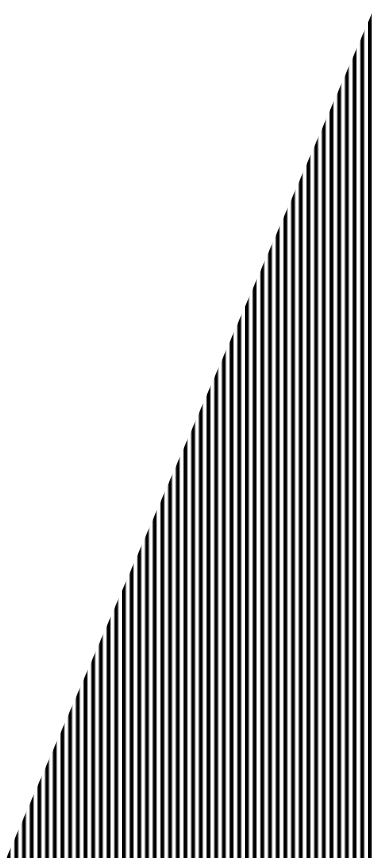


МАТЕМАТИЧЕСКАЯ СТАТИСТИКА



◆ ИЗДАТЕЛЬСТВО ТГТУ ◆

Министерство образования и науки Российской Федерации

Государственное образовательное учреждение
высшего профессионального образования

"Тамбовский государственный технический университет"

**МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА**

Методические разработки и контрольные задания
для студентов 2 курса инженерно-технических специальностей



Тамбов
Издательство ТГТУ
2005

УДК 519.24:51.001.57
ББК В1я73
П 39

Утверждено Редакционно-издательским советом университета

Рецензент
Кандидат физико-математических наук, доцент
Л.И. Ткач

П
39 Математическая статистика: Методические разработки и контрольные задания / Авт.-сост.: С.В. Плотникова. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2005. 52 с.

Приведены краткие теоретические сведения по основным разделам математической статистики: точечные и интервальные оценки параметров распределения, статистическая проверка статистических гипотез, линейная и нелинейная регрессия. Предложены варианты для выполнения курсовых заданий, имеющие прикладной характер.

Предназначены для студентов 2 курса дневного отделения инженерно-технических специальностей.

УДК 519.24:51.001.57
ББК В1я73

© Плотникова С.В.
© Тамбовский государственный
технический университет
(ТГТУ), 2005

Учебное издание

МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Методические разработки и контрольные задания

Автор-составитель
ПЛОТНИКОВА Светлана Валерьевна

Редактор Т.М. Федченко

Компьютерное макетирование И.В. Евсеевой

Подписано к печати 18.02.2005

Гарнитура Times New Roman. Формат 60 × 84/16. Бумага офсетная.

Печать офсетная. Объем: 3,02 усл. печ. л.; 3,00 уч.-изд. л.

Тираж 300 экз. С. 126

Издательско-полиграфический центр ТГТУ
392000, Тамбов, Советская, 106, к. 14

1 ОСНОВНЫЕ ПОНЯТИЯ. ВАРИАЦИОННЫЕ РЯДЫ И ИХ ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ

Генеральной совокупностью называется вся подлежащая изучению совокупность объектов (наблюдений). *Выборочной совокупностью*, или просто *выборкой*, называется совокупность случайно отобранных из генеральной совокупности объектов.

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность. *Бесповторной* называют выборку, при которой отобранный объект в генеральную совокупность не возвращается. На практике обычно пользуются бесповторным случайным отбором.

Объемом совокупности (выборочной или генеральной) называется число объектов в этой совокупности.

Различные наблюдаемые значения признака (случайной величины X) называют *вариантами* (обозначаем их через x_i).

Числа, показывающие, сколько раз встречаются варианты из данного интервала, называются *частотами* (обозначаются n_i), тогда объем выборки можно определить как $n = \sum n_i$. Отношение частот к объему выборки $w_i = n_i/n$ называют относительными частотами.

Последовательность вариантов, записанных в порядке возрастания или убывания с соответствующими им частотами (или относительными частотами) называется *вариационным рядом*.

Вариационный ряд называется *дискретным*, если любые его варианты отличаются на постоянную величину, и *непрерывным (интервальным)*, если его значения могут отличаться одно от другого на сколь угодно малую величину.

Для наглядности строят различные графики статистического распределения, и, в частности, полигон и гистограмму.

Полигон, как правило, служит для изображения дискретного вариационного ряда, и представляет собой ломаную, в которой концы отрезков имеют координаты $(x_i, n_i$, или $x_i, w_i)$.

Гистограмма служит для изображения интервальных вариационных рядов и представляет собой ступенчатую фигуру из прямоугольников с основаниями, равными интервалам значений признака (x_{i-1}, x_i) , и высотами, равными частотам n_i (или относительным частотам w_i) интервалов. Если соединить середины верхних оснований прямоугольников отрезками прямой, то можно получить полигон того же распределения.

2 ТОЧЕЧНЫЕ И ИНТЕРВАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

Пусть требуется изучить количественный признак X генеральной совокупности. Предполагается, что из теоретических соображений известно, какой именно вид имеет распределение признака X , однако неизвестно значение некоторого параметра θ , характеризующего это распределение (например, это параметр λ в распределении Пуассона, или параметр p в биномиальном распределении). Обычно в распоряжении исследователя имеются лишь данные выборки, т.е. значения количественного признака x_1, x_2, \dots, x_n полученные в результате n независимых наблюдений. Через эти данные и выражают оцениваемый параметр. Значения x_1, x_2, \dots, x_n можно рассматривать как реализации n независимых случайных величин X_1, X_2, \dots, X_n , каждая из которых имеет тот же закон распределения, что и сама случайная величина X .

Статистической оценкой θ_n^* неизвестного параметра θ генеральной совокупности называют функцию наблюдений над случайной величиной X : $\theta_n^* = \theta_n^*(X_1, X_2, \dots, X_n)$. Поскольку X_1, X_2, \dots, X_n – случайные величины, то и оценка θ_n^* также является случайной величиной, в отличие от оцениваемого параметра θ , который является величиной неслучайной, детерминированной.

Оценка θ_n^* называется *несмещенной*, если ее математическое ожидание равно оцениваемому параметру, т.е. $M(\theta_n^*) = \theta$. В противном случае оценка называется смещенной.

Несмещенная оценка θ_n^* называется *эффективной*, если она имеет наименьшую дисперсию среди всех возможных несмещенных оценок параметра θ , вычисленных по выборкам одного и того же объема n .

Оценка θ_n^* называется *состоятельной*, если она удовлетворяет закону больших чисел, т.е. сходится по вероятности к оцениваемому параметру:

$$\lim_{n \rightarrow \infty} P(|\theta_n^* - \theta| < \varepsilon) = 1 \text{ для любого } \varepsilon > 0.$$

В качестве статистических оценок параметров генеральной совокупности желательно использовать оценки, удовлетворяющие одновременно требованиям несмещенности, состоятельности и эффективности. Однако иногда для простоты расчетов целесообразно применять оценки, обладающие большей дисперсией по сравнению с эффективными оценками, или незначительно смещенные оценки и т.п.

Генеральной средней \bar{x} конечной генеральной совокупности называют среднее арифметическое значений признака генеральной совокупности. В случае бесконечной генеральной совокупности под генеральной средней понимается математическое ожидание распределения признака X .

Пусть из генеральной совокупности объема n отобрана случайная выборка $X_1, X_2, \dots, X_n \dots$. Определим *выборочную среднюю* \bar{x}_b , которая является несмещенной и состоятельной оценкой генеральной средней: $\bar{x}_b = (X_1 + X_2 + \dots + X_n)/n$. Отметим, что на практике в целях упрощения расчетов используют и другие оценки генеральной средней, например, медиану – значение признака, приходящееся на середину вариационного ряда. При нормальном распределении признака в генеральной совокупности медиана является несмещенной оценкой генеральной средней, однако эффективной оценкой является выборочная средняя. После проведения наблюдений получаем реализацию случайной величины \bar{x}_b , которую для простоты также обозначим \bar{x}_b .

Если все значения x_1, x_2, \dots, x_n признака выборки объема n различны, то

$$\bar{x}_b = (x_1 + x_2 + \dots + x_n) / n. \quad (2.1)$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причем $n = \sum n_i$, то

$$\bar{x}_b = (n_1 x_1 + n_2 x_2 + \dots + n_k x_k) / n. \quad (2.2)$$

В качестве характеристики разброса значений количественного признака X вокруг своего среднего значения используется дисперсия. В случае конечной генеральной совокупности *генеральной дисперсией* $D = \sigma^2$ называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения \bar{x} . В случае бесконечной генеральной совокупности под генеральной дисперсией понимается дисперсия признака X в генеральной совокупности.

Выборочной дисперсией D_b называют среднее арифметическое квадратов отклонений наблюдаемых значений признака от их среднего значения \bar{x}_b . После проведения наблюдений получаем реализацию случайной величины D_b , конкретное значение которой находим по следующим формулам.

Если все значения x_1, x_2, \dots, x_n признака выборки объема n различны, то

$$D_b = \left(\sum_{i=1}^n (x_i - \bar{x}_b)^2 \right) / n. \quad (2.3)$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причем $n = \sum n_i$, то

$$D_b = \left(\sum_{i=1}^k n_i (x_i - \bar{x}_b)^2 \right) / n. \quad (2.4)$$

Если в качестве оценки генеральной дисперсии использовать выборочную дисперсию, то эта оценка будет приводить к систематическим ошибкам, давая заниженное значение генеральной дисперсии. Объясняется это тем, что выборочная дисперсия является смещенной оценкой генеральной дисперсии, т.е.

$$M(D_b) = \frac{n-1}{n} D.$$

Несмещенной оценкой генеральной дисперсии является *исправленная выборочная дисперсия*

$$D_B^* = s^2 = nD_B / (n - 1). \quad (2.5)$$

Если все значения x_1, x_2, \dots, x_n признака выборки объема n различны, то

$$D_B^* = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^2}{n - 1}. \quad (2.6)$$

Если же значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , то

$$D_B^* = s^2 = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n - 1}. \quad (2.7)$$

Обе предложенные оценки – выборочная дисперсия и исправленная выборочная дисперсия – являются состоятельными оценками генеральной дисперсии, и разница между ними заметна лишь при небольшом числе наблюдений n . При $n > 30$ в качестве оценки для D вполне можно использовать D_B . В случае нормально распределенной генеральной совокупности обе эти оценки также являются асимптотически эффективными, т.е. при $n \rightarrow \infty$ они будут стремиться к эффективной оценке.

Точечной называют оценку параметра θ генеральной совокупности, которая определяется одним числом. Оценки, рассмотренные выше – точечные. Однако точечная оценка θ_n^* является лишь приближенным значением неизвестного параметра θ , и при выборке малого объема точечная оценка может значительно отличаться от оцениваемого параметра.

Чтобы получить представление о точности и надежности оценки θ_n^* параметра θ пользуются интервальными оценками.

Интервальной оценкой параметра θ называют числовой интервал $(\theta_n^{*(1)}, \theta_n^{*(2)})$, который с заданной вероятностью γ накрывает неизвестное значение параметра θ (рис. 2.1). Такой интервал называется *доверительным*, а вероятность γ называется *доверительной вероятностью*, или *надежностью* оценки.

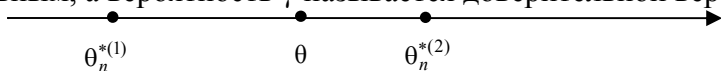


Рис. 2.1

Границы доверительного интервала и его длина находятся по выборочным данным, и являются случайными величинами. Величина доверительного интервала уменьшается с ростом объема выборки n и увеличивается с ростом доверительной вероятности γ .

Если количественный признак генеральной совокупности X имеет нормальное распределение, то доверительный интервал для математического ожидания имеет вид

$$\bar{x} \in (\bar{x}_B - \delta; \bar{x}_B + \delta). \quad (2.8)$$

В случае, когда генеральная дисперсия $D = \sigma^2$ является известной величиной (например, это заданная заранее ошибка измерительного прибора), то точность оценки δ находится по формуле

$$\delta = \frac{t\sigma}{\sqrt{n}}, \quad (2.9)$$

где число t определяется из равенства $\Phi(t) = \gamma/2$, т.е. по таблице функции Лапласа находят значение аргумента t , которому соответствует значение функции Лапласа $\gamma/2$.

