить за кодами команд. При ошибке набрать ШГ, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 1 и 2 соответственно b_0 и b_1	Набрать b_0 х → П 1 b_1 х → П 2	
5	Занести 100 в регистр памяти 8, а остальные регистры, кроме 1 и 2, обнулить	100 х → П 8 Сх х → П 0 х → П 3 .... х → П 7 х → П 9 х → П а .... х → П d	
6	Набрать первое значение исходного ряда x_t и запустить программу с шага 00 После останова машины списать с табло значение δ_{u_1}, из регистра памяти b – расчетное значение $\hat{u}_1$, а из регистра 9 – значение v_1	Набрать x_1 В/О С/П П → х b П → х 9	На табло δ_{u_1} На табло $\hat{u}_1$ v_1
7	Набрать t-е значение исходного ряда x_t и запустить программу После останова машины списать с табло значение δu_t, а из регистров памяти b и 9 выписать соответственно $\hat{u}_t$ и v_t	Набрать x_t С/П П → х b П → х 9	На табло $\hat{u}_t$ v_t
8	Рассчитав последние T-е значения δ_T, $\hat{u}_T$ и v_T, запустить программу с шага 32 После останова выписать из регистров памяти 5 и 6 соответственно s_u^2 и $\bar{\delta}_u$	БП 32 С/П П → х 5 П → х 6	На табло s_u^2 $\bar{\delta}_u$

Примечание. П. 4 выполняется, если после проведения расчетов по программе 8.2.2 микрокалькулятор выключался или в регистрах памяти 1 и 2 не сохранены результаты расчетов по этой программе.

Таблица 8.4. Текст программы вычисления $\hat{u}_t$, $\hat{v}_t$, s^2, δ_t и $\bar{\delta}$

Шаг	Клавиши	Код
00	x → П a	4–
01	П → x 7	67
02	1	01
03	+	10
04	x → П 7	47
05	ПП	53
06	44	44
07	x → П b	4L
08	П → x a	6–
09	П → x b	6L
10	–	11
11	x → П 9	49
12	F x^2	22
13	П → x 0	60
14	+	10
15	x → П 0	40
16	П → x 9	69
17	F x^2	22
18	F $\sqrt{}$	21
19	П → x a	6–
20	:	13
21	П → x 8	68
22	x	12
23	x → П 3	43
24	F x^2	22
25	F $\sqrt{}$	21
26	П → x 4	64
27	+	10
28	x → П 4	44
29	П → x 3	63
30	С/П	50
31	БП	51
32	00	00
33	П → x 7	67
34	F 1/x	23
35	П → x 0	60
36	x	12
37	x → П 5	45
38	П → x 7	67
39	F 1/x	23
40	П → x 4	64
41	x	12
42	x → П 6	46
43	С/П	50
	Подпрограмма 1	
44	П → x 2	62
45	x	12
46	П → x 1	61
47	+	10
48	В/О	52
	Подпрограмма 2	
44	F x^2	22
45	П → x 2	62
46	x	12
47	П → x 1	61
48	+	10
49	В/О	52
	Подпрограмма 3	
44	F 1/x	23
45	П → x 2	62
46	x	12
47	П → x 1	61
48	+	10
49	В/О	52
	Подпрограмма 4	
44	П → x 2	62
45	x	12
46	F e^x	16
47	П → x 1	61
48	+	10
49	В/О	52

8.2.4. Программа вычисления оценок параметров α_i и β_i гармонического ряда V_t

Таблица 8.5. **Инструкция работы с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память. Расчеты проводить в радианах, для чего переключатель перевести в положение "Р"	По тексту программы (табл. 8.6)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем – правильную команду
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти a, b, 0, d соответственно значения $2/T$, $2\pi/T$, 1, 1, а остальные регистры обнулить	Набрать $2/T$ x → П a $2\pi/T$ x → П b 1 x → П 0 x → П d Cx x → П 1 x → П 2 ...... x → П 9 x → П c	Порядок набора $2\pi/T$ 2 F π x T :
5	Ввести в машину первое значение исходного временного ряда V_t и запустить программу с шага 00 После останова машины набрать второе наблюдение ряда – v_2 и запустить программу и т. д. Номер наблюдения (v_t), которое нужно набрать после останова, указан на табло	Набрать v_1 В/О С/П v_2 С/П v_3 С/П .... v_t С/П .... v_T С/П	После запуска каждого значения машина считает 1, 2 мин
6	Выписать из регистров памяти значения оценок параметров α_i и β_i	П → x 1 П → x 2 П → x 3 П → x 4 П → x 5 П → x 6 П → x 7 П → x 8 П → x b	На табло α_1 α_2 α_3 α_4 β_1 β_2 β_3 β_4 $2\pi/T$

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
7	Если программа не работает: 1) перейти в режим программирования 2) проверить правильность набора программы 3) при обнаружении несоответствия кода команды на табло тексту программы 4) при ошибочном наборе номера подпрограммы, например 64, необходимо вернуться на два шага назад и набрать правильный номер подпрограммы 5) проконтролировать правильность ввода исходных данных, возвратившись к п. 4 и п. 5 инструкции	F ПРГ $\overrightarrow{ШГ}$ $\overrightarrow{ШГ}$ . . . $\overrightarrow{ШГ}$ $\overleftarrow{ШГ}$ $\overleftarrow{ШГ}$ $\overleftarrow{ШГ}$ ПП 64	 Сличать высвечиваемые коды с текстом программы (табл 8.6) Набрать правильную команду

Примечание. Если тренд выражен средним значением $\hat{u}_t = \bar{x}$, то в п. 5 вводятся значения исходного временного ряда X_t.

Таблица 8.6. **Текст программы вычисления оценок параметров α_i и β_i**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	x → П 9	49	17	ПП	53
01	П → x 1	61	18	64	64
02	x → П c	4C	19	x → П 3	43
03	ПП	53	20	ПП	53
04	64	64	21	90	90
05	x → П 1	41	22	П → x 4	64
06	ПП	53	23	x → П c	4C
07	90	90	24	ПП	53
08	П → x 2	62	25	64	64
09	x → П c	4C	26	x → П 4	44
10	ПП	53	27	П → x 5	65
11	64	64	28	x → П c	4C
12	x → П 2	42	29	1	01
13	ПП	53	30	x → П d	4Г
14	90	90	31	ПП	53
15	П → x 3	63	32	77	77
16	x → П c	4C	33	x → П 5	45

Продолжение

Шаг	Клавиши	Код
34	ПП	53
35	90	90
36	П → x 6	66
37	x → П c	4C
38	ПП	53
39	77	77
40	x → П 6	46
41	ПП	53
42	90	90
43	П → x 7	67
44	x → П c	4C
45	ПП	53
46	77	77
47	x → П 7	47
48	ПП	53
49	90	90
50	П → x 8	68
51	x → П c	4C
52	ПП	53
53	77	77
54	x → П 8	48
55	1	01
56	x → П d	4Г
57	П → x 0	60
58	1	01
59	+	10
60	x → П 0	40
61	С/П	50
62	БП	51
63	00	00
64	П → x b	6L
65	П → x d	6Г
66	x	12
67	П → x 0	60
68	x	12
69	F cos	1Г
70	П → x 9	69
71	x	12
72	П → x a	6–
73	x	12
74	П → x c	6C
75	+	10
76	В/О	52
77	П → x b	6L
78	П → x d	6Г
79	x	12
80	П → x 0	60
81	x	12
82	F sin	1C
83	П → x 9	69
84	x	12
85	П → x a	6–
86	x	12
87	П → x c	6C
88	+	10
89	В/О	52
90	П → x d	6Г
91	1	01
92	+	10
93	x → П d	4Г
94	В/О	52

8.2.5. Программа вычисления расчетных значений периодической составляющей $\hat{v}_t$

Таблица 8.7. **Инструкция работы с программой**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу в память. Расчеты проводить в радианах, для чего переключатель должен находиться в положении "р"	По тексту программы (табл. 8.8)	Следить за кодами команд. При ошибке набрать $\overleftarrow{\text{ШГ}}$, а затем — правильную команду

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести в регистры памяти 1–8 соответственно оценки параметров α_i и β_i, а в регистр b – значение $2\pi/T$	Набрать α_1 x → П 1 α_2 x → П 2 α_3 x → П 3 α_4 x → П 4 β_1 x → П 5 β_2 x → П 6 β_3 x → П 7 β_4 x → П 8 $2\pi/T$ x → П b	
5	Занести в регистры памяти 0 и d – 1, а регистры a, c, 9 обнулить	Набрать 1 x → П 0 x → П d Cx x → П a x → П c x → П 9	
6	Запустить программу с шага 00	В/О С/П	До останова машина считает около 1 мин
7	После останова машины списать с табло значение $\hat{v}_1$, рассчитать $\hat{x}_1$	Набрать $\hat{u}_1$ +	На табло $\hat{v}_1$ $\hat{x}_1$
8	Для расчета t-го значения $\hat{v}_t$ занести 0 в регистр памяти c и запустить программу ($t = 2, 3, \ldots$). После останова машины списать с табло t-е значение $\hat{v}_t$, рассчитать $\hat{x}_t$	Cx x → П c С/П Набрать $\hat{u}_t$ +	 На табло $\hat{v}_t$ $\hat{x}_t$
9	Пункт 8 повторять, пока последовательно не будут получены все расчетные значения $\hat{v}_2, \hat{v}_3, \hat{v}_4, \ldots, \hat{v}_T$, а также прогнозные значения $\hat{v}_{T+1}, \hat{v}_{T+2}$ и т.д.		

Примечания 1. П. 4 выполняется, если после проведения расчетов по программе 8.2.2 микрокалькулятор выключался или в регистрах памяти 1–8 и b не сохранены результаты расчетов по этой программе.

2. Если расчеты выполняются с момента времени $t = \tau$, то в п. 5 в регистры памяти 0 и d занести значение τ, а в п. 7 – на табло – $\hat{v}_\tau$, в п. 8 – $\hat{v}_{\tau+1}$.

Таблица 8.8. **Текст программы вычисления расчетных значений $\hat{y}_t$**

Шаг	Клавиши	Код	Шаг	Клавиши	Код
00	ПП	53	39	П → x 8	68
01	52	52	40	ПП	53
02	П → x 1	61	41	70	70
03	ПП	53	42	П → x 0	60
04	63	63	43	1	01
05	ПП	53	44	+	10
06	52	52	45	x → П 0	40
07	П → x 2	62	46	1	01
08	ПП	53	47	x → П d	4Г
09	63	63	48	П → x c	6C
10	ПП	53	49	С/П	50
11	52	52	50	БП	51
12	П → x 3	63	51	00	00
13	ПП	53	52	П → x b	6L
14	63	63	53	П → x d	6Г
15	ПП	53	54	x	12
16	52	52	55	П → x 0	60
17	П → x 4	64	56	x	12
18	ПП	53	57	x → П a	4—
19	63	63	58	П → x d	6Г
20	1	01	59	1	01
21	x → П d	4Г	60	+	10
22	ПП	53	61	x → П d	4Г
23	52	52	62	В/О	52
24	П → x 5	65	63	П → x a	6—
25	ПП	53	64	F cos	1Г
26	70	70	65	x	12
27	ПП	53	66	П → x c	6C
28	52	52	67	+	10
29	П → x 6	66	68	x → П c	4C
30	ПП	53	69	В/О	52
31	70	70	70	П → x a	6—
32	ПП	53	71	F sin	1C
33	52	52	72	x	12
34	П → x 7	67	73	П → x c	6C
35	ПП	53	74	+	10
36	70	70	75	x → П c	4C
37	ПП	53	76	В/О	52
38	52	52			

8.2.6. Работа с системой программ

Работа с системой программ заключается в следующем:

1. Выполнение инструкции к программе 8.2.2 (табл. 8.1).

Входными данными этой программы являются значения исходного временного ряда x_t ($t = 1, 2, \ldots, T$).

Вычисления по программе 8.2.2 проводятся в два этапа. На первом этапе в машину последовательно вводятся все значения ряда x_t и после ввода T-го наблюдения осуществляется запуск программы с шага 30. По окончании вычислений по этой программе (табл. 8.2) в адресуемых регистрах памяти сохраняются оценки параметров уравнения тренда b_0 и b_1.

Если на этапе предварительного анализа было отобрано несколько видов уравнений тренда, то после окончания вычислений по программе 8.2.2 для первого из отобранных трендов необходимо перейти в программируемый режим работы и с шага 71 записать подпрограмму, соответствующую уравнению тренда другого вида. Для этого предварительно требуется выполнять следующие команды:

БП 70 В/О F ПРГ

После чего вычислить оценку этого уравнения, возвратившись к п. 3 инструкции.

2. Выполнение инструкции по работе с программой 8.2.3 (табл. 8.3).

Входными данными программы (табл. 8.4) являются значения исходного ряда X_t и параметры b_0 и b_1 уравнения тренда.

При записи этой программы в память с шага 44 нужно набирать подпрограмму, соответствующую выбранному ранее и рассчитанному по программе уравнению тренда.