В случае, когда генеральная дисперсия неизвестна, а известна лишь ее исправленная выборочная оценка $D_B^* = s^2$, то точность оценки δ находится по формуле

$$\delta = \frac{T(1-\gamma; n-1)s}{\sqrt{n}}, \quad (2.10)$$

где значение числа $T(1-\gamma; n-1)$ определяется по таблице критических точек распределения Стьюдента при уровне вероятности $\alpha = 1-\gamma$ и числе степеней свободы $n-1$.

Доверительный интервал для среднеквадратического отклонения σ нормального распределения имеет вид

$$\sigma \in \left(\frac{s\sqrt{n-1}}{\chi_2}; \frac{s\sqrt{n-1}}{\chi_1} \right), \quad (2.11)$$

где значения χ_1^2 и χ_2^2 находятся по таблице критических точек распределения χ^2 при числе степеней свободы $n-1$ и уровнях вероятности $(1+\gamma)/2$ и $(1-\gamma)/2$ соответственно.

3 Проверка гипотез

Статистической гипотезой называется любое предположение о виде неизвестного распределения или о параметрах закона распределения.

Например, статистическими являются гипотезы:

- 1) генеральная совокупность распределена по нормальному закону;
- 2) дисперсии двух нормальных совокупностей равны между собой;
- 3) математическое ожидание признака X нормально распределенной генеральной совокупности равно 25,5.

Нулевой (основной) называют H_0 . Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу. Если выдвинутая гипотеза H_0 будет отвергнута, то имеет место противоречащая ей гипотеза H_1 . Гипотеза H_1 называется конкурирующей (альтернативной).

Для проверки нулевой гипотезы используют специально подобранную случайную величину (статистический критерий). В зависимости от вида распределения критерий обозначают через U (если он распределен по нормальному закону), T (если он распределен по закону Стьюдента), F (если он распределен по закону Фишера–Снедекора) и т.д. Здесь в целях общности обозначим статистический критерий через K . Наблюдаемым значением критерия называют значение критерия, вычисленное по выборкам, т.е. $K_{\text{набл}} = K(x_1, x_2, \dots, x_n)$.

После выбора критерия множество всех его возможных значений разбивают на два подмножества: одно из них содержит значения критерия, при которых нулевая гипотеза отвергается (*критическая область*), а другое содержит те значения критерия, при которых гипотеза принимается (*область принятия гипотезы*).

Основной принцип проверки статистических гипотез можно сформулировать так: если наблюдаемое значение критерия принадлежит критической области, то нулевую гипотезу отвергают в пользу конкурирующей гипотезы; если наблюдаемое значение критерия принадлежит области принятия гипотезы, то нулевую гипотезу принимают. При этом возможны четыре случая (табл. 3.1).

Таблица 3.1

Гипотеза H_0	Принимается	Отвергается
Верна	Правильное решение	Ошибка 1-го рода
Не верна	Ошибка 2-го рода	Правильное решение

Вероятность α допустить ошибку 1-го рода, т.е. отвергнуть гипотезу H_0 , когда она верна, называется *уровнем значимости критерия*.

Вероятность допустить ошибку 2-го рода, т.е. принять гипотезу H_0 , когда она неверна, обозначают β . Вероятность $1-\beta$ не допустить ошибку 2-го рода называется *мощностью критерия*.

Вероятности ошибок 1-го и 2-го рода определяются выбором критической области. При фиксированном объеме выборки уменьшение ошибки 1-го рода α неизбежно приводит к увеличению ошибки 2-го рода β , и наоборот. Лишь при увеличении объема выборки возможно одновременное уменьшение вероятностей α и β .

Поскольку критерий K – одномерная случайная величина, все ее возможные значения принадлежат

некоторому интервалу, поэтому критическую область и область принятия гипотезы можно разделить точками. *Критическими точками* $K_{кр}$ называют точки, отделяющие критическую область от области принятия гипотезы.

Различают одностороннюю (правостороннюю и левостороннюю) и двустороннюю критические области.

Правосторонней называют критическую область, определяемую неравенством $K > K_{кр}$ (рис. 3.1, а). При этом $P(K > K_{кр}) = \alpha$.

Левосторонней называют критическую область, определяемую неравенством $K < K_{кр}$ (рис. 3.1, б). При этом $P(K < K_{кр}) = \alpha$.

Двусторонней называют критическую область, определяемую неравенствами $K < K_{кр1}$ и $K > K_{кр2}$ (рис. 3.1, в). При этом $P(K > K_{кр2}) = P(K < K_{кр1}) = \alpha/2$.

Рассмотрим способы проверки некоторых наиболее часто встречающихся гипотез.

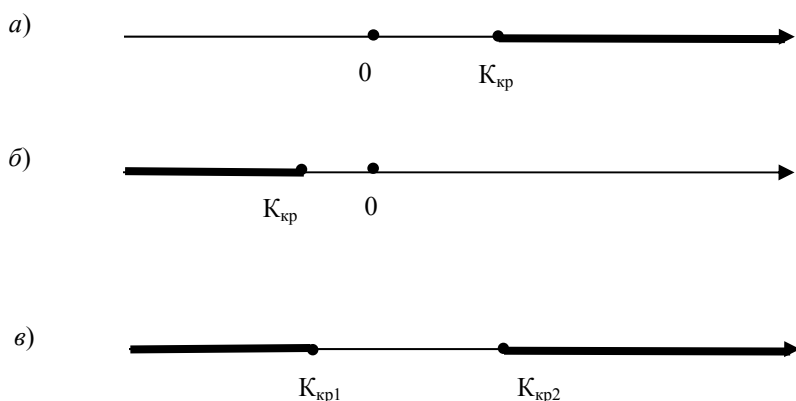


Рис. 3.1

3.1 Гипотеза о равенстве генеральной средней нормальной совокупности заданному числовому значению

Пусть генеральная совокупность X распределена нормально, причем имеются основания предполагать, что генеральная средняя этой совокупности \bar{x} равна некоторому значению a (например, если X – размер деталей, изготавливаемых станком-автоматом, то a – заданный проектный размер).

Предполагаем, что дисперсия генеральной совокупности $D = \sigma^2$ известна (например, может быть найдена теоретически, или вычислена по выборке большого объема). Кроме того, по произведенной выборке объема n найдена выборочная средняя $\bar{x}_в$. Требуется по выборочной средней при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x} = a$.

В качестве критерия проверки нулевой гипотезы принимают случайную величину

$$U = \frac{(\bar{x} - a)\sqrt{n}}{\sigma}, \quad (3.1)$$

которая имеет нормальное распределение.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x} = a$, нужно вычислить наблюдаемое значение критерия

$$U_{набл} = \frac{(\bar{x}_в - a)\sqrt{n}}{\sigma}. \quad (3.2)$$

1) При конкурирующей гипотезе $H_1: \bar{x} \neq a$ критическую точку $U_{кр}$ находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - \alpha) / 2$. Критическая область в этом случае является двусторонней.

Если $|U_{набл}| < U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2) При конкурирующей гипотезе $H_1: \bar{x} > a$ критическую точку $U_{кр}$ правосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - 2\alpha) / 2$.

Если $U_{набл} < U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3) При конкурирующей гипотезе $H_1: \bar{x} < a$ критическую точку $U_{кр}$ левосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - 2\alpha) / 2$.

Если $U_{набл} > -U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

Предположим теперь, что дисперсия генеральной совокупности $D = \sigma^2$ неизвестна, а известна только ее исправленная выборочная оценка $D_b^* = s^2$. При выборке небольшого объема (менее 30 наблюдений) нахождение критической точки по таблице функции Лапласа может привести к существенной погрешности. В качестве критерия проверки нулевой гипотезы тогда принимают случайную величину

$$T = \frac{(\bar{x} - a)\sqrt{n}}{s}, \quad (3.3)$$

имеющую распределение Стьюдента с $n - 1$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x} = a$, нужно вычислить наблюдаемое значение критерия

$$T_{набл} = \frac{(\bar{x}_в - a)\sqrt{n}}{s}. \quad (3.4)$$

1) При конкурирующей гипотезе $H_1: \bar{x} \neq a$ критическую точку $T_{кр}(\alpha, n - 1)$ находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности α . Если $|T_{набл}| < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2) При конкурирующей гипотезе $H_1: \bar{x} > a$ критическую точку $T_{кр}(2\alpha, n - 1)$ правосторонней критической области находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности 2α .

Если $T_{набл} < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3) При конкурирующей гипотезе $H_1: \bar{x} < a$ критическую точку $T_{кр}(2\alpha, n - 1)$ левосторонней критической области находим по таблице критических точек распределения Стьюдента при $n - 1$ степенях свободы и вероятности 2α .

Если $T_{набл} > -T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3.2 Гипотеза о равенстве генеральной дисперсии нормальной совокупности заданному числовому значению

Пусть генеральная совокупность распределена нормально, причем имеются основания предполагать, что генеральная дисперсия равна некоторому заданному значению σ_0^2 . Пусть по выборке объема n получена исправленная выборочная дисперсия $D_b^* = s^2$. Требуется по исправленной выборочной дисперсии при заданном уровне значимости α проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2$.

На практике рассматриваемая гипотеза проверяется, если нужно проверить точность настройки оборудования, устойчивость технологических процессов, точность приборов и т.п.

Для проверки гипотезы H_0 используют случайную величину $(n-1)s^2 / \sigma_0^2$, которая имеет распределение

χ^2 с $n - 1$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \sigma^2 = \sigma_0^2$, нужно вычислить наблюдаемое значение критерия

$$\chi_{\text{набл}}^2 = \frac{(n-1)s^2}{\sigma^2}. \quad (3.5)$$

1) При конкурирующей гипотезе $H_1: \sigma^2 \neq \sigma_0^2$ критическая область является двусторонней. Критические точки находим по таблице критических точек распределения χ^2 с $n - 1$ степенями свободы: левую критическую точку при вероятности $1 - \alpha/2$, а правую критическую точку при вероятности $\alpha/2$. Таким образом, получим точки $\chi_{\text{лев.кр}}^2(1 - \alpha/2, n - 1)$ и $\chi_{\text{пр.кр}}^2(\alpha/2, n - 1)$. Если $\chi_{\text{лев.кр}}^2 < \chi_{\text{набл}}^2 < \chi_{\text{пр.кр}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2) При конкурирующей гипотезе $H_1: \sigma^2 > \sigma_0^2$ критическую точку $\chi_{\text{кр}}^2(\alpha, n - 1)$ правосторонней критической области находим по таблице критических точек распределения χ^2 с $n - 1$ степенями свободы при вероятности α . Если $\chi_{\text{набл}}^2 < \chi_{\text{кр}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3) При конкурирующей гипотезе $H_1: \sigma^2 < \sigma_0^2$ критическую точку $\chi_{\text{кр}}^2(1 - \alpha, n - 1)$ левосторонней критической области находим по таблице критических точек распределения χ^2 с $n - 1$ степенями свободы при вероятности $1 - \alpha$. Если $\chi_{\text{набл}}^2 > \chi_{\text{кр}}^2$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3.3 Гипотеза о равенстве двух средних нормальных генеральных совокупностей

Пусть генеральные совокупности X_1 и X_2 распределены нормально, причем генеральные средние этих совокупностей \bar{x}_1 и \bar{x}_2 нам неизвестны. По произведенным выборкам объемов n_1 и n_2 найдены выборочные средние $\bar{x}_{в1}$ и $\bar{x}_{в2}$.

Предполагаем, что дисперсии обеих генеральных совокупностей известны, и равны σ_1^2 и σ_2^2 (этот же способ можно применять, если известны только выборочные дисперсии, но объемы выборок достаточно большие). Требуется при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$.

Сравнение средних двух совокупностей имеет большое практическое значение. Часто встречается случай, когда средний результат одной серии экспериментов отличается от среднего результата другой серии. Возникает вопрос, можно ли объяснить обнаруженное расхождение средних неизбежными случайными ошибками эксперимента, или оно вызвано некоторыми закономерностями. Задача сравнения средних возникает также при выборочном контроле качества изделий, изготовленных на разных установках или при разных технологических режимах.