Вычисления по программе 8.2.3 проводятся в два этапа. На первом этапе в процессе проведения расчетов после каждого останова машины необходимо выписывать из адресуемых регистров памяти расчетные значения тренда $\hat{u}_t$, исходные значения гармонического ряда v_t и индивидуальные относительные отклонения δ_{u_t} ($t = 1, 2,\dots, T$). Определив все значения рядов $\hat{u}_t$ и v_t, переходим ко второму этапу, для чего запускаем машину с шага 32 и после останова выписываем из адресуемых регистров памяти оценку остаточной дисперсии s_u^2 и среднюю относительную ошибку аппроксимации тренда $\overline{\delta}_u$.

3. Перевод машины в режим работы "Р" – в радианах.

4. Выполнение инструкции по работе с программой 8.2.4 (табл. 8.5).

Входными данными этой программы (табл. 8.6) являются значения гармонического ряда v_t.

После окончания вычислений по программе 8.2.4 в адресуемых регистрах памяти сохраняются оценки параметров α_i и β_i гармонического ряда V_t, которые являются входными данными программы 8.2.5.

5. Выполнение инструкции по работе с программой 8.2.5 (табл. 8.7).

После каждого останова машины необходимо выписывать из регистра памяти с очередное расчетное значение $\hat{v}_t$, а затем регистр обнулить.

6. Расчет "вручную" значения $\hat{x}_t = \hat{u}_t + \hat{v}_t$.

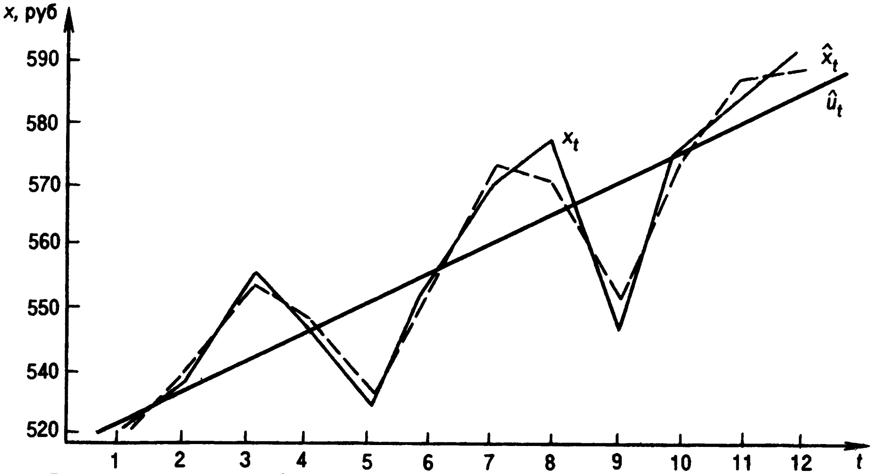
7. Выполнение инструкции для оценки качества построенной модели, приведенной в табл. 8.9.

Таблица 8.9. **Инструкция для оценки качества построенной модели**

№ п/п	Инструкция	Нажимаемые клавиши	Примечание
1	Подготовить машину к записи программы с шага 00	В/О F ПРГ	
2	Записать программу 8.2.3 с шага 00 по шаг 43 в память, заменив на шаге 05 команду ПП на команду x → П 7, а на шаге 06 – команду 44 на команду С/П	По тексту программы (табл. 8.4)	Следить за кодами команд
3	Перевести машину в автоматический режим работы	F АВТ	
4	Занести 100 в регистр памяти 8, а остальные регистры обнулить	Набрать 100 x → П 8 Сx x → П 0 x → П 1 x → П 2 .... x → П 7 x → П 9 .... x → П d	
5	Набрать очередное значение исходного ряда x_t и запустить программу с шага 00 После останова машины набрать расчетное значение данного наблюдения и запустить программу После останова машины списать с табло индивидуальное относительное отклонение	Набрать x_t В/О С/П $\hat{x}_t$ С/П	На табло δ_{x_t}
6	Если введены не все значения рядов X_t и $\hat{X}_t$, то перейти к п. 5, иначе – к п. 7		
7	Запустить программу с шага 32. После останова машины выписать из регистров памяти 5 и 6 оценку остаточной дисперсии s_x^2 и среднюю ошибку аппроксимации $\bar{\delta}_x$	БП 32 С/П П → x 5 П → x 6	На табло s_x^2 $\bar{\delta}_x$

8.3. КОНТРОЛЬНЫЙ ПРИМЕР

Имеются квартальные данные о средней заработной плате одного рабочего производственного объединения за 1983–1985 гг. (табл. 8.10). Графически динамический ряд данного показателя x_t представлен на рис. 8.1. Требуется построить модель ряда с учетом сезонных колебаний.



Решение. Из графического изображения видно, что исходный динамический ряд x_t может быть представлен как сумма постоянной составляющей линейного вида $u_t = \beta_0 + \beta_1 t$ и сезонных колебаний v_t. Определим оценки параметров уравнения тренда $b_0 = 517{,}65151$ и $b_1 = 4{,}7972027$ по программе 8.2.2. По программе 8.2.3 получим расчетные значения тренда $\hat{u}_t$, представленные в табл. 8.10, а также индивидуальные (δ_{u_t}) (табл. 8.10) и среднюю ($\bar{\delta}_u = 1{,}3735943$) ошибки аппроксимации, оценку остаточной дисперсии $s_u^2 = 110{,}39876$. Анализ адекватности модели на основе полученных статистических характеристик позволил остановиться на линейной форме тренда без проведения сопоставлений с другими моделями (параболической и экспоненциальной).

На основе рассчитанных по программе 8.2.3 значений v_t (табл. 8.10) построим гармонический ряд, включающий четыре гармоники ($m = 4$).

В результате расчетов по программе 8.2.4 определим оценки параметров α_i и β_i гармонического ряда:

$\hat{\alpha}_1 = 1{,}585169$; $\hat{\beta}_1 = 2{,}1494528$;
$\hat{\alpha}_2 = 1{,}5361315$; $\hat{\beta}_2 = 4{,}8448953$;
$\hat{\alpha}_3 = 3{,}036134$; $\hat{\beta}_3 = -11{,}369463$;
$\hat{\alpha}_4 = -1{,}6305357$; $\hat{\beta}_4 = 0{,}748942$.

Программа 8.2.5 позволяет получить оценочные значения гармонического ряда ($\hat{v}_t$), которые представлены в табл. 8.10.

Для оценки качества построенной модели с помощью инструкции, приведенной в табл. 8.9, находим индивидуальные относительные отклонения δ_{x_t} (табл. 8.10), а также s_x^2 = 23,064785 и $\bar{\delta}_x$ = 0,6947959. Из табл. 8.10 следует, что относительная ошибка в точках ряда варьируется в пределах от 0,009 до 1,519%. Это свидетельствует об адекватности полученной модели. Учет сезонной составляющей позволил снизить среднюю ошибку аппроксимации с 1,37 до 0,695%, т.е. почти в два раза.

Таблица 8.10. **Исходная информация для анализа и результаты расчетов**

t	x_t	$\hat{u}_t$	δ_{u_t}	v_t	$\hat{v}_t$	$\hat{x}_t$	δ_{x_t}
1	520	522,44871	0,47090575	−2,44871	−2,4942012	519,95451	0,0087480769
2	529	527,24592	0,33158410	1,75408	3,2123336	530,45825	0,2756616000
3	544	532,04312	2,19795560	11,95688	10,3522490	542,39537	0,2949687300
4	537	536,84032	0,029735566	0,15968	0,6050326	537,44535	0,0829329590
5	525	541,63752	3,16905120	−16,63752	−14,9286050	526,70892	0,3255085500
6	546	546,43473	0,79620877	−0,43473	−4,7057085	541,71902	0,7840622700
7	560	551,23193	1,56572670	8,76807	15,3496780	566,58161	1,1752874000
8	568	556,02913	2,10754730	11,97087	3,9764770	560,00561	1,4074630000
9	536	560,82633	4,63177770	−24,82633	−16,6855830	544,14075	1,5187966000
10	565	565,62354	0,11036106	−0,62354	−7,6050283	558,01851	1,2356619000
11	574	570,42074	0,62356444	3,57926	8,4064585	578,82720	0,8409756000
12	582	575,21794	1,16530240	6,78206	4,5269029	579,74484	0,3874845000

ЗАКЛЮЧЕНИЕ

Теперь, уважаемый читатель, вы познакомились с основными методами и алгоритмами многомерного статистического анализа. Программируемый микрокалькулятор избавил вас от трудоемких и довольно однообразных расчетов и позволил сконцентрировать внимание на особенностях применяемых методов и алгоритмов многомерного статистического анализа, содержательной интерпретации получаемых результатов. Однако вы уже убедились, что с помощью программируемых микрокалькуляторов можно решать задачи только ограниченной размерности.

На следующем этапе изучения многомерного статистического анализа и проведения конкретных социально-экономических исследований целесообразно использовать персональные ЭВМ и пакеты стандартных программ по прикладной статистике. ПЭВМ дает возможность применить в процессе исследования комплекс статистических методов анализа зависимостей, снижения размерности и классификации.

Проиллюстрируем возможности многомерного статистического анализа на примере. Расчеты проводились с использованием пакета прикладных программ "АРМ-статистика"[1].

П р и м е р. Исследовать структуру совокупности показателей, характеризующих аптечную службу в 25 регионах. Рассматривалась система пяти показателей:

X_1 – объем реализации медикаментов на одного жителя, руб.;

X_2 – удельный вес городского населения, %;

X_3 – объем товарооборота аптек на одного фармацевта, тыс. руб.;

X_4 – число фармацевтов на 10 тыс. жителей, чел.;

X_5 – оборачиваемость оборотных средств, дн.

Объем реализации медикаментов на одного жителя является одним из основных показателей деятельности аптечной службы региона и поэтому он был выбран в качестве результативного, т.е. $X_1 = Y$.

Первоначально были рассчитаны такие выборочные характеристики показателей, как средняя арифметическая ($\bar{x}$), среднее квадратическое отклонение (s), коэффициент вариации ($V_s = \frac{s}{\bar{x}} \cdot 100\%$), коэффициент асимметрии ($\hat{Ac}$) и эксцесса ($\hat{Ek}$). Значения характеристик приводятся в таблице:

[1] Подробно описание пакета см.: Автоматизированное рабочее место для статистической обработки данных. – М.: Финансы и статистика, 1990. – 189 с.

Показатель	$\bar{x}$	s	V_s	$\hat{Ac}$	$\hat{Ek}$
X_1	19,6	1,05	5,4	0,1	– 0,5
X_2	67,0	7,99	11,9	0,6	– 0,9
X_3	20,8	2,41	11,5	– 0,1	– 1,5
X_4	9,5	1,25	13,2	0,3	– 1,5
X_5	207,9	20,21	9,7	1,1	0,8

Из таблицы следует, что рассматриваемая совокупность наблюдений однородна по всем показателям, так как коэффициент вариации достаточно мал ($V_s < 14\%$). Относительно небольшие значения коэффициентов асимметрии и эксцесса позволяют предположить нормальность распределения показателей.

Матрица парных коэффициентов корреляции

$$R = \begin{array}{c} \\ \end{array}\begin{pmatrix} X_1 & X_2 & X_3 & X_4 & X_5 \\ 1,00 & 0,41 & -0,06 & 0,46 & -0,02 \\ 0,41 & 1,00 & 0,23 & -0,08 & -0,33 \\ -0,06 & 0,23 & 1,00 & -0,91 & -0,16 \\ 0,46 & -0,08 & -0,91 & 1,00 & 0,12 \\ -0,02 & -0,33 & -0,16 & 0,12 & 1,00 \end{pmatrix}$$

показывает, что наиболее тесно с результативным показателем $Y = X_1$ связаны показатели X_4 ($r_{14} = 0,46$) и X_2 ($r_{12} = 0,41$). Что касается факторных признаков (аргументов), то наиболее тесная связь наблюдается между X_3 и X_4 ($r_{34} = - 0,91$), т.е. имеет место мультиколлинеарность. Наличие малоинформативных признаков и мультиколлинеарности требует проверки включения того или иного аргумента в уравнение регрессии.

Наиболее эффективными методами, устраняющими мультиколлинеарность и включающими в уравнение лишь наиболее информативные признаки, являются пошаговые процедуры, построенные на статистически независимых переменных, полученных из исходных признаков на основе компонентного анализа.

Поскольку компонентный анализ требует серьезной работы по содержательной интерпретации построенных главных компонент, то целесообразно начать с пошагового регрессионного анализа. Далее в зависимости от полученных результатов определяется целесообразность проведения компонентного анализа с последующим построением уравнения регрессии на главные компоненты.

Пошаговые процедуры регрессионного анализа обеспечивают отбор информативных признаков и построение уравнения регрессии, которое будет значимым само и включать в себя только значимые переменные.

Так как аргументы X_3 и X_4 тесно связаны между собой ($r_{34} = -0{,}91$), в регрессионную модель результативного показателя X_1 нецелесообразно включать оба этих аргумента. Предпочтение было отдано показателю X_4 (число фармацевтов на 10 тыс. жителей), который включен в модель по экономическим и статистическим соображениям.