В качестве критерия проверки нулевой гипотезы принимают случайную величину

$$U = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}, \quad (3.6)$$

которая имеет нормальное распределение.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$, нужно вычислить наблюдаемое значение критерия

$$U_{\text{набл}} = \frac{\bar{x}_{в1} - \bar{x}_{в2}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}. \quad (3.7)$$

1) При конкурирующей гипотезе $H_1: \bar{x}_1 \neq \bar{x}_2$ критическую точку $U_{\text{кр}}$ находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1 - \alpha)/2$. Критическая область в этом случае является двусторонней.

Если $|U_{\text{набл}}| < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2) При конкурирующей гипотезе $H_1: \bar{x}_1 > \bar{x}_2$ критическую точку $U_{кр}$ правосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - 2\alpha) / 2$.

Если $U_{набл} < U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3) При конкурирующей гипотезе $H_1: \bar{x}_1 < \bar{x}_2$ критическую точку $U_{кр}$ левосторонней критической области находим по таблице функции Лапласа из условия $\Phi(U_{кр}) = (1 - 2\alpha) / 2$.

Если $U_{набл} > -U_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

Приведенное выше правило можно также применять, если генеральные совокупности не имеют нормального распределения, и их дисперсии неизвестны, а выборки имеют большой объем (не менее 30 каждая) и независимы. В этом случае наблюдаемое значение критерия

$$U_{набл} = \frac{\bar{x}_{в1} - \bar{x}_{в2}}{\sqrt{s_1^2 / n_1 + s_2^2 / n_2}} . \quad (3.8)$$

Предположим теперь, что дисперсии обеих генеральных совокупностей неизвестны, а известны только их исправленные выборочные оценки $D_{в1}^* = s_1^2$ и $D_{в2}^* = s_2^2$, а выборки имеют небольшой объем (меньше 30). Предполагается, что дисперсии двух генеральных совокупностей одинаковы (например, дисперсии определяются ошибкой измерительного прибора). Если же нет оснований считать дисперсии одинаковыми, то, прежде чем сравнивать средние, следует проверить гипотезу о равенстве генеральных дисперсий, пользуясь критерием Фишера–Снедекора. В этом случае в качестве критерия проверки нулевой гипотезы принимают случайную величину

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} , \quad (3.9)$$

имеющую распределение Стьюдента с $n_1 + n_2 - 2$ степенями свободы.

Для того, чтобы при заданном уровне значимости α проверить нулевую гипотезу $H_0: \bar{x}_1 = \bar{x}_2$, нужно вычислить наблюдаемое значение критерия

$$T_{набл} = \frac{\bar{x}_{в1} - \bar{x}_{в2}}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} . \quad (3.10)$$

1) При конкурирующей гипотезе $H_1: \bar{x}_1 \neq \bar{x}_2$ критическую точку $T_{кр}(\alpha, n_1 + n_2 - 2)$ находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности α . Если $|T_{набл}| < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

2) При конкурирующей гипотезе $H_1: \bar{x}_1 > \bar{x}_2$ критическую точку $T_{кр}(2\alpha, n_1 + n_2 - 2)$ правосторонней критической области находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности 2α .

Если $T_{набл} < T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

3) При конкурирующей гипотезе $H_1: \bar{x}_1 < \bar{x}_2$ критическую точку $T_{кр}(2\alpha, n_1 + n_2 - 2)$ левосторонней критической области находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности 2α .

Если $T_{набл} > -T_{кр}$, то нет оснований отвергнуть нулевую гипотезу. В противном случае нулевую гипотезу отвергают.

4 ЛИНЕЙНАЯ И НЕЛИНЕЙНАЯ РЕГРЕССИЯ. КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

В естественных науках часто идет речь о функциональной зависимости, когда каждому значению одной переменной соответствует вполне определенное значение другой переменной. Однако на практике между переменными величинами существуют зависимости, когда каждому значению одной переменной может соответствовать множество значений другой переменной, имеющее определенное распределение. Такая зависимость называется статистической.

Статистические связи между переменными можно изучать методами корреляционного и регрессионного анализа. Основной задачей регрессионного анализа является установление формы и изучение зависимости между переменными, корреляционного анализа – выявление связи между случайными переменными и оценка ее тесноты.

В регрессионном анализе рассматривается зависимость случайного результативного признака y от неслучайных факторных признаков x_1, x_2, \dots, x_n . В случае единственного факторного признака x уравнение взаимосвязи двух переменных может быть представлено в виде

$$y = \varphi(x) + \varepsilon,$$

где ε – случайная величина, причем предполагается, что $M(\varepsilon) = 0$, а дисперсия постоянна и не зависит от x . Переменную ε называют *возмущающей переменной* или просто *возмущением*.

В зависимости от вида функции $\varphi(x)$ различают следующие виды регрессий: линейную, гиперболическую, показательную, степенную, логарифмическую, параболическую и т.д.

Предположим, что для оценки параметров регрессии взята выборка, содержащая n пар значений (x_i, y_i) , где $i = 1, 2, \dots, n$. Оценкой предложенных выше уравнений регрессии являются выборочные уравнения регрессии:

- *линейное* $\hat{y} = a_0 + a_1 x$;
- *гиперболическое* $\hat{y} = a_0 + a_1 / x$;
- *показательное* $\hat{y} = a_0 a_1^x$;
- *степенное* $\hat{y} = a_0 x^{a_1}$;
- *логарифмическое* $\hat{y} = a_0 + a_1 \ln x$;
- *параболическое* $\hat{y} = a_0 + a_1 x + a_2 x^2$,

где параметры a_0, a_1, a_2 являются точечными оценками соответствующих параметров исходного уравнения и могут быть найдены на основе метода наименьших квадратов.

Сущность метода наименьших квадратов заключается в нахождении параметров модели a_0, a_1 , при которых минимизируется сумма квадратов отклонений эмпирических (фактических) значений результативного признака от теоретических, полученных по выборочному уравнению регрессии

$$S = \sum (y_i - \hat{y}_i)^2 \rightarrow \min.$$

Для линейной модели

$$S = \sum (y_i - a_0 - a_1 x_i)^2 \rightarrow \min. \quad (4.1)$$

Функция двух переменных $S(a_0, a_1)$ может достигнуть экстремума в том случае, когда ее частные производные равны нулю. Вычисляя эти частные производные, получим систему уравнений для нахождения параметров a_0, a_1 линейного уравнения регрессии $\hat{y} = a_0 + a_1 x$

$$\begin{cases} na_0 + a_1 \sum x_i = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 = \sum x_i y_i. \end{cases} \quad (4.2)$$

В случае, когда возмущающая переменная ε имеет нормальное распределение, коэффициенты a_0, a_1 , полученные методом наименьших квадратов для линейной регрессии, являются несмещенными эффективными оценками параметров α_0, α_1 исходного уравнения.

Для параболического уравнения регрессии $\hat{y} = a_0 + a_1 x + a_2 x^2$ система уравнений для нахождения параметров a_0, a_1, a_2 имеет вид

$$\begin{cases} na_0 + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i; \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum x_i y_i; \\ a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum x_i^2 y_i. \end{cases} \quad (4.3)$$

Заметим, что все предложенные выше виды нелинейных регрессий (кроме параболической) могут быть сведены к линейной путем какой-либо замены переменной. Для гиперболической регрессии вводится переменная $x' = 1/x$, для логарифмической регрессии $x' = \ln x$, уравнения показательной и степенной регрессии предварительно логарифмируют.

Для показательного уравнения получаем $\ln \hat{y} = \ln a_0 + x \ln a_1$, далее делаем замены $\hat{y}' = \ln \hat{y}$, $a'_0 = \ln a_0$, $a'_1 = \ln a_1$. Таким образом, получаем линейное уравнение регрессии $\hat{y}' = a'_0 + a'_1 x$. Найдя коэффициенты этого уравнения методом наименьших квадратов, получим коэффициенты исходного уравнения $a_0 = e^{a'_0}$, $a_1 = e^{a'_1}$.

Для степенного уравнения получаем $\ln \hat{y} = \ln a_0 + a_1 \ln x$, далее делаем замены $x' = \ln x$, $\hat{y}' = \ln \hat{y}$, $a'_0 = \ln a_0$. Таким образом, получаем линейное уравнение регрессии $\hat{y}' = a'_0 + a_1 x'$. Найдя коэффициенты этого уравнения методом наименьших квадратов, вычислим и коэффициент исходного уравнения $a_0 = e^{a'_0}$.

Регрессионную модель удобно представлять графически. Рассмотрим выборку объема 7 (табл. 4.1).

X	1	4	7	11	15	17	22
Y	3	4	10	14	18	24	30

Отложим на координатной плоскости точки $P_i(x_i, y_i)$, ($i = 1, 2, \dots, n$) (рис. 4.1). Полученный график называется диаграммой рассеивания. Видим, что точки группируются около некоторой прямой $y = a_0 + a_1 x$.

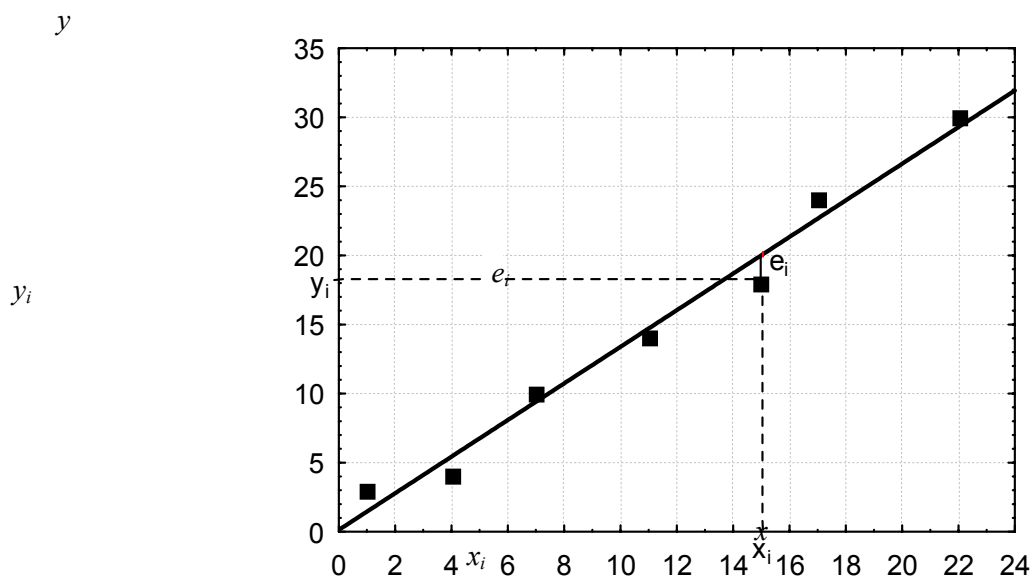


Рис. 4.1

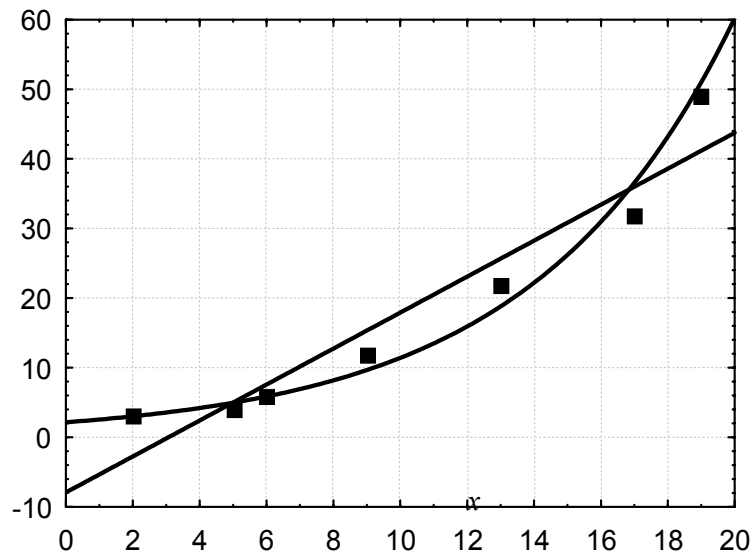


Рис. 4.2

Найдя коэффициенты этой прямой $a_0 = 0,876$; $a_1 = 1,284$ на этом же графике проведем линию, соответствующую уравнению регрессии. Отклонение точки P_i от оцениваемой линии, измеренное по вертикали, будет равно

$$e_i = (y_i - a_0 - a_1 x_i).$$

Тогда функцию $S(a_0, a_1)$, определенную по (4.1), можно представить как $S = \sum e_i^2$, т.е. если фактические данные точно лежали бы на некоторой прямой, то $S = 0$, а чем дальше точки отклоняются от линии регрессии, тем больше значение S .

Построив диаграмму рассеяния, можно подобрать вид уравнения регрессии. На рис. 4.2 для одних и тех же экспериментальных точек построены линейная и показательная регрессии. Видим, что экспериментальные точки "ближе" подходят к линии $y = a_0 a_1^x$, чем к прямой. Следовательно, можно сделать вывод, что показательная регрессия более адекватно описывает фактические данные, чем линейная.

Однако по графику можно только приближенно сделать вывод о качестве той или иной модели. Существуют способы более точной оценки адекватности (значимости) уравнения регрессии. Проверить значимость уравнения регрессии – значит установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным.

Для проверки значимости уравнения регрессии на уровне значимости α вычисляют наблюдаемое значение случайной величины

$$F_{\text{набл}} = \frac{\sigma_y^2 (n-2)}{\sigma_{\text{ост}}^2}, \quad (4.4)$$

где *остаточная дисперсия* $\sigma_{\text{ост}}^2$ и *дисперсия уравнения регрессии* σ_y^2 находятся по формулам

$$\begin{aligned} \sigma_{\text{ост}}^2 &= \frac{\sum (y_i - \hat{y}_i)^2}{n-1}; \\ \sigma_y^2 &= \frac{\sum (\bar{y} - \hat{y}_i)^2}{n-1}. \end{aligned} \quad (4.5)$$

Учитывая смысл величин $\sigma_{\text{ост}}^2$ и σ_y^2 , можно сказать, что значение $F_{\text{набл}}$ показывает, насколько лучше уравнение регрессии оценивает значение результативного признака по сравнению с его средней. Далее находим критическое значение критерия $F(\alpha, 1; n-2)$ по таблице критических точек распределения Фишера при $k_1 = 1$, $k_2 = n-2$ степенях свободы и уровне значимости α . Если $F_{\text{набл}} > F(\alpha, 1; n-2)$, то уравнение регрессии признается значимым, в противном случае уравнение регрессии признается незначимым, т.е. статистически подтверждается отсутствие линейной связи между факторным и результативным признаком.

По расположению точек на диаграмме рассеяния не всегда можно принять окончательное решение о виде уравнения регрессии. Если теоретические соображения не могут подсказать точного решения, то следует сделать расчеты по двум или более уравнениям. Предпочтение следует отдать той модели, для которой меньше величина остаточной дисперсии, или, что то же самое, меньше величина остаточной суммы квадратов $S = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$. Однако, при незначительных расхождениях в остаточных дисперсиях, следует остановиться на более простом уравнении.

Рассмотрим более подробно линейное уравнение регрессии. Коэффициент a_1 показывает величину изменения результативного признака при изменении факторного признака на единицу. Следовательно, коэффициент a_1 зависит от единиц измерения переменных. Поэтому в качестве универсального показателя тесноты связи между величинами x и y используется *выборочный линейный коэффициент корреляции*

$$r = a_1 \frac{s_x}{s_y}. \quad (4.6)$$

Здесь s_x и s_y – средние квадратические отклонения соответствующих признаков (факторного и результативного). Они вычисляются по формулам

$$s_x^2 = \frac{\sum (x_i - \bar{x}_в)^2}{n-1}; \quad s_y^2 = \frac{\sum (y_i - \bar{y}_в)^2}{n-1}, \quad (4.7)$$

где $\bar{x}_в, \bar{y}_в$ – выборочные средние значения.

Линейный коэффициент корреляции изменяется в пределах от -1 до 1 , т.е. $-1 \leq r \leq 1$. Если $r > 0$, то связь между переменными x и y прямая, если $r < 0$, то связь между переменными x и y обратная. В зависимости от того, насколько $|r|$ приближается к 1 , различают связь слабую, умеренную и сильную. При $|r| < 0,2$ связь между переменными практически отсутствует. При $|r| = 1$ связь между x и y функциональная, т.е. наблюдаемые значения располагаются точно на прямой.

Выборочный коэффициент корреляции r является оценкой генерального коэффициента корреляции ρ , который определяется по формуле

$$\rho = \frac{M(xy) - M(x)M(y)}{\sigma_x \sigma_y}.$$

Величина ρ характеризует тесноту связи между переменными x и y в генеральной совокупности. Однако в практических исследованиях о тесноте корреляционной зависимости между переменными судят фактически не по величине генерального коэффициента корреляции ρ , а по величине его выборочного аналога r .

Пусть вычисленное значение $r \neq 0$. Возникает вопрос, объясняется ли это действительно существующей линейной связью между переменными x и y , или является следствием случайности отбора переменных в выборку. Обычно в этих случаях проверяется гипотеза H_0 об отсутствии линейной корреляционной связи между переменными, т.е. $H_0: \rho = 0$ при альтернативной гипотезе $H_0: \rho \neq 0$. Для проверки этой гипотезы на уровне значимости α вычисляют наблюдаемое значение критерия

$$T_{\text{набл}} = \frac{|r|}{\sqrt{1-r^2}} \sqrt{n-2}. \quad (4.8)$$

Критическое значение критерия $T(1 - \alpha, n - 2)$ находят по таблице критических точек распределения Стьюдента для числа степеней свободы $n - 2$ и уровня значимости α . Если $T_{\text{набл}} < T(\alpha, n - 2)$, то гипотеза H_0 принимается, в противном случае гипотеза H_0 отвергается, т.е. коэффициент корреляции признается существенно отличающимся от нуля.

5 ЗАДАНИЯ ДЛЯ ВЫПОЛНЕНИЯ КУРСОВОЙ РАБОТЫ

(Уровень значимости α , доверительная вероятность γ , номера варианта N1 и N2 задаются преподавателем).

Задача 1 Размер обработанных на некотором станке деталей может быть рассмотрен как случайная величина X , распределенная по нормальному закону. Для контроля качества деталей было произведено 50 измерений. Результаты измерений приведены в табл. 5.1.1.

1.1 Провести группировку данных, разбив варианты на 8 интервалов.

1.2 Для сгруппированного ряда построить гистограмму частот.

2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднее квадратическое отклонение случайной величины X .

2.2 Построить доверительный интервал для генеральной средней и генерального среднее квадратического отклонения с заданным уровнем доверительной вероятности γ .

3.1 Проектный размер детали должен быть равен a (табл. 5.1.2). При уровне значимости α проверить утверждение производителя о совпадении размера произведенных деталей с проектным размером.

3.2. При уровне значимости α проверить, является ли статистически обоснованным утверждение производителя о равенстве среднее квадратического отклонения размера детали заданному значению σ_0 (табл. 5.1.3).

4.1 Была исследована зависимость случайной величины Y (срок службы произведенных деталей) от величины X (размер детали). В результате проведения 10 измерений были получены следующие результаты (табл. 5.1.4). По этим данным построить диаграмму рассеяния.

4.2 Построить линейное уравнение регрессии.

4.3 Построить параболическое уравнение регрессии.

4.4 Для линейной модели проверить адекватность по F -критерию.

4.5 По модели с наименьшей остаточной дисперсией вычислить прогнозируемое значение y^* при заданном значении x^* (табл. 5.1.5).

4.6 Вычислить выборочный линейный коэффициент корреляции.

4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.1.1

N1	1	2	3	4	5	6
i	x_i					
1	1,41 +	1,94 +	2,21 +	1,78 +	4,20 +	-1,38 +
2	N2	N2	N2	N2	N2	N2
3	2,74 +	2,05 +	2,27 +	2,71 +	4,31 +	-0,73 +
4	N2	N2	N2	N2	N2	N2
5	3,77 +	2,09 +	3,18 +	2,88 +	4,39 +	-0,36 +
	N2	N2	N2	N2	N2	N2
	4,24 +	2,50 +	3,35 +	3,42 +	4,58 +	0,07 +
	N2	N2	N2	N2	N2	N2
	4,66 +	2,54 +	3,38 +	3,61 +	4,67 +	0,20 +
	N2	N2	N2	N2	N2	N2

Продолжение табл. 5.1.1

N1	1	2	3	4	5	6
i	x_i					
6	4,91 +	2,71 +	3,55 +	3,65 +	4,69 +	0,46 +
7	N2	N2	N2	N2	N2	N2
8	4,99 +	2,78 +	3,87 +	3,67 +	4,79 +	0,79 +
9	N2	N2	N2	N2	N2	N2
10	5,24 +	2,83 +	4,09 +	3,87 +	4,87 +	1,02 +
11	N2	N2	N2	N2	N2	N2
12	5,47 +	2,98 +	4,27 +	3,94 +	5,17 +	1,14 +
13	N2	N2	N2	N2	N2	N2
14	6,33 +	3,07 +	4,43 +	3,98 +	5,24 +	1,30 +
15	N2	N2	N2	N2	N2	N2

16	6,42 +	3,07 +	4,55 +	4,07 +	5,28 +	1,43 +
17	N2	N2	N2	N2	N2	N2
18	6,42 +	3,18 +	4,62 +	4,16 +	5,44 +	1,68 +
19	N2	N2	N2	N2	N2	N2
20	6,55 +	3,23 +	4,73 +	4,17 +	5,45 +	1,74 +
	N2	N2	N2	N2	N2	N2
	6,64 +	3,26 +	5,07 +	4,23 +	5,57 +	2,15 +
	N2	N2	N2	N2	N2	N2
	6,64 +	3,40 +	5,31 +	4,43 +	5,74 +	2,49 +
	N2	N2	N2	N2	N2	N2
	6,66 +	3,49 +	5,32 +	4,50 +	5,83 +	2,58 +
	N2	N2	N2	N2	N2	N2
	6,86 +	3,49 +	5,35 +	4,65 +	5,88 +	2,74 +
	N2	N2	N2	N2	N2	N2
	6,91 +	3,51 +	5,47 +	4,68 +	5,92 +	3,05 +
	N2	N2	N2	N2	N2	N2
	6,98 +	3,51 +	5,53 +	4,78 +	5,95 +	3,13 +
	N2	N2	N2	N2	N2	N2
	7,48 +	3,62 +	5,64 +	4,81 +	6,00 +	3,13 +
	N2	N2	N2	N2	N2	N2
21	7,69 +	3,70 +	5,64 +	4,89 +	6,05 +	3,17 +
22	N2	N2	N2	N2	N2	N2
23	7,74 +	3,82 +	5,73 +	4,90 +	6,13 +	3,20 +
24	N2	N2	N2	N2	N2	N2
25	7,79 +	3,83 +	5,84 +	4,99 +	6,13 +	3,25 +
26	N2	N2	N2	N2	N2	N2
27	7,84 +	3,84 +	5,92 +	5,00 +	6,23 +	3,34 +
28	N2	N2	N2	N2	N2	N2
29	7,95 +	3,87 +	5,94 +	5,01 +	6,34 +	3,57 +
30	N2	N2	N2	N2	N2	N2
31	7,97 +	3,89 +	5,99 +	5,04 +	6,37 +	3,61 +
32	N2	N2	N2	N2	N2	N2
33	8,44 +	4,02 +	6,02 +	5,05 +	6,53 +	3,80 +
34	N2	N2	N2	N2	N2	N2
35	8,52 +	4,21 +	6,15 +	5,24 +	6,90 +	3,85 +
36	N2	N2	N2	N2	N2	N2
37	8,58 +	4,24 +	6,20 +	5,24 +	7,05 +	3,94 +
38	N2	N2	N2	N2	N2	N2
39	9,09 +	4,33 +	6,28 +	5,25 +	7,12 +	4,18 +
40	N2	N2	N2	N2	N2	N2
41	9,28 +	4,43 +	6,33 +	5,26 +	7,14 +	4,25 +
42	N2	N2	N2	N2	N2	N2
43	9,63 +	4,55 +	6,50 +	5,34 +	7,22 +	4,43 +
44	N2	N2	N2	N2	N2	N2
45	9,96 +	4,56 +	6,59 +	5,42 +	7,33 +	4,45 +
46	N2	N2	N2	N2	N2	N2
47	10,37 +	4,60 +	6,67 +	5,44 +	7,36 +	4,56 +
48	N2	N2	N2	N2	N2	N2
49	10,51 +	4,75 +	6,79 +	5,46 +	7,49 +	4,61 +
50	N2	N2	N2	N2	N2	N2
	10,80 +	4,80 +	6,82 +	5,54 +	7,73 +	4,90 +
	N2	N2	N2	N2	N2	N2
	10,86 +	4,82 +	7,24 +	5,67 +	7,78 +	4,92 +
	N2	N2	N2	N2	N2	N2
	11,02 +	4,84 +	7,32 +	5,83 +	8,11 +	5,04 +
	N2	N2	N2	N2	N2	N2
	11,21 +	4,84 +	7,69 +	5,97 +	8,21 +	5,34 +
	N2	N2	N2	N2	N2	N2