По алгоритму пошагового регрессионного анализа исходя из предположения линейности была получена следующая модель:

$$\hat{y} = 11{,}7 + 0{,}059X_2 + 0{,}419X_4.$$

Уравнение статистически значимо по F-критерию, так как $F_{набл} = 3{,}54$ превосходит критическое значение $F_{кр} = 3{,}44$, соответствующее уровню значимости $\alpha = 0{,}05$ и числам степеней свободы $\nu_1 = 2$ и $\nu_2 = 22$.

Значимость коэффициентов регрессии проверялась по t-критерию. Оба коэффициента значимы при $\alpha = 0{,}1$ и $\nu = 22$, так как наблюдаемые значения $t_2 = 1{,}82$ и $t_4 = 2{,}05$ превосходят по абсолютной величине $t_{кр} = 1{,}72$.

Показатель X_5 не включен в уравнение регрессии на основании t-критерия ($t_5 = 0{,}27$).

Полученная модель достаточно адекватна исследуемому явлению, так как средняя относительная ошибка аппроксимации $\bar{\delta} = 3{,}4$ %. Множественный коэффициент детерминации $r_1^2 = 0{,}41$ показывает, что только 41% вариации результативного показателя объясняется влиянием аргументов, включенных в модель.

Заметим, что к низкой информативности модели ($r_1^2 = 0{,}41$) могло привести и исключение показателей X_3 и X_5. В этой связи целесообразно построить регрессионную модель на главных компонентах.

Компонентный анализ проведен по результатам 25 наблюдений ($n = 25$) над четырьмя показателями (X_2, X_3, X_4, X_5).

В результате ортогонального преобразования корреляционной матрицы получены собственные значения: $\lambda_1 = 2{,}03$; $\lambda_2 = 1{,}21$; $\lambda_3 = 0{,}68$ и $\lambda_4 = 0{,}08$. Вклад в суммарную дисперсию первой главной компоненты равен 50,7%; второй – 30,3, третьей – 17,0 и четвертой – 2,0%.

С целью снижения размерности факторного пространства и последующей классификации объектов следует ограничиться двумя первыми главными компонентами, общий вклад которых в суммарную дисперсию составляет 81,0%.

Для экономической интерпретации главных компонент рассмотрим матрицу факторных нагрузок:

$$\mathbf{A} = \begin{pmatrix} f_1 & f_2 & f_3 & f_4 \\ -0{,}396 & -0{,}718^* & 0{,}573 & -0{,}034 \\ -0{,}950^* & 0{,}236 & 0{,}048 & 0{,}201 \\ 0{,}912^* & -0{,}357 & 0{,}060 & 0{,}196 \\ 0{,}370 & 0{,}718^* & 0{,}587 & -0{,}003 \end{pmatrix}$$

(звездочкой обозначены показатели, используемые для интерпретации главной компоненты).

Наибольшую нагрузку на первую главную компоненту имеют показатели X_3 и X_4 (a_{31} = – 0,950; a_{41} = 0,912), т.е. первая главная компонента f_1 имеет тесную отрицательную связь с показателем X_3 и положительную связь с X_4. Компонента f_1 была интерпретирована как уровень обеспеченности населения фармацевтическими кадрами.

Вторая главная компонента f_2 наиболее тесно связана с показателями X_2 и X_5. Причем отрицательная связь имеет место с X_2 и положительная – с X_5 (a_{22} = – 0,718; a_{52} = 0,718).

При интерпретации второй главной компоненты мы исходили из того, что уровень потребления медикаментов у городского населения значительно выше, чем у сельского. Поэтому при плановой системе снабжения населения на основе нормативов потребления чем выше удельный вес городского населения X_2, тем ниже уровень превышения норматива товарных запасов и меньше оборачиваемость оборотных средств X_5. В этой связи вторая главная компонента была названа уровнем превышения норматива товарных запасов.

Далее была проведена классификация регионов по двум первым главным компонентам, вклад которых в суммарную вариацию признаков составляет 81%. Применялась агломеративная иерархическая процедура кластерного анализа. При классификации использовалось обычное евклидово расстояние, а расстояние между кластерами определялось по принципу "ближайшего соседа".

В результате классификации совокупность из n = 25 регионов была разбита на три кластера. В первый кластер попало n_1 = 10 регионов, во второй n_2 = 14 и в третий – n_3 = 1 регион, который следует считать аномальным наблюдением. Анализ средних значений показателей для трех кластеров (см. таблицу) позволяет наполнить формальные результаты классификации экономическим содержанием.

№ кластера (i)	n_i	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$
1	10	18,8	68,1	22,1	8,6	199,4
2	14	19,8	64,1	20,3	9,9	217,3
3	1	21,8	81,6	19,9	11,0	185

Так, в первый кластер входят районы с низким уровнем обеспечения населения медикаментами ($\bar{x}_1$ = 18,6 руб./чел.) и фармацевтическими кадрами ($\bar{x}_4$ = 8,6). В районах, входящих во второй кластер, аптечная служба поставлена лучше. Особое положение занимает район, входящий в третий кластер. В нем самые высокие доля городского населения (81,6%) и объем реализации медикаментов (21,8 руб./чел.), а также самая низкая оборачиваемость оборотных средств (185 дней).

Как уже отмечалось, необходимость включения в регрессионную модель объема реализации медикаментов всех показателей-аргументов приводит к построению уравнения регрессии на главных компонентах. Идея этого метода заключается в том, что рассчитываются значения главных компонент для каждого наблюдения, а затем к этим "новым обобщенным признакам" и выбранному результативному показателю применяется один из методов регрессионного анализа.

При пошаговой регрессии на главные компоненты не обязательно самые весомые компоненты являются и самыми информативными для результативного признака. Поэтому при построении уравнения рассматриваются все главные компоненты.

В результате пошагового регрессионного анализа была получена следующая модель

$$\hat{y} = 19{,}6 - 0{,}409 f_2 + 0{,}821 f_4.$$

Уравнение значимо, так как $F_{\text{набл}} = 22{,}7$ больше, чем $F_{\text{кр}} = 3{,}44$, найденное по таблице F-распределения для уровня значимости $\alpha = 0{,}05$ и чисел степеней свободы $\nu_1 = 2$ и $\nu_2 = 22$.

Коэффициенты регрессии также значимы, так как $|t_2| = 3{,}00$ и $|t_4| = 6{,}02$ больше $t_{\text{кр}} = 2{,}07$, найденного при $\alpha = 0{,}05$ и $\nu = 22$ по таблицам t-распределения.

Используя матрицу факторных нагрузок **A**, можно выразить вторую и четвертую главные компоненты через исходные показатели-аргументы, воспользовавшись соотношением $f_\nu = \sum_{j=1}^{k} a_{j\nu} z_j$, где $a_{j\nu}$ – элемент матрицы факторных нагрузок **A**; $z_j = \frac{x_j - \bar{x}_j}{s_j}$.

Окончательно уравнение регрессии имеет вид:

$$\hat{y} = 22{,}3 + 0{,}034X_2 + 0{,}028X_3 + 0{,}246X_4 - 0{,}015X_5.$$

В отличие от уравнения регрессии, полученного непосредственно по исходным аргументам, последняя модель более адекватна рассматриваемому явлению. Ей соответствует средняя относительная ошибка аппроксимации $\bar{\delta} = 1{,}62\%$ (ранее было $\bar{\delta} = 3{,}4\%$), множественный коэффициент детерминации $r_1^2 = 0{,}819$ (ранее $r_1^2 = 0{,}41$), т.е. на 81,9 % вариация Y объясняется аргументами, входящими в модель.

Уважаемый читатель, авторы надеются, что при решении задач с применением совокупности методов многомерного статистического анализа и интерпретации полученных результатов данный справочник окажется вам полезным.

Приложение 1

МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ ТАБЛИЦЫ

Методические указания к использованию таблиц

В таблице П. 1.1 табулирована функция

$$\Phi(t) = \frac{2}{\sqrt{2\pi}} \int_0^t e^{-\frac{x^2}{2}} dx,$$

изображенная на рис. П.1; $f(t)$ – плотность нормированной нормально распределенной случайной величины $T \in N(0, 1)$.

Вероятность попадания случайной величины T в интервал от t_1 до t_2 вычисляется по формуле

$$P(t_1 < T < t_2) = \frac{1}{2}[\Phi(t_2) - \Phi(t_1)].$$

$\Phi(t)$ обладает следующими свойствами:

$$\Phi(-t) = -\Phi(t);\ \Phi(\infty) = 1;\ \Phi(3) = 0{,}9973.$$

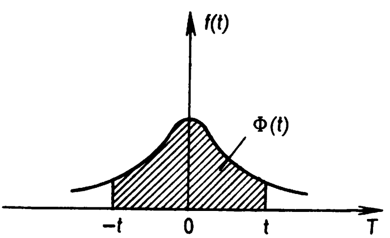
Пример. Определить вероятность:

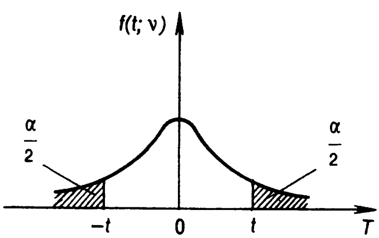
$$P(-1{,}36 < T < 2{,}15) = \frac{1}{2}[\Phi(2{,}15) - \Phi(-1{,}36)] = \frac{1}{2}[0{,}9684 + 0{,}8262] = 0{,}8973.$$

В таблице П.1.2 табулирована вероятность выхода за пределы интервала от $-t$ до $+t$ случайной величины, имеющей распределение Стьюдента (t-распределение) с числом степеней свободы ν (рис. П.2)

$$\alpha = St(t; \nu) = P(|T| > t),$$

$f(t;\ \nu)$ – плотность распределения Стьюдента с числом степеней свободы ν.





Вероятность попадания случайной величины T в интервал от t_1 до t_2 вычисляется по формуле:

$$P(t_1 < T < t_2) = \frac{1}{2}[St(t_1) - St(t_2)].$$

Функция $St(t)$ обладает следующими свойствами:

$$St(-t) = 2 - St(t);\ St(\infty) = 0;\ St(-\infty) = 2;\ St(0) = 1.$$

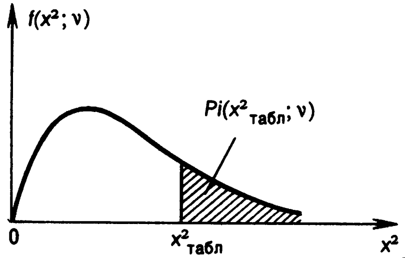
Пример. При $\nu = 10$ определить вероятность:

$$P(-1{,}36 < T < 2{,}15) = \frac{1}{2}[St(-1{,}36) - St(2{,}15)] =$$
$$= \frac{1}{2}[2 - St(1{,}36) - St(2{,}15)] \approx \frac{1}{2}[2 - St(1{,}372) - St(2{,}228)] =$$
$$= \frac{1}{2}[2 - 0{,}2 - 0{,}05] = 0{,}875.$$

Чтобы не прибегать к интерполяции, в строке, соответствующей $\nu = 10$, берем ближайшие к заданным значениям 1,36 и 2,15.

Каждая строка таблицы отвечает t-распределению с соответствующим числом степеней свободы ν.

В таблице П. 1.3 табулирована вероятность того, что наблюдаемое значение случайной величины χ^2, имеющей распределение Пирсона (хи-квадрат-распределение) с числом степеней свободы ν, превысит табличное значение $\chi^2_{табл}(\nu)$.



На рис. П.3 представлен график функции $f(\chi^2_{табл}\nu)$ – плотности χ^2-распределения с числом степеней свободы ν.

Вероятность попадания случайной величины χ^2 в интервал от χ^2_1 до χ^2_2 вычисляется по формуле:

$$P(\chi^2_1 < \chi^2 < \chi^2_2) = P(\chi^2 > \chi^2_1) - P(\chi^2 > \chi^2_2) = Pi(\chi^2_1) - Pi(\chi^2_2).$$

Функция $Pi(\chi^2_{табл})$ обладает следующими свойствами:

$$Pi(0) = 1;\ Pi(\infty) = 0.$$

Пример. При $\nu = 10$ определить

$$P(2{,}5 < \chi^2 < 19{,}0) = Pi(2{,}5) - Pi(19{,}0) \approx Pi(2{,}558) - Pi(18{,}307) = 0{,}99 - 0{,}05.$$

Чтобы не прибегать к интерполяции в строке таблицы, соответствующей $\nu = 10$, берем ближайшие к заданным значениям 2,5 и 19,0.

Каждая строка таблицы отвечает χ^2-распределению с соответствующим числом степеней свободы ν.

В таблице П. 1.4 для случайной величины F, имеющей закон распределения Фишера–Снедекора (F-распределение) с числами степеней свободы числителя ν_1 и знаменателя ν_2, протабулированы три табличных значения, соответствующие трем вероятностям (уровням значимости):

$$\alpha = P(F > F_{\text{табл}}) = 0{,}05;\ 0{,}01 \text{ и } 0{,}001.$$

Пример. Уровню значимости $\alpha = 0{,}01$ и числам степеней свободы числителя $\nu_1 = 5$ и знаменателя $\nu_2 = 7$ соответствует $F_{\text{табл}} = 7{,}46$.