11,34 +	4,88 +	7,69 +	6,08 +	8,42 +	5,57 +
N2	N2	N2	N2	N2	N2
11,45 +	4,91 +	7,76 +	6,10 +	8,57 +	6,13 +
N2	N2	N2	N2	N2	N2
11,71 +	5,02 +	8,13 +	6,15 +	8,62 +	6,21 +
N2	N2	N2	N2	N2	N2
11,83 +	5,10 +	8,21 +	6,25 +	9,12 +	6,24 +
N2	N2	N2	N2	N2	N2
11,89 +	5,29 +	8,33 +	6,27 +	9,14 +	6,30 +
N2	N2	N2	N2	N2	N2
12,91 +	5,37 +	8,88 +	6,50 +	9,93 +	6,40 +
N2	N2	N2	N2	N2	N2
13,46 +	5,42 +	9,62 +	6,50 +	10,17 +	6,59 +
N2	N2	N2	N2	N2	N2
13,54 +	5,42 +	9,67 +	6,53 +	10,25 +	6,71 +
N2	N2	N2	N2	N2	N2
13,55 +	5,49 +	9,73 +	6,56 +	10,28 +	7,87 +
N2	N2	N2	N2	N2	N2
15,92 +	5,62 +	10,59 +	6,67 +	10,29 +	8,19 +
N2	N2	N2	N2	N2	N2
16,92 +	5,82 +	10,68 +	6,76 +	11,34 +	8,74 +
N2	N2	N2	N2	N2	N2

Таблица 5.1.2

N1	1	2	3	4	5	6
<i>a</i>	10 + N2	4 + N2	6 + N2	8 + N2	8 + N2	4 + N2

Таблица 5.1.3

N1	1	2	3	4	5	6
σ_0	3	1	2,5	2	1	2

Таблица 5.1.4

N1	1	2	3	4	5	6
<i>i</i>	x_i					
1	2,41 +	2,08 +	2,69 +	3,07 +	5,24 +	2,27 +
2	N2	N2	N2	N2	N2	N2
3	5,72 +	3,28 +	3,12 +	3,48 +	5,32 +	2,43 +
4	N2	N2	N2	N2	N2	N2
5	6,46 +	3,28 +	3,88 +	4,46 +	6,46 +	2,78 +
6	N2	N2	N2	N2	N2	N2
7	7,21 +	3,50 +	5,94 +	5,20 +	6,49 +	3,57 +
8	N2	N2	N2	N2	N2	N2
9	7,62 +	4,06 +	6,79 +	5,20 +	6,79 +	3,63 +
10	N2	N2	N2	N2	N2	N2
11	7,93 +	4,70 +	6,89 +	5,40 +	6,83 +	3,73 +
12	N2	N2	N2	N2	N2	N2
13	8,13 +	4,85 +	7,19 +	6,40 +	7,22 +	4,31 +
14	N2	N2	N2	N2	N2	N2
15	8,39 +	4,99 +	7,37 +	7,40 +	7,86 +	4,55 +
16	N2	N2	N2	N2	N2	N2

	10,45 + N2	5,09 + N2	7,39 + N2	7,53 + N2	11,12 + N2	5,88 + N2
	11,83 + N2	5,88 + N2	12,00 + N2	10,35 + N2	11,77 + N2	6,22 + N2
i	y_i					
1	69,07	51,63	28,35	63,99	37,76	73,23
2	94,81	59,62	33,78	65,59	40,84	80,45
3	97,66	58,77	40,44	66,49	44,86	77,35
4	99,93	59,31	50,14	71,77	43,61	82,86
5	99,50	59,31	48,72	67,27	43,05	88,74
6	99,94	59,60	48,04	68,67	42,46	83,85
7	99,75	58,18	47,48	67,74	45,45	82,69
8	99,84	58,01	46,03	58,16	45,03	80,94
9	93,98	58,43	44,70	58,68	14,10	75,10
1	85,01	52,80	11,41	10,99	2,85	75,73
0						

Таблица 5.1.5

N1	1	2	3	4	5	6
x^*	12 + N2	6 + N2	11 + N2	11 + N2	10 + N2	8 + N2

Задача 2 Суточное потребление электроэнергии на рабочем месте в цехе может быть рассмотрено как случайная величина X , распределенная по нормальному закону. Было произведено 40 измерений этой величины (табл. 5.2.1).

- 1.1 Провести группировку данных, разбив варианты на 7 интервалов.
- 1.2 Для сгруппированного ряда построить гистограмму частот.
- 2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднее квадратическое отклонение случайной величины X .
- 2.2 Построить доверительный интервал для генеральной средней и генерального среднее квадратического отклонения с заданным уровнем доверительной вероятности γ .
- 3.1 При уровне значимости α проверить утверждение энергослужб предприятия, что среднесуточное потребление электроэнергии на рабочем месте равно a (табл. 5.2.2).
- 3.2 После переналадки оборудования, произведенной с целью уменьшения потребления электроэнергии, были проведены новые измерения, и получена выборка объема 10 (табл. 5.2.3). При уровне значимости α проверить, является ли статистически обоснованным утверждение производителя об уменьшении среднего потребления электроэнергии.
- 4.1 Была исследована зависимость случайной величины Y (производительность труда рабочего) от величины X (среднесуточное потребление электроэнергии). В результате проведения 8 измерений были получены следующие результаты (табл. 5.2.4). По этим данным построить диаграмму рассеяния.
- 4.2 Построить линейное уравнение регрессии.
- 4.3 Построить показательное уравнение регрессии.
- 4.4 Для построенных моделей проверить адекватность по F -критерию.
- 4.5 По адекватной модели вычислить прогнозируемое значение y^* при заданном значении x^* (табл. 5.2.5).
- 4.6 Вычислить выборочный линейный коэффициент корреляции.
- 4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.2.1

N1	1	2	3	4	5	6
i	x_i					
1	2,06 + N2	3,16 + N2	5,63 + N2	0,98 + N2	1,44 + N2	1,10 + N2
2	3,45 + N2	4,07 + N2	5,98 + N2	1,12 + N2	1,52 + N2	1,94 + N2
3	3,72 + N2	4,28 + N2	6,32 + N2	1,29 + N2	1,58 + N2	2,50 + N2

4	3,77 + N2	4,35 + N2	6,89 + N2	1,37 + N2	2,23 + N2	2,61 + N2
5	3,88 + N2	4,45 + N2	7,05 + N2	1,43 + N2	2,56 + N2	3,86 + N2

Продолжение табл. 5.2.1

N1	1	2	3	4	5	6
<i>i</i>	<i>x_i</i>					
6	4,11 + N2	4,47 + N2	7,21 + N2	1,46 + N2	2,60 + N2	4,08 + N2
7	4,15 + N2	4,48 + N2	7,36 + N2	1,46 + N2	2,94 + N2	4,25 + N2
8	4,24 + N2	4,92 + N2	7,44 + N2	1,48 + N2	3,14 + N2	4,64 + N2
9	4,28 + N2	5,46 + N2	7,49 + N2	1,57 + N2	3,21 + N2	5,04 + N2
1	4,49 + N2	5,47 + N2	7,51 + N2	1,61 + N2	3,55 + N2	5,11 + N2
0	4,52 + N2	5,48 + N2	7,76 + N2	1,76 + N2	3,65 + N2	5,42 + N2
1	4,64 + N2	5,54 + N2	8,40 + N2	1,77 + N2	3,67 + N2	6,24 + N2
1	4,65 + N2	5,81 + N2	8,53 + N2	1,84 + N2	3,86 + N2	6,50 + N2
1	4,75 + N2	5,85 + N2	9,02 + N2	1,85 + N2	4,27 + N2	6,71 + N2
2	4,78 + N2	6,30 + N2	9,09 + N2	1,95 + N2	4,41 + N2	7,16 + N2
1	4,88 + N2	6,46 + N2	9,24 + N2	1,99 + N2	4,41 + N2	7,23 + N2
3	5,08 + N2	6,52 + N2	9,66 + N2	2,02 + N2	4,41 + N2	7,92 + N2
1	5,09 + N2	6,59 + N2	9,70 + N2	2,03 + N2	4,44 + N2	7,98 + N2
4	5,18 + N2	6,62 + N2	9,76 + N2	2,07 + N2	4,67 + N2	8,31 + N2
1	5,22 + N2	6,71 + N2	9,77 + N2	2,08 + N2	4,73 + N2	8,41 + N2
5	5,22 + N2	6,87 + N2	10,17 +	2,10 + N2	4,73 + N2	9,01 + N2
1	5,27 + N2	6,96 + N2	N2	2,19 + N2	4,95 + N2	9,63 + N2
6	5,29 + N2	7,13 + N2	10,19 +	2,20 + N2	5,13 + N2	10,00 +
1	5,44 + N2	7,32 + N2	N2	2,22 + N2	5,16 + N2	N2
7	5,44 + N2	7,32 + N2	10,39 +	2,22 + N2	5,16 + N2	N2
1	5,63 + N2	7,36 + N2	N2	2,27 + N2	5,24 + N2	10,01 +
8	5,65 + N2	7,39 + N2	10,39 +	2,27 + N2	5,24 + N2	10,01 +
1	5,68 + N2	7,41 + N2	N2	2,28 + N2	5,25 + N2	N2
9	5,68 + N2	7,41 + N2	10,48 +	2,30 + N2	5,27 + N2	10,17 +
9	5,75 + N2	7,55 + N2	N2	2,31 + N2	5,31 + N2	N2
2	5,8 + N2	7,56 + N2	10,84 +	2,33 + N2	5,76 + N2	10,37 +
0	5,88 + N2	7,86 + N2	N2	2,42 + N2	5,85 + N2	N2
2	5,91 + N2	8,30 + N2	11,33 +	2,44 + N2	6,28 + N2	10,38 +
1	5,91 + N2	8,30 + N2	N2	2,44 + N2	6,28 + N2	N2
2	5,91 + N2	8,65 + N2	11,57 +	2,52 + N2	6,59 + N2	10,46 +
2	6,24 + N2	8,66 + N2	N2	2,61 + N2	6,76 + N2	N2
2	6,25 + N2	9,02 + N2	11,71 +	2,63 + N2	6,80 + N2	10,59 +
2	6,27 + N2	9,07 + N2	N2	2,70 + N2	7,21 + N2	N2
3	6,32 + N2	9,20 + N2	12,00 +	2,76 + N2	7,54 + N2	10,62 +
2	6,4 + N2	9,45 + N2	N2	2,81 + N2	7,56 + N2	N2
4	6,51 + N2	10,26 +	12,04 +	2,91 + N2	8,57 + N2	10,77 +
2	6,51 + N2	N2	N2	2,91 + N2	8,57 + N2	10,77 +
5	6,78 + N2	N2	N2	2,98 + N2	9,11 + N2	N2
2	7,05 + N2	10,50 +	12,28 +	3,09 + N2	10,71 +	11,23 +
6		N2	N2		N2	N2
2		14,21 +	12,49 +			11,25 +
7		N2	N2			N2
2			12,57 +			11,35 +
8			N2			N2
2			12,71 +			11,35 +
9			N2			N2
3			12,94 +			11,76 +
0			N2			N2
3			13,60 +			11,87 +
1			N2			N2

3			14,21 +			12,10 +
2			N2			N2
3			14,35 +			12,57 +
3			N2			N2
3			15,64 +			13,02 +
4			N2			N2
3			16,46 +			
5			N2			
3						
6						
3						
7						
3						
8						
3						
9						
4						
0						