Статистика F строится таким образом, чтобы наблюдаемое значение было не меньше единицы.

Таблица П. 1.1. **Нормальное распределение**

t	Значение функции $\Phi(t) = P(\lvert T \rvert \leqslant t_{\text{табл}})$									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0797	0876	0955	1034	1113	1192	1271	1350	1428	1507
0,2	1585	1663	1741	1819	1897	1974	2051	2128	2205	2282
0,3	2358	2434	2510	2586	2661	2737	2812	2886	2960	3035
0,4	3108	3182	3255	3328	3401	3473	3545	3616	3688	3759
0,5	3829	3899	3969	4039	4108	4177	4245	4313	4381	4448
0,6	4515	4581	4647	4713	4778	4843	4907	4971	5035	5098
0,7	5161	5223	5285	5346	5407	5467	5527	5587	5646	5705
0,8	5763	5821	5878	5935	5991	6047	6102	6157	6211	6265
0,9	6319	6372	6424	6476	6528	6579	6629	6679	6729	6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	7287	7330	7373	7415	7457	7499	7540	7580	7620	7660
1,2	7699	7737	7775	7813	7850	7887	7923	7959	7994	8029
1,3	8064	8098	8132	8165	8198	8230	8262	8293	8324	8355
1,4	8385	8415	8444	8473	8501	8529	8557	8584	8611	8638
1,5	8664	8690	8715	8740	8764	8789	8812	8836	8859	8882
1,6	8904	8926	8948	8969	8990	9011	9031	9051	9070	9090
1,7	9109	9127	9146	9164	9181	9199	9216	9233	9249	9265
1,8	9281	9297	9312	9327	9342	9357	9371	9385	9399	9412
1,9	9426	9439	9451	9464	9476	9488	9500	9512	9523	9534
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	9643	9651	9660	9668	9676	9684	9692	9700	9707	9715
2,2	9722	9729	9736	9743	9749	9756	9762	9768	9774	9780
2,3	9786	9791	9797	9802	9807	9812	9817	9822	9827	9832
2,4	9836	9841	9845	9849	9853	9857	9861	9865	9869	9872

Продолжение

t	Значение функции $\Phi(t) = P(\lvert T \rvert \leq t_{табл})$									
	0	1	2	3	4	5	6	7	8	9
2,5	9876	9879	9883	9886	9889	9892	9895	9898	9901	9904
2,6	9907	9910	9912	9915	9917	9920	9922	9924	9926	9928
2,7	9931	9933	9935	9937	9939	9940	9942	9944	9946	9947
2,8	9949	9951	9952	9953	9955	9956	9958	9959	9960	9961
2,9	9963	9964	9965	9966	9967	9968	9969	9970	9971	9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9980
3,1	9981	9981	9982	9983	9983	9984	9984	9985	9985	9986
3,5	9995	9996	9996	9996	9996	9996	9996	9996	9997	9997
3,6	9997	9997	9997	9997	9997	9997	9997	9998	9998	9998
3,7	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998
3,8	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
3,9	9999	9999	9999	9999	9999	9999	9999	9999	9999	9999
4,0	0,999936	9999	9999	9999	9999	9999	9999	9999	9999	9999
4,5	0,999994	–	–	–	–	–	–	–	–	–
5,0	0,99999994	–	–	–	–	–	–	–	–	–

Таблица П. 1.2. **Распределение Стьюдента** (t-распределение)

ν	Значения t при вероятности $\alpha = St(t) = P(\lvert T \rvert > t_{табл})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,941
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,043	6,859
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,583
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,833
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850

ν	Значения t при вероятности $\alpha = St(t) = P(\lvert T\rvert > t_{табл})$												
	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1	0,05	0,02	0,01	0,001
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,868	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,402	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
60	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

Таблица П. 1.3. **Распределение Пирсона** (χ^2-распределение)

ν	Значения $\chi^2_{табл}$ для вероятностей $P(\chi^2 > \chi^2_{табл})$, равных										
	0,999	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
1	0,05157	0,04393	0,03157	0,03628	0,03982	0,00393	0,0158	0,0642	0,102	0,148	0,455
2	0,00200	0,0100	0,0201	0,0404	0,0506	0,103	0,211	0,446	0,575	0,713	1,386
3	0,0243	0,0717	0,115	0,185	0,216	0,352	0,584	1,005	1,213	1,424	2,366
4	0,0908	0,207	0,297	0,429	0,484	0,711	1,064	1,649	1,923	2,195	3,357
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	2,675	3,000	4,351
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	3,455	3,828	5,348
7	0,598	0,989	1,239	1,564	1,690	2,167	2,833	3,822	4,255	4,671	6,346
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	5,071	5,527	7,344
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	5,899	6,393	8,343
10	1,479	2,156	2,558	3,059	3,247	3,240	4,865	6,179	6,737	7,267	9,342
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	7,584	8,148	10,341
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	8,438	9,034	11,340
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	9,299	9,926	12,340
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	10,165	10,821	13,339
15	3,483	4,601	5,229	5,985	6,262	7,261	8,547	10,307	11,036	11,721	14,339
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	11,912	12,624	15,338
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	12,892	13,531	16,338
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	13,675	14,440	17,338
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352	18,338
20	5,921	7,434	8,260	9,237	9,591	10,871	12,443	14,578	15,452	16,266	19,337
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	16,344	17,182	20,337
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	17,240	18,101	21,337
23	7,529	9,260	10,196	11,293	11,688	13,091	14,848	17,187	18,137	19,021	22,337
24	8,035	9,886	10,856	11,992	12,401	13,848	15,659	18,062	19,037	19,943	23,337
25	8,649	10,520	11,524	12,697	13,120	14,611	16,173	18,940	19,939	20,887	24,337
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	20,843	21,792	25,336
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	21,749	22,719	26,136
28	10,391	12,461	13,565	14,847	15,308	16,928	18,937	21,588	22,657	23,617	27,336
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	23,567	24,577	28,336
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	24,478	25,508	29,336

Продолжение

ν	Значения $\chi^2_{\text{табл}}$ для вероятностей $P(\chi^2 > \chi^2_{\text{табл}})$, равных									
	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
1	1,074	1,323	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	2,408	2,773	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	3,665	4,108	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,268
4	4,878	5,385	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,465
5	6,064	6,626	7,289	9,236	11,070	12,839	13,388	15,086	16,750	20,517
6	7,231	7,841	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	8,383	9,037	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	9,524	10,219	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,125
9	10,656	11,389	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	11,781	12,549	13,412	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	12,899	13,701	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	14,011	14,845	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	15,119	15,984	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	16,222	17,117	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	17,322	18,245	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	18,418	19,369	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	19,511	20,489	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	20,601	21,605	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	21,689	22,718	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	22,775	23,828	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	23,858	24,935	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	24,939	26,039	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	26,018	27,141	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	27,096	28,241	29,553	33,196	36,415	39,364	40,270	42,980	45,558	51,170
25	28,172	29,339	30,675	34,382	37,652	40,046	41,566	44,314	46,928	52,620
26	29,246	30,434	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	30,319	31,528	32,912	36,741	40,113	43,194	44,140	46,963	49,645	55,476
28	31,391	32,620	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,893
29	32,461	33,711	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,302
30	33,530	34,800	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,703

Таблица П. 1.4. Распределение Фишера — Снедекора (F-распределение)
(Первое значение соответствует вероятности 0,05;
второе — вероятности 0,01 и третье — вероятности 0,001;
ν_1 — число степеней свободы числителя, ν_2 — знаменателя)

ν_2 \ ν_1	Значения $F_{табл}$, удовлетворяющие условию $P(F > F_{табл})$										
	1	2	3	4	5	6	8	12	24	∞	t
1	161,4	199,5	215,7	224,6	230,2	234,0	238,9	243,9	249,0	253,3	12,71
	4052	4999	5403	5625	5764	5859	5981	6106	6234	6366	63,66
	406523	500016	536700	562527	576449	585953	598149	610598	623432	636535	636,2
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50	4,30
	98,49	99,01	99,17	99,25	99,30	99,33	99,36	99,42	99,46	99,50	9,92
	998,46	999,00	999,20	999,20	999,20	999,20	999,40	999,60	999,40	999,40	31,00
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53	3,18
	34,12	30,81	29,46	28,71	28,24	27,91	27,49	27,05	26,60	26,12	5,84
	67,47	148,51	141,10	137,10	134,60	132,90	130,60	128,30	125,90	123,50	12,94
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63	2,78
	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,37	13,93	13,46	4,60
	74,13	61,24	56,18	53,43	51,71	50,52	49,00	47,41	45,77	44,05	8,61
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36	2,57
	16,26	13,27	12,06	11,39	10,97	10,67	10,27	9,89	9,47	9,02	4,03
	47,04	36,61	33,20	31,09	20,75	28,83	27,64	26,42	25,14	23,78	6,86
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67	2,45
	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,72	7,31	6,88	3,71
	35,51	26,99	23,70	21,90	20,81	20,03	19,03	17,99	16,89	15,75	5,96
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23	2,36
	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,47	6,07	5,65	3,50
	29,22	21,69	18,77	17,19	16,21	15,52	14,63	13,71	12,73	11,70	5,40

ν_2 \ ν_1	Значения $F_{табл}$, удовлетворяющие условию $P(F > F_{табл})$										
	1	2	3	4	5	6	8	12	24	∞	t
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,99	2,31
	11,26	8,65	7,59	7,10	6,63	6,37	6,03	5,67	5,28	4,86	3,36
	25,42	18,49	15,83	14,39	13,49	12,86	12,04	11,19	10,30	9,35	5,04
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71	2,26
	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,11	4,73	4,31	3,25
	22,86	16,39	13,90	12,56	11,71	11,13	10,37	9,57	8,72	7,81	4,78
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54	2,23
	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,71	4,33	3,91	3,17
	21,04	14,91	12,55	11,28	10,48	9,92	9,20	8,45	7,64	6,77	4,59
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40	2,20
	9,65	7,20	6,22	5,67	5,32	5,07	4,74	4,40	4,02	3,60	3,11
	19,69	13,81	11,56	10,35	9,58	9,05	8,35	7,62	6,85	6,00	4,49
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30	2,18
	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,16	3,78	3,36	3,06
	18,64	12,98	10,81	9,63	8,89	8,38	7,71	7,00	6,25	5,42	4,32
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21	2,16
	9,07	6,70	5,74	5,20	4,86	4,62	4,30	3,96	3,59	3,16	3,01
	17,81	12,31	10,21	9,07	8,35	7,86	7,21	6,52	5,78	4,97	4,12
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13	2,14
	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,80	3,43	3,00	2,98
	17,14	11,78	9,73	8,62	7,92	7,44	6,80	6,13	5,41	4,60	4,14
15	4,45	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07	2,13
	8,68	6,36	5,42	4,89	4,56	4,32	4,00	3,67	3,29	2,87	2,95
	16,59	11,34	9,34	8,25	7,57	7,09	6,47	5,81	5,10	4,31	4,07

Продолжение

ν_2 \ ν_1	Значения $F_{табл}$, удовлетворяющие условию $P(F > F_{табл})$										
	1	2	3	4	5	6	8	12	24	∞	t
16	4,41	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01	2,12
	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,55	3,18	2,75	2,92
	16,12	10,97	9,01	7,94	7,27	6,80	6,20	5,55	4,85	4,06	4,02
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96	2,11
	8,40	6,11	5,18	4,67	4,34	4,10	3,79	3,45	3,08	2,65	2,90
	15,72	10,66	8,73	7,68	7,02	6,56	5,96	5,32	4,63	3,85	3,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92	2,10
	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,37	3,01	2,57	2,88
	15,38	10,39	8,49	7,46	6,81	6,35	5,76	5,13	4,45	3,67	3,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88	2,09
	8,18	5,93	5,01	4,50	4,17	3,94	3,63	3,30	2,92	2,49	2,86
	15,08	10,16	8,28	7,26	6,61	6,18	5,59	4,97	4,29	3,52	3,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84	2,09
	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,23	2,86	2,42	2,84
	14,82	9,95	8,10	7,10	6,46	6,02	5,44	4,82	4,15	3,38	3,85
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,82	2,08
	8,02	5,78	4,87	4,37	4,04	3,81	3,51	3,17	2,80	2,36	2,83
	14,62	9,77	7,94	6,95	6,32	5,88	5,31	4,70	4,03	3,26	3,82
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78	2,07
	7,94	5,72	4,82	4,31	3,99	3,75	3,45	3,12	2,75	2,30	2,82
	14,38	9,61	7,80	6,81	6,19	5,76	5,19	4,58	3,92	3,15	3,79
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76	2,07
	7,88	5,66	4,76	4,26	3,94	3,71	3,41	3,07	2,70	2,26	2,81
	14,19	9,46	7,67	6,70	6,08	5,56	5,09	4,48	3,82	3,05	3,77

Продолжение

ν_2 \ ν_1	Значения $F_{табл}$, удовлетворяющие условию $P(F > F_{табл})$										
	1	2	3	4	5	6	8	12	24	∞	t
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73	2,06
	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,03	2,66	2,21	2,80
	14,03	9,34	7,55	6,59	5,98	5,55	4,99	4,39	3,74	2,97	3,75
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71	2,06
	7,77	5,57	4,68	4,18	3,86	3,63	3,32	2,99	2,62	2,17	2,79
	13,88	9,22	7,45	6,49	5,89	5,46	4,91	4,31	3,66	2,89	3,72
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69	2,06
	7,72	5,53	4,64	4,14	3,82	3,59	3,29	2,96	2,58	2,13	2,78
	13,74	9,12	7,36	6,41	5,80	5,38	4,83	4,24	3,59	2,82	3,71
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67	2,05
	7,68	5,49	4,60	4,11	3,78	3,56	3,26	2,93	2,55	2,10	2,77
	13,61	9,02	7,27	6,33	5,73	5,31	4,76	4,17	3,52	2,76	3,69
28	4,19	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65	2,05
	7,64	5,45	4,57	4,07	3,75	3,53	3,23	2,90	2,52	2,06	2,76
	13,50	8,93	7,18	6,25	5,66	5,24	4,69	4,11	3,46	2,70	3,67
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64	2,05
	7,60	5,42	4,54	4,04	3,73	3,50	3,20	2,87	2,49	2,03	2,76
	13,39	8,85	7,12	6,19	5,59	5,18	4,65	4,05	3,41	2,64	3,66
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62	2,04
	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,84	2,47	2,01	2,75
	13,29	8,77	7,05	6,12	5,53	5,12	4,58	4,00	3,36	2,59	3,64
60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39	2,00
	7,08	4,98	4,13	3,65	3,34	3,12	2,82	2,50	2,12	1,60	2,66
	11,97	7,76	6,17	5,31	4,76	4,37	3,87	3,31	2,76	1,90	3,36
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,03	1,96
	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,18	1,79	1,04	2,58
	10,83	6,91	5,42	4,62	4,10	3,74	3,27	2,74	2,13	1,05	3,29

Таблица П. 1.5. **Таблица Фишера—Иейтса**
(H_1: $\rho \neq 0$; $\nu = n - 2$ в случае парной корреляции и $\nu = n - l - 2$,
где l — число исключенных величин, в случае частной корреляции)

ν	Значения $r_{кр}$ для уровня значимости α				ν	Значения $r_{кр}$ для уровня значимости α			
	0,05	0,02	0,01	0,001		0,05	0,02	0,01	0,001
1	0,997	1,000	1,000	1,000	16	0,468	0,543	0,590	0,708
2	0,950	0,980	0,990	0,999	17	0,456	0,529	0,575	0,693
3	0,878	0,934	0,959	0,991	18	0,444	0,516	0,561	0,679
4	0,811	0,882	0,917	0,974	19	0,433	0,503	0,549	0,665
5	0,754	0,833	0,875	0,951	20	0,423	0,492	0,537	0,652
6	0,707	0,789	0,834	0,925	25	0,381	0,445	0,487	0,597
7	0,666	0,750	0,798	0,898	30	0,349	0,409	0,449	0,554
8	0,632	0,715	0,765	0,872	35	0,325	0,381	0,418	0,519
9	0,602	0,685	0,735	0,847	40	0,304	0,358	0,393	0,490
10	0,576	0,658	0,708	0,823	45	0,288	0,338	0,372	0,465
11	0,553	0,634	0,684	0,801	50	0,273	0,322	0,354	0,443
12	0,532	0,612	0,661	0,780	60	0,250	0,295	0,325	0,408
13	0,514	0,592	0,641	0,760	70	0,232	0,274	0,302	0,380
14	0,497	0,574	0,623	0,742	80	0,217	0,257	0,283	0,338
15	0,482	0,558	0,606	0,725	90	0,205	0,242	0,267	0,338
					100	0,195	0,230	0,254	0,321

Таблица П. 1.6. **Таблица Z-преобразования Фишера**

r	Значения $Z = \frac{1}{2}\{\ln(1+r) - \ln(1-r)\}$									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0101	0,0200	0,0300	0,0400	0,0501	0,0601	0,0701	0,0802	0,0902
1	0,1003	0,1104	0,1206	0,1308	0,1409	0,1511	0,1614	0,1717	0,1820	0,1923
2	0,2027	0,2132	0,2237	0,2342	0,2448	0,2554	0,2661	0,2769	0,2877	0,2986
3	0,3095	0,3205	0,3316	0,3428	0,3541	0,3654	0,3767	0,3884	0,4001	0,4118
4	0,4236	0,4356	0,4477	0,4599	0,4722	0,4847	0,4973	0,5101	0,5230	0,5361
5	0,5493	0,5627	0,5764	0,5901	0,6042	0,6184	0,6328	0,6475	0,6625	0,6777
6	0,6932	0,7089	0,7250	0,7414	0,7582	0,7753	0,7928	0,8107	0,8291	0,8480
7	0,8673	0,8872	0,9077	0,9287	0,9505	0,9730	0,9962	1,0203	1,0454	1,0714
8	1,0986	1,1270	1,1568	1,1881	1,2212	1,2562	1,2933	1,3331	1,3758	1,4219
9	1,4722	1,5275	1,5890	1,6584	1,7381	1,8318	1,9459	2,0923	2,2976	2,6467
0,99	2,6466	2,6996	2,7587	2,8257	2,9031	2,9945	3,1063	3,2504	3,4534	3,8002

Таблица П. 1.7. **Пяти- и однопроцентные пределы для отношения G наибольшей выборочной дисперсии к сумме l выборочных дисперсий, полученных из l независимых выборок объема n**

(Первое значение соответствует уровню значимости $\alpha = 0{,}05$, второе — 0,01)

l \ $n-1$	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0,998	0,975	0,939	0,906	0,877	0,853	0,833	0,816	0,801	0,788	0,734	0,660	0,518	0,500
	0,999	0,995	0,979	0,959	0,937	0,917	0,809	0,882	0,867	0,854	0,795	0,700	0,606	0,500
3	0,967	0,871	0,798	0,746	0,707	0,677	0,653	0,633	0,617	0,603	0,547	0,475	0,403	0,333
	0,993	0,942	0,883	0,834	0,903	0,761	0,734	0,711	0,691	0,674	0,606	0,515	0,423	0,333
4	0,906	0,768	0,684	0,629	0,590	0,560	0,537	0,518	0,502	0,488	0,437	0,372	0,309	0,250
	0,968	0,864	0,781	0,721	0,676	0,641	0,613	0,590	0,570	0,554	0,488	0,406	0,325	0,250
5	0,841	0,684	0,598	0,544	0,507	0,478	0,456	0,439	0,424	0,412	0,365	0,307	0,251	0,200
	0,928	0,789	0,696	0,633	0,588	0,553	0,526	0,504	0,485	0,470	0,409	0,335	0,264	0,200
6	0,781	0,616	0,532	0,480	0,445	0,418	0,398	0,382	0,368	0,357	0,314	0,261	0,212	0,167
	0,883	0,722	0,626	0,564	0,520	0,487	0,461	0,440	0,423	0,408	0,353	0,286	0,223	0,167
7	0,727	0,561	0,480	0,431	0,397	0,373	0,354	0,338	0,326	0,315	0,276	0,228	0,183	0,143
	0,838	0,664	0,569	0,508	0,466	0,435	0,411	0,391	0,375	0,362	0,311	0,249	0,193	0,143
8	0,680	0,516	0,438	0,391	0,360	0,336	0,319	0,304	0,293	0,283	0,246	0,202	0,162	0,125
	0,795	0,615	0,521	0,463	0,423	0,393	0,370	0,352	0,337	0,325	0,278	0,221	0,170	0,125
9	0,639	0,478	0,403	0,358	0,329	0,307	0,290	0,277	0,266	0,257	0,223	0,182	0,145	0,111
	0,754	0,573	0,481	0,425	0,387	0,359	0,338	0,321	0,307	0,295	0,251	0,199	0,152	0,111

Приложение 2

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

П. 2.1. РЕГРЕССИОННЫЙ И КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

Варианты задач 1 – 25 с указанием результативного y и факторных x_1, x_2 признаков даны в табл. П. 2.1.

По выборочным данным, представленным в приложении П. 2.5, исследовать на основе линейной регрессионной модели зависимость одного из результативных признаков от показателей производственно-хозяйственной деятельности предприятий машиностроения.

Для этого требуется:

1) найти оценку уравнения регрессии вида

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2, \text{ т.е. вектор } \mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix};$$

2) проверить значимость уравнения регрессии при $\alpha = 0{,}05$ или $\alpha = 0{,}01$;

3) проверить значимость отдельных коэффициентов регрессии β_0, β_1, β_2;

4) построить интервальные оценки для значимых коэффициентов регрессии при $\gamma = 1 - \alpha$;

5) дать экономическую интерпретацию коэффициентам регрессии и оценить адекватность полученной модели по величине абсолютных e_i и относительных δ_i отклонений;

6) при необходимости перейти к алгоритму пошагового регрессионного анализа, отбросив один из незначимых коэффициентов регрессии;

7) построить матрицы парных и частных коэффициентов корреляции;

8) найти множественные коэффициенты корреляции и детерминации;

9) проверить значимость частных и множественных коэффициентов корреляции;

10) построить интервальные оценки частных коэффициентов корреляции;

11) провести содержательный экономический анализ полученных результатов.

Таблица П. 2.1. **Варианты задач для самостоятельной работы по регрессионному и корреляционному анализу**

Номер варианта	Результативный признак	Факторные признаки	Номер варианта	Результативный признак	Факторные признаки
1	Y_1	X_1, X_3	14	Y_3	X_1, X_{14}
2	Y_2	X_1, X_5	15	Y_2	X_5, X_9
3	Y_2	X_1, X_7	16	Y_3	X_8, X_{10}
4	Y_2	X_1, X_{11}	17	Y_3	X_7, X_{14}
5	Y_2	X_1, X_{10}	18	Y_3	X_3, X_6
6	Y_1	X_3, X_4	19	Y_3	X_1, X_{14}
7	Y_2	X_3, X_{11}	20	Y_1	X_2, X_6
8	Y_2	X_{11}, X_{15}	21	Y_1	X_3, X_7
9	Y_1	X_3, X_5	22	Y_2	X_5, X_8
10	Y_2	X_{11}, X_{16}	23	Y_2	X_9, X_{10}
11	Y_2	X_1, X_6	24	Y_3	X_4, X_{11}
12	Y_2	X_1, X_{12}	25	Y_3	X_1, X_{12}
13	Y_2	X_1, X_2			

П. 2.2. КОМПОНЕНТНЫЙ АНАЛИЗ

Для выполнения самостоятельной работы в табл. П. 2.2 приводится перечень вариантов и порядок их формирования из набора показателей производственно-хозяйственной деятельности предприятий машиностроения и их значений, приведенных в Приложении П. 2.5.

Требуется:

1) найти оценку матрицы **R** парных коэффициентов корреляции;

2) рассчитать матрицу **Λ** собственных значений. Определить вклад компонент в суммарную дисперсию признаков;

3) получить матрицу **V** нормируемых собственных векторов и матрицу **А** факторных нагрузок. Дать экономическую интерпретацию полученным главным компонентам;

4) вычислить матрицу **F** значений главных компонент. Провести классификацию объектов по двум первым главным компонентам. Дать интерпретацию полученным результатам.

Таблица П. 2.2. **Варианты задач для самостоятельной работы по компонентному анализу**

Номер варианта	Показатели	Номер варианта	Показатели
1	X_1, X_2, X_{10}	13	X_3, X_6, X_{13}
2	X_2, X_9, X_{11}	14	X_3, X_{11}, X_{12}
3	X_2, X_4, X_7	15	X_3, X_{11}, X_{14}
4	X_2, X_7, X_{12}	16	X_2, X_4, X_6
5	X_3, X_7, X_{14}	17	X_1, X_4, X_6
6	X_4, X_6, X_9	18	X_1, X_2, X_4
7	X_4, X_5, X_{14}	19	X_1, X_7, X_{12}
8	X_5, X_7, X_{14}	20	X_1, X_7, X_{13}
9	X_5, X_7, X_{12}	21	X_2, X_7, X_{14}
10	X_6, X_7, X_{13}	22	X_2, X_7, X_{13}
11	X_4, X_7, X_{11}	23	X_2, X_{11}, X_{14}
12	X_3, X_6, X_{12}	24	X_2, X_{11}, X_{13}
		25	X_2, X_{11}, X_{12}

П. 2.3. ДИСКРИМИНАНТНЫЙ АНАЛИЗ

Имеются 15 предприятий, характеризуемых тремя экономическими показателями (см. табл. П. 2.3): Y_1 – производительность труда, тыс. руб./чел.; X_6 – удельный вес потерь от брака, %; X_7 – фондоотдача активной части основных производственных фондов, руб./руб.