Таблица 5.2.2

N1	1	2	3	4	5	6
<i>a</i>	5 + N2	7 + N2	8 + N2	2 + N2	4 + N2	9 + N2

Таблица 5.2.3

N1	1	2	3	4	5	6
<i>i</i>	x_i					
1	2,47 +	4,44 +	1,25 +	1,07 +	2,49 +	0,86 + N2
2	N2	N2	N2	N2	N2	3,05 + N2
3	3,85 +	4,94 +	3,52 +	1,08 +	3,28 +	3,51 + N2
4	N2	N2	N2	N2	N2	4,80 + N2
5	4,34 +	5,97 +	3,73 +	1,72 +	4,49 +	4,83 + N2
6	N2	N2	N2	N2	N2	5,45 + N2
7	4,35 +	6,28 +	5,32 +	1,83 +	4,73 +	7,01 + N2
8	N2	N2	N2	N2	N2	8,07 + N2
9	4,47 +	6,31 +	5,66 +	1,83 +	5,18 +	8,74 + N2
10	N2	N2	N2	N2	N2	14,24 + N2
	4,59 +	6,94 +	6,38 +	1,86 +	6,30 +	
	N2	N2	N2	N2	N2	
	4,79 +	7,40 +	9,43 +	1,88 +	6,62 +	
	N2	N2	N2	N2	N2	
	5,16 +	7,40 +	11,4 +	2,52 +	6,83 +	
	N2	N2	N2	N2	N2	
	5,42 +	8,50 +	11,8 +	2,53 +	6,90 +	
	N2	N2	N2	N2	N2	
	5,88 +	9,17 +	13,8 +	2,69 +	7,89 +	
	N2	N2	N2	N2	N2	

Таблица 5.2.4

N1	1	2	3	4	5	6
<i>i</i>	x_i					

1	3,53 + N2	3,37 + N2	2,39 + N2	0,79 + N2	0,80 + N2	7,72 + N2
2	4,29 + N2	6,24 + N2	6,23 + N2	1,13 + N2	3,74 + N2	9,75 + N2
3	4,60 + N2	6,30 + N2	7,83 + N2	1,41 + N2	5,01 + N2	11,04 +
4	4,70 + N2	6,35 + N2	8,46 + N2	1,73 + N2	5,46 + N2	N2
5	5,79 + N2	7,97 + N2	11,59 + N2	1,77 + N2	5,78 + N2	11,57 +
6	5,86 + N2	8,45 + N2	12,71 + N2	1,98 + N2	6,87 + N2	N2
7	6,74 + N2	9,80 + N2	13,34 + N2	2,05 + N2	7,09 + N2	11,70 +
8	6,75 + N2	11,20 + N2	13,42 + N2	2,44 + N2	9,06 + N2	N2
						12,42 +
						N2
						12,46 +
						N2
						14,08 +
						N2
<i>i</i>	<i>y_i</i>					
1	30,89	12,68	12,02	4,67	11,89	26,69 +
2	41,85	20,63	18,19	5,01	12,33	N2
3	46,17	15,73	22,70	8,21	15,31	30,41 +
4	45,76	20,87	25,23	9,49	17,38	N2
5	70,81	20,87	27,41	6,98	20,05	34,57 +
6	75,85	23,70	35,25	8,31	19,20	N2
7	94,52	25,83	32,69	9,34	22,16	31,18 +
8	92,79	24,35	42,47	11,39	23,24	N2
						33,40 +
						N2
						33,59 +
						N2
						35,92 +
						N2
						35,75 +
						N2

Таблица 5.2.5

N1	1	2	3	4	5	6
<i>x</i> *	7 + N2	10 + N2	14 + N2	0,5 + N2	10 + N2	13 + N2

Задача 3 При контроле технологического процесса было произведено 50 измерений температуры продукта на выходе из реактора (табл. 5.3.1). Пусть температура продукта на выходе из реактора может быть рассмотрена как случайная величина X , распределенная по нормальному закону.

1.1 Провести группировку данных, разбив варианты на 7 интервалов.

1.2 Для сгруппированного ряда построить гистограмму частот.

2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднее квадратическое отклонение случайной величины X .

2.2 Построить доверительный интервал для генеральной средней и генерального среднее квадратического отклонения с заданным уровнем доверительной вероятности γ .

3.1 При уровне значимости α проверить утверждение о равенстве средней температуры продукта на выходе из реактора номинальному значению a (табл. 5.3.2).

3.2 По техническим условиям среднее квадратическое отклонение случайной величины X не должно превышать заданного числа σ_0 (табл. 5.3.3). При уровне значимости α проверить, выполняется ли это требование.

4.1 Была исследована зависимость случайной величины Y (показатель качества произведенного продукта) от величины X (температуры продукта на выходе из реактора). В результате проведения 15 измерений были получены следующие результаты (табл. 5.3.4). По этим данным построить диаграмму

рассеяния.

4.2 Построить линейное уравнение регрессии

4.3 Построить гиперболическое уравнение регрессии.

4.4 Для построенных моделей проверить адекватность по F -критерию.

4.5 По адекватной модели вычислить прогнозируемое значение y^* при заданном значении x^* (табл.

5.3.5).

4.6 Вычислить выборочный линейный коэффициент корреляции.

4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.3.1

N1	1	2	3	4	5	6
i	x_i					
1	1,32 + N2	1,97 + N2	2,47 + N2	1,09 + N2	0,72 + N2	2,23 + N2
2	2,63 + N2	2,12 + N2	2,47 + N2	1,39 + N2	0,73 + N2	2,27 + N2
3	2,76 + N2	2,34 + N2	4,70 + N2	1,48 + N2	0,94 + N2	2,44 + N2
4	2,83 + N2	2,47 + N2	5,16 + N2	1,63 + N2	1,78 + N2	2,44 + N2
5	3,09 + N2	2,51 + N2	5,48 + N2	1,77 + N2	2,06 + N2	2,64 + N2
6	3,39 + N2	3,09 + N2	5,70 + N2	1,83 + N2	2,69 + N2	2,66 + N2
7	3,57 + N2	3,09 + N2	5,79 + N2	1,88 + N2	2,91 + N2	2,72 + N2
8	3,88 + N2	3,90 + N2	6,13 + N2	1,89 + N2	2,95 + N2	3,16 + N2

Продолжение табл. 5.3.1

N1	1	2	3	4	5	6
i	x_i					
9	4,18 + N2	4,06 + N2	6,31 + N2	1,96 + N2	3,29 + N2	3,19 + N2
10	4,35 + N2	4,63 + N2	6,40 + N2	2,06 + N2	3,57 + N2	3,25 + N2
11	4,49 + N2	4,79 + N2	6,47 + N2	2,28 + N2	3,75 + N2	3,26 + N2
12	4,71 + N2	5,02 + N2	6,54 + N2	2,30 + N2	3,89 + N2	3,35 + N2
13	4,97 + N2	5,46 + N2	6,63 + N2	2,37 + N2	3,92 + N2	3,37 + N2
14	5,00 + N2	5,47 + N2	6,63 + N2	2,39 + N2	3,95 + N2	3,40 + N2
15	5,26 + N2	5,49 + N2	6,75 + N2	2,40 + N2	4,06 + N2	3,40 + N2
16	5,30 + N2	5,65 + N2	7,09 + N2	2,41 + N2	4,17 + N2	3,40 + N2
17	5,36 + N2	5,75 + N2	7,16 + N2	2,42 + N2	4,17 + N2	3,51 + N2
18	5,41 + N2	5,80 + N2	7,23 + N2	2,44 + N2	4,65 + N2	3,64 + N2
19	5,59 + N2	6,17 + N2	7,34 + N2	2,51 + N2	4,83 + N2	3,70 + N2
20	5,67 + N2	6,40 + N2	7,36 + N2	2,54 + N2	4,86 + N2	3,75 + N2
21	5,69 + N2	6,46 + N2	7,43 + N2	2,58 + N2	5,06 + N2	3,77 + N2
22	5,75 + N2	7,06 + N2	7,61 + N2	2,61 + N2	5,12 + N2	3,78 + N2
23	5,82 + N2	7,28 + N2	7,66 + N2	2,75 + N2	5,57 + N2	3,86 + N2
24	5,88 + N2	7,90 + N2	7,72 + N2	2,84 + N2	5,75 + N2	3,91 + N2
25	5,94 + N2	8,37 + N2	7,73 + N2	2,89 + N2	5,75 + N2	4,00 + N2
26	5,96 + N2	8,67 + N2	7,81 + N2	2,92 + N2	5,76 + N2	4,03 + N2
27	6,01 + N2	8,67 + N2	7,91 + N2	2,94 + N2	5,76 + N2	4,07 + N2
28	6,28 + N2	8,67 + N2	8,09 + N2	2,99 + N2	5,80 + N2	4,08 + N2
29	6,80 + N2	8,69 + N2	8,17 + N2	3,03 + N2	5,89 + N2	4,12 + N2
30	6,94 + N2	8,73 + N2	8,22 + N2	3,14 + N2	5,92 + N2	4,16 + N2
31	6,96 + N2	8,74 + N2	8,60 + N2	3,16 + N2	5,92 + N2	4,19 + N2
32	6,98 + N2	9,00 + N2	8,64 + N2	3,21 + N2	6,04 + N2	4,25 + N2
33	7,33 + N2	9,01 + N2	8,80 + N2	3,26 + N2	6,05 + N2	4,25 + N2
34	7,51 + N2	9,13 + N2	9,01 + N2	3,29 + N2	6,06 + N2	4,27 + N2
35	7,63 + N2	9,16 + N2	9,03 + N2	3,32 + N2	6,22 + N2	4,34 + N2
36	7,68 + N2	9,17 + N2	9,06 + N2	3,40 + N2	6,50 + N2	4,35 + N2
37	7,70 + N2	9,46 + N2	9,10 + N2	3,48 + N2	6,72 + N2	4,49 + N2
38	8,06 + N2	9,90 + N2	9,14 + N2	3,52 + N2	7,02 + N2	4,50 + N2
39	8,22 + N2	9,93 + N2	9,34 + N2	3,56 + N2	7,15 + N2	4,51 + N2
40	8,29 + N2	9,94 + N2	9,38 + N2	3,62 + N2	7,17 + N2	4,54 + N2
41	8,33 + N2	10,02 +	9,39 + N2	3,63 + N2	7,20 + N2	4,76 + N2

42	8,33 + N2	N2	9,48 + N2	3,69 + N2	7,23 + N2	4,78 + N2
43	8,56 + N2	10,52 +	9,60 + N2	3,81 + N2	7,43 + N2	4,94 + N2
44	8,92 + N2	N2	9,64 + N2	3,84 + N2	7,56 + N2	5,00 + N2
45	8,94 + N2	10,77 +	10,03 +	3,92 + N2	7,67 + N2	5,14 + N2
46	9,01 + N2	N2	N2	4,13 + N2	8,20 + N2	5,33 + N2
47	9,60 + N2	10,80 +	10,10 +	4,16 + N2	8,26 + N2	5,58 + N2
48	10,06 +	N2	N2	4,27 + N2	8,27 + N2	5,61 + N2
49	N2	10,90 +	10,22 +	4,88 + N2	8,70 + N2	5,89 + N2
50	11,01 +	N2	N2	5,69 + N2	9,00 + N2	6,06 + N2
	N2	11,12 +	10,99 +			
	12,06 +	N2	N2			
	N2	11,45 +	11,26 +			
		N2	N2			
		11,91 +	12,34 +			
		N2	N2			
		12,58 +				
		N2				
		13,70 +				
		N2				