В каждом варианте (табл. П. 2.3) даны две обучающие выборки, первая из которых включает четыре предприятия группы *A*, а вторая – пять предприятий группы *B*. Требуется:

1) найти оценки векторов средних $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$ и ковариационных матриц $\mathbf{S}_x, \mathbf{S}_y$;

2) определить несмещенную оценку суммарной ковариационной матрицы $\hat{\mathbf{S}}$ и обратной матрицы $\hat{\mathbf{S}}^{-1}$;

3) получить вектор оценок коэффициентов дискриминантной функции;

4) найти оценки значений дискриминантной функции $\hat{\mathbf{u}}_X$ и $\hat{\mathbf{u}}_Y$ для матриц исходных данных **X** и **Y**;

5) определить оценку константы $\hat{c}$;

6) вычислить оценки значений дискриминантной функции для оставшихся предприятий и провести их дискриминацию;

7) дать экономическую интерпретацию результатов дискриминации, охарактеризовав эффективность деятельности предприятий, вошедших в группу *A* и группу *B*.

Таблица П. 2.3. **Варианты задач для самостоятельной работы по дискриминантному анализу**

Номер варианта	Номер предприятия		Номер варианта	Номер предприятия	
	группа A	группа B		группа A	группа B
1	1, 2, 3, 10	5, 6, 7, 8, 9	14	1, 3, 10, 13	4, 5, 6, 7, 11
2	1, 2, 3, 13	5, 6, 7, 8, 11	15	2, 3, 13, 15	4, 5, 6, 7, 12
3	1, 2, 3, 15	6, 7, 8, 11, 12	16	1, 2, 3, 13	4, 5, 6, 7, 14
4	1, 2, 10, 13	6, 7, 8, 11, 14	17	1, 2, 10, 13	4, 5, 6, 9, 11
5	1, 2, 10, 15	6, 7, 8, 9, 11	18	1, 2, 13, 15	4, 5, 6, 9, 12
6	1, 2, 13, 15	5, 6, 7, 9, 14	19	1, 3, 10, 15	4, 5, 6, 9, 14
7	1, 3, 13, 15	5, 6, 7, 9, 11	20	1, 10, 13, 15	4, 5, 6, 8, 9
8	1, 3, 10, 15	5, 6, 7, 9, 12	21	2, 3, 10, 13	4, 5, 7, 8, 9
9	1, 10, 13, 15	5, 6, 7, 9, 14	22	2, 3, 10, 15	4, 5, 7, 8, 11
10	2, 10, 13, 15	7, 8, 9, 11, 12	23	3, 10, 13, 15	4, 5, 7, 8, 12
11	2, 3, 10, 13	7, 8, 9, 11, 14	24	1, 3, 10, 13	4, 6, 7, 8, 9
12	2, 3, 10, 15	4, 5, 6, 7, 8	25	1, 2, 3, 10	4, 6, 7, 8, 11
13	3, 10, 13, 15	4, 5, 6, 7, 9			

П. 2.4. КЛАСТЕРНЫЙ АНАЛИЗ

Для классификации шести предприятий, характеризуемых четырьмя экономическими показателями (Y_1, Y_2, Y_3, X_1), требуется:

1) найти матрицу нормированных значений исходных данных Z;

2) построить матрицу расстояний между наблюдениями;

3) реализовать иерархическую агломеративную процедуру кластерного анализа;

4) построить дендрограмму;

Значения показателей производственно-хозяйственной деятельности предприятий машиностроения приводятся в приложении П. 2.5.

Варианты задач даны в табл. П. 2.4.

Таблица П. 2.4. **Варианты для самостоятельной работы по кластерному анализу**

Номер варианта	Номер предприятия	Номер варианта	Номер предприятия
1	3, 5, 6, 7, 8, 10	10	7, 9, 10, 11, 12, 13
2	2, 3, 5, 6, 7, 13	11	2, 3, 7, 11, 13, 14
3	1, 3, 6, 7, 8, 15	12	3, 4, 6, 8, 10, 15
4	2, 6, 7, 8, 10, 13	13	4, 5, 6, 10, 13, 15
5	1, 8, 9, 10, 11, 15	14	1, 3, 4, 5, 7, 13
6	2, 7, 9, 13, 14, 15	15	2, 3, 4, 5, 12, 13
7	3, 6, 7, 9, 13, 15	16	1, 2, 4, 6, 13, 14
8	1, 6, 7, 10, 12, 15	17	2, 4, 5, 10, 11, 13
9	2, 5, 9, 13, 15, 14	18	1, 4, 6, 9, 13, 15

Номер варианта	Номер предприятия	Номер варианта	Номер предприятия
19	1, 4, 9, 10, 14, 15	23	3, 4, 8, 10, 12, 15
20	4, 8, 9, 10, 13, 15	24	1, 4, 6, 7, 10, 13
21	3, 4, 7, 8, 10, 13	25	1, 2, 3, 6, 8, 11
22	2, 5, 8, 10, 11, 15		

П. 2.5. ОСНОВНЫЕ ПОКАЗАТЕЛИ ПРОИЗВОДСТВЕННО-ХОЗЯЙСТВЕННОЙ ДЕЯТЕЛЬНОСТИ МАШИНОСТРОИТЕЛЬНЫХ ПРЕДПРИЯТИЙ

Обозначения и наименования показателей:

Y_1 – производительность труда, тыс. руб./чел.;
Y_2 – индекс снижения себестоимости продукции;
Y_3 – рентабельность, %;
X_1 – трудоемкость единицы продукции, чел.-ч;
X_2 – удельный вес рабочих в составе промышленно-производственного персонала;
X_3 – удельный вес покупных изделий;
X_4 – коэффициент сменности оборудования, смен;
X_5 – премии и вознаграждения на одного работника ППП, тыс. руб.;
X_6 – удельный вес потерь от брака, %;
X_7 – фондоотдача активной части ОПФ, руб./руб.;
X_8 – среднегодовая численность промышленно-производственного персонала, чел.;
X_9 – среднегодовая стоимость основных производственных фондов, млн. руб.;
X_{10} – среднегодовой фонд заработной платы промышленно-производственного персонала, тыс. руб.;
X_{11} – фондовооруженность труда, тыс. руб./чел.;
X_{12} – оборачиваемость нормируемых оборотных средств, дн.;
X_{13} – оборачиваемость ненормируемых оборотных средств, дн.;
X_{14} – непроизводительные расходы, тыс. руб.

Таблица П. 2.5. **Значение показателей производственно-хозяйственной деятельности машиностроительных предприятий**

Номер предприятия	Y_1	Y_2	Y_3	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
1	9,4	62,0	10,6	0,23	0,62	0,40	1,35	0,88	0,15	1,91	7394	39,53	14257	5,35	173,9	11,88	28,13
2	9,9	53,1	9,1	0,43	0,76	0,19	1,39	0,57	0,34	1,68	11586	40,41	22661	3,90	162,3	12,60	17,55
3	9,1	56,5	23,4	0,26	0,71	0,44	1,27	1,70	0,09	1,89	7801	37,02	14903	4,88	101,2	8,28	19,52
4	5,5	30,1	9,7	0,43	0,74	0,25	1,10	0,84	0,05	1,02	6371	41,08	12973	5,65	177,8	17,28	18,13
5	6,6	18,1	9,1	0,38	0,72	0,02	1,23	1,04	0,48	0,88	4210	42,39	6920	8,85	93,2	13,32	21,21
6	4,3	13,6	5,4	0,42	0,68	0,06	1,39	0,66	0,41	0,62	3557	37,39	5736	8,52	126,7	17,28	22,97
7	7,4	89,8	9,9	0,30	0,77	0,15	1,38	0,86	0,62	1,09	14148	101,78	26705	7,19	91,8	9,72	16,38
8	6,6	76,6	19,1	0,37	0,77	0,24	1,35	1,27	0,50	1,32	15118	81,32	28025	5,38	70,6	8,64	16,66
9	5,5	32,3	6,6	0,34	0,72	0,11	1,24	0,68	1,20	0,68	6462	59,92	11049	9,27	97,2	9,00	20,09
10	9,4	199,6	14,2	0,23	0,79	0,47	1,40	0,86	0,21	2,30	24628	107,34	45893	4,36	80,3	14,76	15,98
11	5,7	90,8	8,0	0,41	0,71	0,20	1,28	0,45	0,66	1,43	1948	80,83	36813	4,16	128,5	10,44	22,76
12	5,2	82,1	17,5	0,41	0,79	0,24	1,33	0,74	0,74	1,82	18963	59,42	33956	3,13	94,7	14,76	15,41
13	10,0	76,2	17,2	0,22	0,76	0,54	1,22	1,03	0,32	2,62	9185	36,96	17016	4,02	85,3	20,52	19,35
14	6,7	37,1	12,9	0,31	0,79	0,29	1,35	0,96	0,39	1,24	6391	37,21	11688	5,82	85,3	7,92	14,63
15	9,4	51,6	13,2	0,24	0,70	0,56	1,20	0,98	0,28	2,03	6555	32,87	12243	5,01	116,6	18,72	22,62

Задача № 1 (варианты № 1 – 5).
Варианты № 1 + i* (i = 0, 1, 2, 3, 4).

По данным, приведенным в табл. П. 2.6 и П. 2.7, исследовать влияние коэффициента сменности оборудования (фактор A) на производительность труда.

Таблица П. 2.6. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Номер наблюдения	Коэффициент сменности оборудования		
	1,0 – 1,3	1,3 – 1,6	1,6 – 2,0
1	6,3	6,5	7,9
2	6,3	6,8	7,9
3	6,4	6,9	8,0
4	7,2	7,9	8,5 + 0,2i
5	7,3	7,9	8,4 + 0,2i
6	7,5	8,0	8,4 + 0,2i

Таблица П. 2.7. **Производительность труда на предприятиях машиностроения** (комплекс с неравным числом наблюдений), тыс. руб./чел.

Номер наблюдения	Коэффициент сменности оборудования		
	1,0 – 1,3	1,3 – 1,6	1,6 – 2,0
1	6,3	6,5	7,9
2	6,3	6,8	7,9
3	6,4	6,9	8,0
4	7,2	7,9	8,5 + 0,2i
5	7,3		8,4 + 0,2i
6			8,4 + 0,2i

---

* Подставляя значения числа i, можно обеспечить каждого обучающегося индивидуальным вариантом задания.

Задача № 2 (варианты № 6 – 10).
Варианты № 6 + i (i = 0, 1, 2, 3, 4).

По данным, приведенным в табл. П. 2.8 и П. 2.9, исследовать влияние на производительность труда среднесписочной численности промышленно-производственного персонала (фактор A).

Таблица П. 2.8. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Номер наблюдения	Численность ППП, человек		
	500 – 1000	1000 – 2000	2000 – 5000
1	4,0	4,0	4,5 + 0,2i
2	4,1	5,1	5,5 + 0,2i
3	3,9	5,2	6,0 + 0,2i
4	5,2	6,1	6,5 + 0,2i
5	5,6	6,9	6,0 + 0,2i
6	6,0	6,5	5,9 + 0,2i

Таблица П. 2.9. **Производительность труда на предприятиях машиностроения** (комплекс с неравным числом наблюдений), тыс. руб./чел.

Номер наблюдения	Численность ППП, человек		
	500 – 1000	1000 – 2000	2000 – 5000
1	4,0	4,0	4,5 + 0,2i
2	4,1	5,1	5,5 + 0,2i
3	3,9	5,2	6,0 + 0,2i
4	5,2	6,1	6,5 + 0,2i
5	5,6	6,9	
6	6,0		

Задача № 3 (варианты № 11 – 15).
Варианты № 11 + i (i = 0, 1, 2, 3, 4).

По данным, приведенным в табл. П. 2.10 и П. 2.11, исследовать влияние фондовооруженности рабочего (фактор A) на себестоимость товарной продукции.

Таблица П. 2.10. **Себестоимость товарной продукции,** руб.

Номер наблюдения	Фондовооруженность рабочего, тыс. руб.		
	15 – 25	25 – 35	35 – 45
1	65	62	52 + i
2	60	63	51 + i
3	65	60	50 + i
4	62	58	51 + i
5	60	58	52 + i
6	59	56	50 + i

Таблица П. 2.11. **Себестоимость товарной продукции**
(комплекс с неравным числом наблюдений), руб.

Номер наблюдения	Фондовооруженность рабочего, тыс. руб.		
	15 – 25	25 – 35	35 – 45
1	65	62	52 + i
2	60	63	51 + i
3	65	60	50 + i
4	62	58	51 + i
5		58	52 + i
6		56	

Задача № 4 (варианты № 16 – 20).
Варианты № 16 + i (i = 0, 1, 2, 3, 4).

По данным, приведенным в табл. П. 2.12 и П. 2.13, исследовать влияние на толщину никелевого покрытия детали "экран" химической активности электролита, характеризуемой различными периодами его эксплуатации (фактор A).

Таблица П. 2.12. **Толщина никелевого покрытия детали,** мкм

Номер наблюдения	Химическая активность электролита в период эксплуатации		
	начальный	средний	конечный
1	5,3	6,5	7,8 + 0,2i
2	5,4	6,7	7,9 + 0,2i
3	5,5	6,8	8,0 + 0,2i
4	5,6	7,5	8,3 + 0,2i
5	5,8	7,8	8,4
6	5,9	7,9	8,5

Таблица П. 2.13. **Толщина никелевого покрытия детали**
(комплекс с неравным числом наблюдений)

Номер наблюдения	Химическая активность электролита в период эксплуатации		
	начальный	средний	конечный
1	5,3	6,5	7,8 + 0,2i
2	5,5	7,5	7,9 + 0,2i
3	5,6	7,8	8,0 + 0,2i
4	5,8	7,9	8,3 + 0,2i
5	5,9		8,4
6			8,5

Задача № 5 (варианты № 21 – 25).
Варианты № 21 + i (i = 0, 1, 2, 3, 4).

По данным, приведенным в табл. П. 2.14 и П. 2.15, исследовать влияние коэффициента сменности оборудования (фактор A) на производительность труда.

Таблица П. 2.14. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Номер наблюдения	Коэффициент сменности оборудования		
	1,0 – 1,3	1,3 – 1,6	1,6 – 2,0
1	6,4	6,4	8,0 + 0,2i
2	6,4	6,9	8,1 + 0,2i
3	6,5	7,0	8,3 + 0,2i
4	7,4	8,1	9,0 + 0,2i
5	7,3	8,2	8,8
6	7,5	8,3	8,7

Таблица П. 2.15. **Производительность труда на предприятиях машиностроения** (комплекс с неравным числом наблюдений), тыс. руб./чел.

Номер наблюдения	Коэффициент сменности оборудования		
	1,0 – 1,3	1,3 – 1,6	1,6 – 2,0
1	6,4	6,4	8,0 + 0,2i
2	6,5	6,9	8,1 + 0,2i
3	7,4	7,0	8,3 + 0,2i
4	7,3	8,1	9,0 + 0,2i
5	7,5	8,2	
6		8,3	

Задача № 1 (варианты № 1 – 5).
Варианты № $1 + i$ ($i = 0, 1, 2, 3, 4$).

По данным табл. П. 2.16 исследовать влияние коэффициента сменности оборудования (фактор *A*), фонда заработной платы (фактор *B*) и их взаимодействия на производительность труда.

Таблица П. 2.16. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Коэффициент сменности оборудования	Фонд заработной платы, тыс. руб.					
	1000 – 3000			3000 – 5000		
1 – 1,3	6,3	6,3	6,4	7,2	7,3	7,5
1,3 – 1,6	6,5	6,8	6,9	7,9	7,9	8,0
1,6 – 2,0	7,9	7,9	8,0	$8{,}5 + 0{,}2i$	$8{,}4 + 0{,}2i$	$8{,}4 + 0{,}2i$

Задача № 2 (варианты № 6 – 10).
Варианты № $6 + i$ ($i = 0, 1, 2, 3, 4$).

По данным табл. П. 2.17 исследовать влияние среднесписочной численности промышленно-производственного персонала (фактор *A*), среднегодовой стоимости основных производственных фондов (фактор *B*) и их взаимодействия на производительность труда.

Таблица П. 2.17. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Численность ППП, чел.	Стоимость основных фондов, тыс. руб.					
	1500 – 5000			5000 – 15000		
500 – 1000	4,0	4,1	3,9	5,2	5,6	6,0
1000 – 2000	4,0	5,1	5,2	6,1	6,9	6,5
2000 – 5000	4,5	5,5	6,0	$6{,}5 + 0{,}2i$	$6{,}0 + 0{,}2i$	$5{,}9 + 0{,}2i$

Задача № 3 (варианты № 11 – 15).
Варианты № 11 + i (i = 0, 1, 2, 3, 4).

По данным табл. П. 2.18 исследовать влияние фондовооруженности рабочего (фактор A), производительности труда (фактор B) и их взаимодействия на себестоимость товарной продукции.

Таблица П. 2.18. **Себестоимость товарной продукции,** руб.

Фондовооруженность рабочего, тыс. руб.	Производительность труда, тыс. руб.					
	5 – 6			6 – 7		
15 – 25	65	60	65	62	60	59
25 – 35	62	63	60	58	58	58
35 – 45	52 + i	51 + i	50 + i	51	52	50

Задача № 4 (варианты № 16 – 20).
Варианты № 16 + i (i = 0, 1, 2, 3, 4).

По данным табл. П. 2.19 исследовать влияние периода эксплуатации электролита (фактор A), места расположения точки контроля на поверхности детали (фактор B) и их взаимодействия на толщину никелевого покрытия детали "экран".

Таблица П. 2.19. **Толщина никелевого покрытия детали,** мкм

Период эксплуатации электролита	Место расположения точки контроля (см от центра поверхности детали)					
	12			6		
Начальный	5,3	5,4	5,5	5,6	5,8	5,9
Средний	6,5	6,7	6,8	7,5	7,8	7,9
Конечный	7,8	7,9	8,0	8,3 + 0,2i	8,4 + 0,2i	8,5 + 0,2i

Задача № 5 (варианты № 21 – 25).
Варианты № 21 + i (i = 0, 1, 2, 3, 4).

По данным табл. П. 2.20 исследовать влияние коэффициента сменности работы оборудования (фактор A), фонда заработной платы (фактор B) и их взаимодействия на производительность труда.

Таблица П. 2.20. **Производительность труда на предприятиях машиностроения,** тыс. руб./чел.

Коэффициент сменности оборудования	Фонд заработной платы, тыс. руб.					
	1000 – 3000			3000 – 5000		
1,0 – 1,3	6,4	6,4	6,5	7,4	7,3	7,5
1,3 – 1,6	6,4	6,9	7,0	8,1	8,2	8,3
1,6 – 2,0	8,0	8,1	8,3	9,0 + 0,2i	8,8 + 0,2i	8,7 + 0,2i

П.2.8. ОЦЕНИВАНИЕ ПАРАМЕТРОВ И ПРОВЕРКА ГИПОТЕЗЫ О НОРМАЛЬНОМ ЗАКОНЕ РАСПРЕДЕЛЕНИЯ

По выборочным данным, представленным ниже в 25 вариантах, требуется:

1) построить интервальный вариационный ряд распределения;

2) вычислить выборочные характеристики по вариационному ряду: среднюю арифметическую ($\bar{x}$), центральные моменты ($\hat{\mu}_k$, $k = \overline{1,4}$), дисперсию (s^2), среднее квадратическое отклонение (s), коэффициенты асимметрии ($\widehat{\mathrm{Ac}}$) и эксцесса ($\widehat{\mathrm{Ek}}$), медиану ($\widehat{\mathrm{Me}}$), моду ($\widehat{\mathrm{Mo}}$), коэффициент вариации (V_s);

3) построить гистограмму, полигон и кумуляту;

4) сделать вывод о форме ряда распределения по виду гистограммы и полигона, а также по значениям коэффициентов $\widehat{\mathrm{Ac}}$ и $\widehat{\mathrm{Ek}}$;

5) рассчитать плотность и интегральную функцию теоретического нормального распределения и построить эти кривые на графиках гистограммы и кумуляты соответственно;

6) проверить гипотезу о нормальном законе распределения по критерию согласия Пирсона (χ^2).

Задача № 1 (варианты № 1 – 5).
Варианты № 1 + i (i = 0, 1, 2, 3, 4).

Урожайность пшеницы (ц/га) на полях колхозов района составила:

20,4	19,5	14,3	18,1	25,7	30,1	20,1	18,4	13,5 + i
19,0	13,1	11,5	32,1	33,2	31,5	32,0	29,5	25,1 + i
19,5	19,1	15,1	22,1	21,1	24,5	23,7	13,5	28,1
13,7	20,5	23,9	18,6	22,5	26,1	27,5	27,9	22,4 + i
23,1	23,2	20,1	21,4	25,3	20,5	21,4	24,5	23,5 + i
27,4	33,1	30,1	27,3	23,8	23,1	23,0	26,2	31,5
30,1	25,4	29,3	20,8	23,1	21,3	28,1	23,4	28,5
22,5	20,6	20,5	27,1	24,1	26,1	20,3	29,3	22,1 + i
23,1	25,1	29,1	25,7	25,1	30,7	24,0	21,9	30,1
24,1	25,3	26,1	21,3	24,0	21,3	24,2	21,0	28,4
32,0	24,5	36,5	20,1	23,1	30,4	21,3	22,0	24,3 + i
24,4								

Задача № 2 (варианты № 6 – 10).
Варианты № 6 + i (i = 0, 1, 2, 3, 4).

Производительность труда по предприятиям на 1 работающего (тыс. руб.) за некоторый период составила:

11,70	9,03	13,70	12,31	6,68	5,60	8,06	12,90
7,35	7,76	12,30	5,91	6,23	12,37	11,50	8,69 + i
11,35	13,70	11,11	9,74	12,33	14,75	6,86	12,90
13,90	9,70	12,00	13,56	6,67	12,75	15,33	9,73 + i
11,00	15,30	9,50	11,99	14,40	10,36	13,00	10,60 + i
9,75	10,79	14,10	12,05	11,25	15,67	14,67	15,95
15,21	16,00	12,41	9,02	16,20	9,32	8,81	10,11 + i
13,57	10,32	13,85	13,60	16,60	15,05	12,97	13,60
9,21	17,00	12,80	17,60	10,81	16,95	9,85	10,70 + i
14,90	15,95	13,40	16,80	6,96	12,03	12,00	11,50
12,90	7,39	16,10	9,35	13,75	8,80	13,01	8,64 + i
11,80	10,48	15,85	11,56	12,56	11,67	12,27	12,07
10,51	12,09	12,31	9,76				

Задача № 3 (варианты № 11 – 15).
Варианты № 11 + i (i = 0, 1, 2, 3, 4).

Продолжительность горения электролампочек (ч) следующая:

750	750	756	769	757	767	760	743	745	$759 + 2i$
750	750	739	751	746	758	750	758	753	$747 + 2i$
751	762	748	750	752	763	739	744	764	$755 + 2i$
751	750	733	752	750	763	749	754	745	$747 + 2i$
762	751	758	766	757	769	739	746	750	$753 + 2i$
738	735	760	738	747	752	747	750	746	$748 + 2i$
742	742	758	751	752	762	740	753	758	$754 + 2i$
737	743	748	747	754	754	750	753	754	760
740	756	741	752	747	749	745	757	755	764
756	764	751	759	754	745	752	755	765	762

Задача № 4 (варианты № 16 – 20).
Варианты № 16 + i (i = 0, 1, 2, 3, 4).

Оплата труда колхозников одного из колхозов деньгами и натурой (руб.) за некоторый период времени следующая:

338	336	312	381	302	296	360	342	334	$322 + i$
348	304	323	310	368	314	298	312	322	$350 + i$
304	302	336	334	304	292	324	331	324	$334 + i$
314	338	324	292	298	262	338	331	275	$324 + i$
326	314	312	362	368	324	352	304	302	$332 + i$
314	308	312	381	290	322	326	316	328	$340 + i$
324	320	364	304	340	290	318	332	354	$324 + i$
304	324	356	366	324	332	304	282	330	$314 + i$
342	322	362	298	316	298	332	342	316	326
308	321	302	304	322	296	322	338	324	323

Задача № 5 (варианты № 21 – 25).
Варианты № 21 + i (i = 0, 1, 2, 3, 4).

Высота крышек трубопроводных вентилей (мм) следующая:

109,6	108,9	107,3	105,0	106,5	113,3	109,2	110,9 + i
106,8	108,6	112,9	109,5	107,3	107,6	111,6	115,7
110,3	107,9	107,8	111,8	116,6	108,6	110,9	106,9 + i
106,7	110,8	104,1	110,2	107,3	108,6	108,7	110,3
104,5	106,3	111,2	111,2	107,2	108,6	109,4	113,4
116,6	107,1	104,1	110,8	113,6	116,8	104,7	112,3
114,9	104,7	112,0	112,6	111,8	109,7	105,3	115,5
113,3	112,6	110,4	109,4	112,9	111,3	112,1	110,8 + i
110,4	113,4	111,9	113,5	111,0	108,6	110,2	114,7
110,8	109,1	109,6	111,2	110,3	109,9	109,9	108,6 + i
112,6	111,4	105,1	107,4	106,9	107,8	111,0	107,3 + i
106,9	108,6	109,7	113,3	106,4	112,1	107,9	109,7 + i
114,5	106,1	110,0	104,0				

П. 2.9. МОДЕЛИРОВАНИЕ РЯДОВ ДИНАМИКИ С ИСПОЛЬЗОВАНИЕМ ГАРМОНИЧЕСКОГО АНАЛИЗА

По рядам динамики показателей деятельности предприятий производственного объединения электротехнической промышленности за пятилетку по кварталам (табл. П. 2.21) построить модель временного ряда с учетом тренда и сезонных колебаний. При этом требуется:

1) изобразить график исходного временного ряда X_t;

2) выбрать вид, оценить параметры и построить уравнение тренда $\hat{u}_t$ на графике временного ряда;

3) вычислить остаточную дисперсию и среднюю относительную ошибку аппроксимации;

4) рассчитать оценки параметров α_i и β_i (i = 1, 2, 3, 4) гармонического ряда $V_t = X_t - \hat{U}_t$;

5) получить расчетные значения $\hat{x}_t = \hat{u}_t + \hat{v}_t$ и изобразить их на графике;

6) оценить качество полученной аппроксимации по величинам s_x^2 и $\bar{\delta}_x$;

7) провести содержательную интерпретацию полученной модели.