Таблица 5.3.2

N1	1	2	3	4	5	6
<i>a</i>	6 + N2	7 + N2	9 + N2	3 + N2	8 + N2	4 + N2

Таблица 5.3.3

N1	1	2	3	4	5	6
σ_0	2	2	2	0,5	2	1

Таблица 5.3.4

N1	1	2	3	4	5	6
<i>i</i>	x_i					
1	1,80 + N2	3,42 + N2	5,88 + N2	1,23 + N2	1,72 + N2	1,84 + N2
2	2,19 + N2	5,05 + N2	6,47 + N2	1,38 + N2	2,68 + N2	2,69 + N2
3	4,73 + N2	6,26 + N2	6,57 + N2	1,91 + N2	2,91 + N2	2,79 + N2
4	4,82 + N2	6,65 + N2	6,59 + N2	2,72 + N2	3,59 + N2	2,85 + N2
5	5,15 + N2	8,03 + N2	6,78 + N2	2,78 + N2	4,03 + N2	3,35 + N2
6	5,68 + N2	8,45 + N2	7,09 + N2	3,11 + N2	4,20 + N2	3,44 + N2
7	5,80 + N2	8,57 + N2	7,30 + N2	3,12 + N2	4,33 + N2	3,64 + N2
8	5,84 + N2	8,77 + N2	7,42 + N2	3,28 + N2	4,47 + N2	3,81 + N2
9	6,27 + N2	9,24 + N2	8,51 + N2	3,37 + N2	4,67 + N2	4,15 + N2
10	6,72 + N2	9,26 + N2	8,71 + N2	3,50 + N2	5,29 + N2	4,31 + N2
11	7,30 + N2	9,49 + N2	8,74 + N2	3,50 + N2	5,61 + N2	4,54 + N2
12	7,34 + N2	9,51 + N2	9,08 + N2	3,54 + N2	5,95 + N2	4,59 + N2
13	7,47 + N2	9,75 + N2	9,76 + N2	3,64 + N2	7,27 + N2	4,95 + N2
14	9,22 + N2	10,16 +	9,84 + N2	3,71 + N2	7,46 + N2	5,62 + N2
15	9,38 + N2	N2	10,19 +	4,71 + N2	7,54 + N2	5,97 + N2
		10,49 +	N2			
		N2				
<i>i</i>	y_i					
1	12,78	13,23	8,43	5,35	6,33	21,48
2	15,73	12,76	8,50	5,23	3,18	19,28
3	20,87	10,08	8,39	5,09	6,03	24,22
4	19,90	10,68	8,34	5,15	8,07	20,32
5	22,00	5,74	8,24	5,12	7,52	29,14
6	20,80	0,37	8,16	5,14	7,69	24,64

7	22,78	0,56	8,32	5,19	9,57	28,03
8	20,73	0,87	8,39	5,15	9,79	26,91
9	19,45	1,13	8,23	5,13	5,59	26,26
10	20,99	6,15	8,19	5,10	8,74	28,68
11	27,35	2,43	8,15	5,11	6,33	36,49
12	30,90	0,33	8,07	5,07	6,08	29,35
13	28,98	3,27	8,17	5,11	6,90	31,73
14	29,45	0,42	8,08	5,04	6,12	35,03
15	30,57	2,52	8,09	5,00	7,63	40,25

Таблица 5.3.5

N1	1	2	3	4	5	6
x^*	$8 + N2$	$11 + N2$	$10 + N2$	$5 + N2$	$7 + N2$	$3 + N2$

Задача 4 Предприятие получает сырье, характеризующееся процентным содержанием в нем полезного вещества A , а также стабильностью этого показателя. Для поступившей партии сырья были произведены анализы концентрации вещества A по 40 опытам (таблица 5.4.1). Пусть концентрация вещества A может быть рассмотрена как случайная величина X , распределенная по нормальному закону.

1.1 Провести группировку данных, разбив варианты на 7 интервалов.

1.2 Для сгруппированного ряда построить гистограмму частот.

2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднее квадратическое отклонение случайной величины X .

2.2 Построить доверительный интервал для генеральной средней и генерального среднее квадратического отклонения с заданным уровнем доверительной вероятности γ .

3.1 При уровне значимости α проверить утверждение поставщика о том, что его сырье имеет среднюю концентрацию a % (табл. 5.4.2).

3.2 По техническим условиям среднее квадратическое отклонение случайной величины X не должно превышать заданного числа σ_0 (табл. 5.4.3). При уровне значимости α проверить, выполняется ли это требование.

4.1 Была исследована зависимость случайной величины Y (показатель качества произведенного продукта) от величины X (концентрации вещества A в сырье). В результате проведения 12 измерений были получены следующие результаты (табл. 5.4.4). По этим данным построить диаграмму рассеяния.

4.2 Построить линейное уравнение регрессии.

4.3 Построить степенное уравнение регрессии.

4.4 Для построенных моделей проверить адекватность по F -критерию.

4.5 По адекватной модели вычислить прогнозируемое значение y^* при заданном значении x^* (табл. 5.4.5).

4.6 Вычислить выборочный линейный коэффициент корреляции.

4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.4.1

N1	1	2	3	4	5	6
i	x_i					
1	$2,17 + N2$	$0,74 + N2$	$1,83 + N2$	$1,65 + N2$	$1,84 + N2$	$1,16 + N2$
2	$2,27 + N2$	$3,68 + N2$	$2,74 + N2$	$2,26 + N2$	$1,92 + N2$	$1,73 + N2$
3	$2,60 + N2$	$4,72 + N2$	$3,27 + N2$	$2,49 + N2$	$1,98 + N2$	$3,27 + N2$
4	$2,76 + N2$	$5,32 + N2$	$3,79 + N2$	$2,50 + N2$	$2,09 + N2$	$3,50 + N2$

5	2,81 + N2	5,77 + N2	5,75 + N2	2,92 + N2	2,14 + N2	3,79 + N2
---	-----------	-----------	-----------	-----------	-----------	-----------

Продолжение табл. 5.4.1

N1	1	2	3	4	5	6
i	x_i					
6	3,00 + N2	6,27 + N2	5,78 + N2	2,93 + N2	2,15 + N2	3,82 + N2
7	3,05 + N2	6,51 + N2	5,90 + N2	3,01 + N2	2,43 + N2	4,39 + N2
8	3,06 + N2	6,51 + N2	5,96 + N2	3,05 + N2	2,45 + N2	4,45 + N2
9	3,22 + N2	6,54 + N2	6,01 + N2	3,10 + N2	2,51 + N2	4,73 + N2
10	3,28 + N2	6,97 + N2	6,12 + N2	3,19 + N2	2,52 + N2	5,19 + N2
11	3,37 + N2	7,50 + N2	6,19 + N2	3,40 + N2	2,65 + N2	5,40 + N2
12	3,60 + N2	7,85 + N2	6,27 + N2	3,41 + N2	2,65 + N2	5,59 + N2
13	3,63 + N2	7,94 + N2	6,35 + N2	3,55 + N2	2,71 + N2	5,65 + N2
14	3,68 + N2	8,20 + N2	6,88 + N2	3,63 + N2	2,71 + N2	5,72 + N2
15	3,69 + N2	8,21 + N2	7,04 + N2	3,75 + N2	2,74 + N2	5,98 + N2
16	3,90 + N2	8,42 + N2	7,12 + N2	3,82 + N2	2,74 + N2	6,60 + N2
17	4,27 + N2	8,52 + N2	7,84 + N2	3,82 + N2	2,79 + N2	6,63 + N2
18	4,39 + N2	8,79 + N2	7,96 + N2	3,82 + N2	2,92 + N2	7,18 + N2
19	4,67 + N2	9,10 + N2	8,14 + N2	3,86 + N2	2,95 + N2	7,22 + N2
20	4,69 + N2	9,14 + N2	8,26 + N2	3,94 + N2	2,95 + N2	7,24 + N2
21	4,84 + N2	9,26 + N2	8,42 + N2	4,16 + N2	3,00 + N2	7,29 + N2
22	5,02 + N2	9,39 + N2	8,82 + N2	4,22 + N2	3,00 + N2	7,35 + N2
23	5,02 + N2	9,41 + N2	8,84 + N2	4,25 + N2	3,01 + N2	7,41 + N2
24	5,02 + N2	9,46 + N2	8,86 + N2	4,27 + N2	3,03 + N2	7,49 + N2
25	5,17 + N2	9,51 + N2	9,16 + N2	4,31 + N2	3,06 + N2	7,54 + N2
26	5,29 + N2	9,67 + N2	9,58 + N2	4,31 + N2	3,06 + N2	7,62 + N2
27	5,35 + N2	9,84 + N2	9,71 + N2	4,37 + N2	3,07 + N2	7,65 + N2
28	5,61 + N2	9,95 + N2	9,74 + N2	4,38 + N2	3,08 + N2	7,76 + N2
29	5,62 + N2	10,30 + N2	9,75 + N2	4,39 + N2	3,08 + N2	7,95 + N2
30	5,62 + N2	N2	10,36 + N2	4,45 + N2	3,20 + N2	8,13 + N2
31	6,01 + N2	10,41 + N2	N2	4,68 + N2	3,31 + N2	8,35 + N2
32	6,64 + N2	N2	10,56 + N2	4,75 + N2	3,36 + N2	8,39 + N2
33	6,68 + N2	10,42 + N2	N2	4,76 + N2	3,41 + N2	8,67 + N2
34	7,07 + N2	N2	10,68 + N2	5,00 + N2	3,51 + N2	9,24 + N2
35	7,12 + N2	10,50 + N2	N2	5,22 + N2	3,57 + N2	9,27 + N2
36	7,35 + N2	N2	11,24 + N2	5,30 + N2	3,62 + N2	9,31 + N2
37	7,36 + N2	11,09 + N2	N2	5,42 + N2	3,75 + N2	9,44 + N2
38	7,40 + N2	N2	11,29 + N2	5,45 + N2	3,84 + N2	10,80 + N2
39	7,99 + N2	11,11 + N2	N2	5,63 + N2	3,91 + N2	N2
40	8,19 + N2	N2	11,34 + N2	6,35 + N2	3,91 + N2	10,87 + N2
		11,27 + N2	N2			N2
		N2	11,98 + N2			11,49 + N2
		12,17 + N2	N2			N2
		N2	13,71 + N2			
		12,46 + N2	N2			
		N2	14,10 + N2			
		13,25 + N2	N2			
		N2	14,53 + N2			
		13,67 + N2	N2			
		N2	15,16 + N2			
		15,95 + N2	N2			
		N2				

Таблица 5.4.2

N1	1	2	3	4	5	6
a	$4 + N2$	$9 + N2$	$8 + N2$	$5 + N2$	$2 + N2$	$2,5 + N2$

Таблица 5.4.3

N1	1	2	3	4	5	6
σ_0	2	2	3	1	2,5	2

Таблица 5.4.4

N1	1	2	3	4	5	6
i	x_i					
1	$1,76 + N2$	$5,14 + N2$	$4,53 + N2$	$2,49 + N2$	$1,65 + N2$	$2,35 + N2$
2	$2,39 + N2$	$7,66 + N2$	$5,12 + N2$	$2,98 + N2$	$1,90 + N2$	$3,28 + N2$
3	$2,68 + N2$	$8,07 + N2$	$5,37 + N2$	$3,09 + N2$	$2,15 + N2$	$3,83 + N2$
4	$3,36 + N2$	$8,75 + N2$	$5,46 + N2$	$3,44 + N2$	$2,34 + N2$	$6,13 + N2$
5	$3,61 + N2$	$9,36 + N2$	$6,39 + N2$	$3,49 + N2$	$2,42 + N2$	$6,54 + N2$
6	$4,58 + N2$	$9,48 + N2$	$8,58 + N2$	$3,58 + N2$	$2,55 + N2$	$6,64 + N2$
7	$4,83 + N2$	$10,70 + N2$	$9,83 + N2$	$3,60 + N2$	$2,67 + N2$	$6,82 + N2$
8	$5,68 + N2$	$10,80 + N2$	$10,36 + N2$	$3,61 + N2$	$2,74 + N2$	$6,86 + N2$
9	$5,77 + N2$	$10,80 + N2$	$10,48 + N2$	$4,09 + N2$	$3,00 + N2$	$7,22 + N2$
10	$6,68 + N2$	$10,94 + N2$	$10,48 + N2$	$4,29 + N2$	$3,06 + N2$	$7,74 + N2$
11	$7,89 + N2$	$10,94 + N2$	$10,96 + N2$	$5,08 + N2$	$3,18 + N2$	$9,32 + N2$
12	$8,58 + N2$	$11,21 + N2$	$11,01 + N2$	$5,26 + N2$	$3,23 + N2$	$10,80 + N2$
		$11,25 + N2$	$11,23 + N2$			
		$12,10 + N2$				
i	y_i					
1	1,20	4,99	$1,52 + N2$	16,73	15,70	$19,54 + N2$
2	2,70	11,76	$1,33 + N2$	17,10	14,19	$20,68 + N2$
3	2,30	11,87	$2,66 + N2$	16,53	14,71	$22,78 + N2$
4	5,33	16,05	$2,43 + N2$	16,86	16,36	$26,83 + N2$
5	6,72	16,33	$3,15 + N2$	18,22	13,83	$27,85 + N2$
6	9,30	16,16	$4,68 + N2$	17,59	15,64	$29,20 + N2$
7	12,41	24,95	$5,55 + N2$	18,49	15,97	$29,49 + N2$
8	16,46	23,45	$5,83 + N2$	18,40	15,34	$31,00 + N2$
9	16,00	24,67	$5,98 + N2$	18,70	16,86	$31,74 + N2$
10	23,33	25,39	$4,98 + N2$	18,21	16,17	$35,10 + N2$
11	32,25	24,61	$7,02 + N2$	21,19	16,74	
12	37,78	29,11	$5,73 + N2$	21,69	17,79	

						N2 38,64 + N2
--	--	--	--	--	--	---------------------

Таблица 5.4.5

N1	1	2	3	4	5	6
x^*	10 + N2	6 + N2	5 + N2	6 + N2	2 + N2	8,5 + N2

Задача 5 Анализируется выброс химическим предприятием в атмосферу вредных примесей. Концентрация вредных примесей может быть рассмотрена как случайная величина X , распределенная по нормальному закону. Были произведены анализы концентрации вредных примесей на территории химического предприятия в течение 40 дней (табл. 5.5.1).