Показатели производственно-хозяйственной деятельности:

X_1 – доля рабочих в общей численности промышленно-производственного персонала, %;

X_2 – фондовооруженность рабочего, тыс. руб.;

X_3 – выработка товарной продукции на одного рабочего, тыс. руб.;
X_4 – заработная плата одного инженерно-технического работника, руб.;
X_5 – численность промышленно-производственного персонала, тыс. чел.;
X_6 – затраты на рубль товарной продукции, коп.;
X_7 – удельный вес продукции высшей категории качества, %;
X_8 – заработная плата одного рабочего, руб.;
X_9 – фондоемкость товарной продукции;
X_{10} – прибыль от реализации, тыс. руб.;
X_{11} – фондоотдача;
X_{12} – доля ИТР в общей численности промышленно-производственного персонала, %.

Таблица П. 2.21. **Значения показателей производственно-хозяйственной деятельности предприятий и объединения за пятилетку по кварталам**

Год	Квартал	Объединение в целом				Завод № 1			Завод № 2			
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_5	x_2	x_7
		Номер варианта										
		1	2	3	4	5	6	7	8	9	10	11
1981	1	75,5	5,93	2,57	489	1,43	91,8	85,3	401	685	7,55	68,9
	2	75,7	6,03	2,67	498	1,42	87,4	85,0	501	620	9,68	63,8
	3	75,2	6,27	2,69	515	1,39	86,3	83,9	536	607	13,25	71,1
	4	74,4	6,47	2,69	517	1,40	88,1	84,3	531	619	11,14	72,1
1982	1	74,7	7,13	2,80	508	1,39	86,7	85,7	514	632	11,12	64,9
	2	74,6	6,54	2,89	518	1,37	84,7	83,9	553	610	11,72	68,9
	3	76,4	6,42	2,82	513	1,37	82,3	85,1	540	635	11,45	60,5
	4	74,2	6,61	2,84	534	1,38	87,7	82,6	552	670	11,37	57,5
1983	1	73,3	7,12	2,93	520	1,40	86,7	85,7	556	705	13,06	61,4
	2	73,9	7,28	3,02	529	1,37	87,8	83,9	555	710	12,86	62,7
	3	73,8	7,26	3,0	544	1,36	88,0	85,7	561	735	12,65	60,3
	4	73,5	7,49	2,98	527	1,37	93,1	77,2	532	781	12,20	56,0
1984	1	73,4	7,71	3,13	525	1,37	85,0	83,3	541	766	13,23	60,6
	2	73,9	7,76	3,32	546	1,35	89,8	78,9	548	787	13,16	58,4
	3	73,7	7,92	3,20	560	1,33	86,2	73,6	549	778	12,81	58,8
	4	74,0	8,04	3,23	568	1,30	87,8	84,6	548	817	13,16	58,6
1985	1	74,3	9,92	3,16	536	1,30	86,7	84,8	548	954	25,5	54,0
	2	74,2	10,01	3,19	565	1,28	87,3	77,4	559	939	25,4	49,5
	3	73,8	9,96	3,29	574	1,27	87,4	75,0	592	1020	23,57	43,4
	4	73,4	10,07	3,18	582	1,27	88,0	76,5	607	1029	22,48	42,0

Продолжение

Год	Квартал	Завод № 3							Завод № 4			Завод № 5			
		x_8	x_6	x_{10}	x_4	x_{11}	x_2	x_7	x_{12}	x_8	x_7	x_5	x_7	x_6	x_9
		12	13	14	15	16	17	18	19	20	21	22	23	24	25
1981	1	443	73,1	423	393	25,1	12,0	58,7	15,4	436	20,2	3,26	82,7	24,3	3,98
	2	450	73,0	435	406	24,4	12,1	59,8	15,2	450	20,2	3,22	82,7	25,2	4,10
	3	473	69,6	422	415	26,7	11,9	66,1	15,6	450	24,5	3,25	83,0	24,0	3,75
	4	470	76,0	381	411	24,4	12,3	64,8	15,7	443	25,5	3,25	86,3	23,5	4,09
1982	1	467	80,3	297	407	20,9	12,3	73,7	15,5	430	21,7	3,23	90,5	21,5	4,79
	2	487	76,7	333	423	22,1	13,0	75,8	15,3	432	22,5	3,19	90,1	18,1	4,53
	3	501	74,2	397	417	20,2	12,9	72,6	15,1	447	23,0	3,19	90,8	17,4	4,96
	4	520	76,8	332	398	20,4	13,5	77,3	14,3	437	23,6	3,19	88,4	13,2	4,92
1983	1	499	77,2	299	374	21,3	13,3	77,0	14,3	443	21,3	3,21	89,3	20,0	4,67
	2	490	80,8	321	379	18,9	13,2	71,3	14,6	448	20,7	3,18	88,4	18,5	5,31
	3	542	72,5	399	404	21,9	13,6	68,9	14,6	450	21,9	3,14	89,3	19,2	4,57
	4	540	87,4	428	404	21,2	13,6	64,6	14,6	455	24,1	3,16	88,1	23,5	4,72
1984	1	505	76,1	383	409	20,9	14,2	67,5	14,4	446	21,8	3,16	87,1	22,5	4,79
	2	518	73,6	381	413	20,5	14,4	65,1	14,4	486	22,5	3,16	86,3	22,3	4,88
	3	548	69,1	498	424	22,6	15,1	64,3	14,3	469	28,1	3,15	86,3	22,1	4,43
	4	559	80,5	263	407	20,8	15,2	69,8	14,0	473	14,2	3,14	83,4	20,8	4,80
1985	1	516	69,2	522	396	22,1	15,4	75,9	13,9	463	13,0	3,18	84,4	26,5	4,46
	2	539	69,0	552	413	21,9	15,3	75,4	14,2	495	11,9	3,16	81,4	27,1	4,57
	3	553	66,2	601	407	22,7	15,3	70,5	14,4	466	12,2	3,15	82,8	28,6	4,41
	4	565	78,8	430	426	20,2	15,2	66,3	14,8	473	17,6	3,15	83,4	22,1	4,96

ЛИТЕРАТУРА

1. *Айвазян С.А., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Основы моделирования и первичной обработки данных. – М.: Финансы и статистика, 1983. – 471 с.
2. *Айвазян С.А., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Исследование зависимостей. – М.: Финансы и статистика, 1985. – 487 с.
3. *Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Классификация и снижение размерности. – М.: Финансы и статистика, 1989 г. – 607 с.
4. *Болч Б., Хуань К.Дж.* Многомерные статистические методы для экономики. – М.: Финансы и статистика, 1979. – 317 с.
5. *Дубров А.М.* Обработка статистических данных методом главных компонент. – М.: Статистика, 1978. – 135 с.
6. *Епанечников В.А., Цветков А.Н.* Справочник по прикладным программам для микрокалькуляторов. – М.: Финансы и статистика, 1988. – 320 с.
7. *Дьяконов В.П.* Справочник по расчетам на микрокалькуляторах. – М.: Наука, 1985. – 224 с.
8. *Иберла К.* Факторный анализ. – М.: Статистика, 1980. – 398 с.
9. *Мюллер П., Нойман П., Шторм Р.* Таблицы по математической статистике. – М.: Финансы и статистика, 1982. – 271 с.

ОГЛАВЛЕНИЕ

Математико-статистический анализ на программируемых мик-
М34 рокалькуляторах: Справ. пособие/Под ред. В.В.Шуракова. – М.: Финансы и статистика, 1991. –176 с.: ил. – (Статистика и информатика).

ISBN 5-279-00528-2.

Изложены основные методы математико-статистического анализа. Приведены примеры решения экономических задач с помощью корреляционного, регрессионного, компонентного, кластерного, дискриминантного и дисперсионного анализа, а также анализа временных рядов на отечественных программируемых микрокалькуляторах. Более глубокому усвоению методов поможет прилагаемый комплект задач для самостоятельного решения.

Для практических и научных работников, аспирантов и студентов, использующих математико-статистические методы в социально-экономических и технических приложениях.

$М\frac{1404000000-059}{010(01)-91}125-91$ ББК 22.172

В УСЛОВИЯХ РЫНОЧНОЙ ЭКОНОМИКИ ЭФФЕКТИВНОЙ РАБОТЕ ЭКОНОМИСТОВ И СТАТИСТИКОВ ПОМОГУТ КНИГИ СЕРИИ "СТАТИСТИКА И ИНФОРМАТИКА", ВЫХОДЯЩИЕ В ИЗДАТЕЛЬСТВЕ "ФИНАНСЫ И СТАТИСТИКА" В 1991 г.

Информатика: данные, технология, маркетинг/Божко В.П., Брага В.В., Бубнова Н.Г. и др. – 12 л.

Статистическая информационная система – основа формирования данных для всех уровней управления народным хозяйством. Поэтому важно знать, что такое статистическая информатика, какие вычислительные и программно-технологические средства она использует, как они эксплуатируются, как организуется коммерческая деятельность по распространению статистической информации, что такое информационный маркетинг.

Н.Б.Мироносецкий, В.И.Псарев, А.А.Суханов
Территориальные информационно-вычислительные системы. – 15 л.

Территориальная информационно-вычислительная система (ТИВС) рассматривается авторами как эффективный и неотъемлемый компонент территориального управления экономической и социальной сферами жизни региона в условиях рыночной экономики.

Книги будут полезны также студентам, аспирантам и преподавателям экономических специальностей.

В IV кв. 1991 г. В ИЗДАТЕЛЬСТВЕ
"ФИНАНСЫ И СТАТИСТИКА"
В СЕРИИ "СТАТИСТИКА И ИНФОРМАТИКА"
ВЫХОДИТ КНИГА

Ципис Я.Л. Компьютеризация и статистика. – 10 л.

Статистическая информационная система – одна из крупнейших в стране систем обработки данных (СОД). На основе тридцатилетнего опыта работы в этой системе автор рассматривает практические аспекты возникающих проблем и возможные варианты их решения. И хотя это опыт статистической информационной системы, рассматриваемые проблемы и пути развития СОД – общие для информационных систем различного назначения, с ними неизбежно сталкиваются все специалисты, соприкасающиеся с информационным производством.

Основное внимание в книге уделяется проблемам, возникающим на стыках технического, программного и информационного обеспечения, и влиянию этих проблем на технологию применения ЭВМ при обработке информации. Интересен раздел, посвященный развитию архитектуры систем, построенный на основе концепций автора, существенно отличающихся от традиционных.

Автор рассматривает также такие актуальные практические вопросы, как проектирование программного обеспечения, методы управления качеством программного обеспечения, развитие АРМ и связанные с ними проблемы локальных и распределенных сетей; информационный маркетинг и др.

Для специалистов в области проектирования СОД, информатики, работников предприятий и организаций информационного обслуживания.

Коллективные заявки на книги серии "Статистика и информатика" можно направлять по адресу:

101000, Москва, ул. Чернышевского, 7
издательство "Финансы и статистика"

Отдел маркетинга

Справочное издание

Дубров Абрам Моисеевич
Мхитарян Владимир Сергеевич
Трошин Лев Иванович
Масленченко Ирина Викторовна

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ НА ПРОГРАММИРУЕМЫХ МИКРОКАЛЬКУЛЯТОРАХ

Редактор *Л. Н. Вылегжанина*
Худож. редактор *С. Л. Витте*
Техн. редактор *Л. Г. Челышева*
Корректоры *Т. Г. Кочеткова, М. М. Виноградова*
Обложка художника *Ф. Г. Миллера*

Набрано на "Типотайпере"
оператором *И. В. Витте*

ИБ № 2630

Сдано в набор 2.0490. Подписано в печать 22.03.91.
Формат 60x88 1/16. Бум. офсетная. Гарнитура "Литературная".
Печать офсетная. Усл.п.л. 10,78. Усл.кр.-отт. 11,03 Уч.-изд.л. 11,03
Тираж 20 000 экз. Заказ № 325. Цена 1 р. 80 к.

Издательство "Финансы и статистика", 101000, Москва,
ул. Чернышевского, 7.

Отпечатано в типографии им. Котлякова
издательства "Финансы и статистика"
Государственного комитета СССР по печати
195273, г. Ленинград, ул. Руставели, 13

1 р. 80 к.

СТАТИСТИКА И ИНФОРМАТИКА

СТАТИСТИКИ, ЭКОНОМИСТЫ, СТУДЕНТЫ И АСПИРАНТЫ! ВАМ АДРЕСОВАНА НОВАЯ СЕРИЯ "СТАТИСТИКА И ИНФОРМАТИКА". КНИГИ СЕРИИ ПОМОГУТ ВАМ УГЛУБИТЬ И РАСШИРИТЬ ЗНАНИЯ В ОБЛАСТИ ОБРАБОТКИ И АНАЛИЗА МАССОВЫХ СТАТИСТИЧЕСКИХ ДАННЫХ С ПОМОЩЬЮ СОВРЕМЕННЫХ СРЕДСТВ ВТ И ЭКОНОМИКО-МАТЕМАТИЧЕСКИХ МЕТОДОВ И МОДЕЛЕЙ.

ФС