1.1 Провести группировку данных, разбив варианты на 7 интервалов.

1.2 Для сгруппированного ряда построить гистограмму частот.

2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднеквадратическое отклонение случайной величины X .

2.2 Построить доверительный интервал для генеральной средней и генерального среднеквадратического отклонения с заданным уровнем доверительной вероятности γ .

3.1 При уровне значимости α проверить утверждение, что его средняя концентрация вредных примесей не превышает величину $a\%$ (табл. 5.5.2).

3.2 После установки новых очистных сооружений были проведены новые 10 замеров концентрации вредных примесей (табл. 5.5.3). При уровне значимости α проверить, является ли статистически обоснованным утверждение об уменьшении выброса в атмосферу вредных примесей.

4.1 Была исследована зависимость случайной величины Y (количество больничных листов в расчете на одного работающего) от величины X (концентрация вредных примесей на территории предприятия). Были получены следующие результаты (табл. 5.5.4). По этим данным построить диаграмму рассеяния.

4.2 Построить линейное уравнение регрессии.

4.3 Построить показательное уравнение регрессии.

4.4 Проверить адекватность моделей по F -критерию.

4.5 По обеим моделям вычислить прогнозируемое значение y^* при заданном значении x^* (табл. 5.5.5).

4.6 Вычислить выборочный линейный коэффициент корреляции.

4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.5.1

N1	1	2	3	4	5	6
i	x_i					
1	2,00 + N2	0,07 + N2	1,17 + N2	0,49 + N2	2,54 + N2	1,05 + N2
2	2,88 + N2	0,30 + N2	1,50 + N2	1,06 + N2	2,84 + N2	1,33 + N2
3	2,90 + N2	0,38 + N2	1,51 + N2	2,58 + N2	3,17 + N2	1,47 + N2
4	3,14 + N2	1,47 + N2	1,61 + N2	3,26 + N2	4,36 + N2	1,48 + N2
5	3,18 + N2	1,94 + N2	1,69 + N2	4,17 + N2	4,41 + N2	1,49 + N2
6	3,19 + N2	1,99 + N2	1,89 + N2	4,25 + N2	4,96 + N2	1,49 + N2

Продолжение табл. 5.5.1

N1	1	2	3	4	5	6
i	x_i					
7	3,22 + N2	2,01 + N2	2,02 + N2	4,26 + N2	5,42 + N2	1,53 + N2
8	3,28 + N2	2,32 + N2	2,07 + N2	4,29 + N2	5,68 + N2	1,76 + N2
9	3,31 + N2	2,48 + N2	2,12 + N2	4,59 + N2	5,86 + N2	1,78 + N2
1	3,38 + N2	2,64 + N2	2,39 + N2	4,75 + N2	6,00 + N2	1,83 + N2
0	3,47 + N2	2,77 + N2	2,42 + N2	4,82 + N2	6,44 + N2	1,83 + N2
1	3,91 + N2	3,03 + N2	2,53 + N2	4,86 + N2	6,66 + N2	1,84 + N2

1	3,99 + N2	3,16 + N2	2,60 + N2	4,94 + N2	6,95 + N2	1,84 + N2
1	4,04 + N2	3,22 + N2	2,62 + N2	5,28 + N2	7,00 + N2	1,88 + N2
2	4,22 + N2	3,25 + N2	2,66 + N2	5,37 + N2	7,13 + N2	1,89 + N2
1	4,30 + N2	3,78 + N2	2,76 + N2	5,75 + N2	7,19 + N2	1,96 + N2
3	4,54 + N2	3,79 + N2	2,81 + N2	6,13 + N2	7,77 + N2	1,98 + N2
1	4,80 + N2	3,86 + N2	2,82 + N2	6,16 + N2	8,12 + N2	2,01 + N2
4	4,95 + N2	4,04 + N2	2,85 + N2	6,66 + N2	8,17 + N2	2,03 + N2
1	5,24 + N2	4,25 + N2	2,85 + N2	6,69 + N2	8,42 + N2	2,05 + N2
5	5,34 + N2	4,26 + N2	2,89 + N2	6,95 + N2	8,56 + N2	2,09 + N2
1	5,37 + N2	4,35 + N2	3,09 + N2	7,02 + N2	8,64 + N2	2,09 + N2
6	5,55 + N2	4,37 + N2	3,12 + N2	7,24 + N2	9,16 + N2	2,11 + N2
1	5,77 + N2	4,52 + N2	3,32 + N2	7,24 + N2	9,19 + N2	2,11 + N2
7	5,81 + N2	4,57 + N2	3,33 + N2	8,24 + N2	9,36 + N2	2,14 + N2
1	5,90 + N2	4,65 + N2	3,33 + N2	8,37 + N2	9,41 + N2	2,18 + N2
8	6,00 + N2	4,82 + N2	3,36 + N2	8,52 + N2	9,73 + N2	2,18 + N2
1	6,02 + N2	4,86 + N2	3,71 + N2	8,63 + N2	9,90 + N2	2,21 + N2
9	6,20 + N2	5,33 + N2	3,79 + N2	8,73 + N2	10,09 +	2,26 + N2
2	6,22 + N2	5,36 + N2	3,83 + N2	8,74 + N2	N2	2,26 + N2
0	6,25 + N2	5,48 + N2	3,91 + N2	8,81 + N2	10,60 +	2,27 + N2
2	6,42 + N2	5,52 + N2	4,13 + N2	8,84 + N2	N2	2,29 + N2
1	6,43 + N2	5,53 + N2	4,15 + N2	9,07 + N2	10,71 +	2,29 + N2
2	6,64 + N2	5,95 + N2	4,30 + N2	9,22 + N2	N2	2,38 + N2
2	6,65 + N2	6,96 + N2	4,39 + N2	9,61 + N2	10,94 +	2,38 + N2
2	7,08 + N2	6,98 + N2	4,44 + N2	10,16 +	N2	2,38 + N2
3	7,11 + N2	7,11 + N2	4,46 + N2	N2	11,79 +	2,41 + N2
2	7,57 + N2	7,28 + N2	4,54 + N2	10,66 +	N2	2,48 + N2
4	7,59 + N2	7,68 + N2	4,57 + N2	N2	11,84 +	2,48 + N2
2	8,48 + N2	8,88 + N2	4,63 + N2	11,82 +	N2	3,03 + N2
5				N2	11,89 +	
2				12,93 +	N2	
6				N2	12,14 +	
2				13,74 +	N2	
7				N2	13,03 +	
2					N2	
8					14,11 +	
2					N2	
9					14,40 +	
3					N2	
0					15,06 +	
3					N2	
1						
3						
2						
3						
3						
3						
4						
3						
5						
3						
6						
3						
7						
3						
8						
3						

9						
4						
0						

Таблица 5.5.2

N1	1	2	3	4	5	6
a	7 + N2	4 + N2	3 + N2	8 + N2	8 + N2	2 + N2

Таблица 5.5.3

N1	1	2	3	4	5	6
<i>i</i>	x_i					
1	0,43 + N2	0,79 + N2	0,51 + N2	0,04 + N2	2,65 + N2	0,54 + N2
2	1,86 + N2	2,05 + N2	0,74 + N2	0,30 + N2	4,91 + N2	1,61 + N2
3	2,94 + N2	2,15 + N2	0,99 + N2	3,47 + N2	7,39 + N2	1,83 + N2
4	3,01 + N2	3,34 + N2	1,19 + N2	3,66 + N2	7,72 + N2	2,00 + N2
5	4,13 + N2	4,19 + N2	1,74 + N2	3,76 + N2	8,91 + N2	2,09 + N2
6	4,72 + N2	5,10 + N2	1,74 + N2	5,62 + N2	8,98 + N2	2,15 + N2
7	6,00 + N2	5,42 + N2	2,06 + N2	5,66 + N2	9,69 + N2	2,20 + N2
8	7,47 + N2	6,04 + N2	2,30 + N2	6,40 + N2	10,06 +	2,24 + N2
9	7,52 + N2	6,57 + N2	2,41 + N2	7,44 + N2	N2	2,39 + N2
1	7,82 + N2	6,59 + N2	3,28 + N2	8,74 + N2	10,09 +	2,95 + N2
0					N2	
					11,84 +	
					N2	

Таблица 5.5.4

N1	1	2	3	4	5	6
<i>i</i>	x_i					
1	2,26 + N2	0,43 + N2	2,48 + N2	2,25 + N2	4,20 + N2	1,44 + N2
2	4,02 + N2	1,56 + N2	2,48 + N2	3,10 + N2	5,90 + N2	1,57 + N2
3	4,17 + N2	1,80 + N2	2,52 + N2	5,25 + N2	5,93 + N2	2,12 + N2
4	4,41 + N2	3,17 + N2	2,92 + N2	5,81 + N2	6,70 + N2	2,18 + N2
5	4,90 + N2	3,19 + N2	3,20 + N2	6,28 + N2	7,06 + N2	2,25 + N2
6	5,01 + N2	4,31 + N2	3,34 + N2	8,99 + N2	7,15 + N2	2,40 + N2
7	5,37 + N2	5,29 + N2	4,22 + N2	10,23 +	9,44 + N2	2,54 + N2
8	8,48 + N2	6,94 + N2	4,60 + N2	N2	10,46 +	2,81 + N2
				10,51 +	N2	
				N2		
<i>i</i>	y_i					
1	4,47	6,72	9,17	5,05	3,79	6,18
2	3,63	5,96	8,67	3,99	5,08	7,02
3	5,24	9,58	8,96	8,77	3,67	7,15
4	5,65	6,70	9,03	9,24	4,72	8,33
5	4,66	7,13	9,70	10,70	4,70	7,37
6	5,37	8,69	9,62	18,96	3,79	7,92
7	5,23	10,14	10,60	23,66	3,57	7,76
8	6,99	13,14	11,34	24,53	4,70	7,55

Таблица 5.5.5

N1	1	2	3	4	5	6
----	---	---	---	---	---	---

x^*	7 + N2	7 + N2	2 + N2	4 + N2	10 + N2	3 + N2
-------	--------	--------	--------	--------	------------	--------

Задача 6 Напряжение на входе в узел прибора может быть рассмотрено как случайная величина X , распределенная по нормальному закону. Было произведено 40 измерений входного напряжения. Результаты измерений приведены в табл. 5.6.1.

- 1.1 Провести группировку данных, разбив варианты на 8 интервалов.
- 1.2 Для сгруппированного ряда построить гистограмму частот.
- 2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднее квадратическое отклонение случайной величины X .
- 2.2 Построить доверительный интервал для генеральной средней и генерального среднее квадратического отклонения с заданным уровнем доверительной вероятности γ .
- 3.1 При уровне значимости α проверить утверждение о равенстве среднего напряжения на входе в узел прибора номинальному значению a (таблица 5.6.2).
- 3.2 При уровне значимости α проверить утверждение о том, что дисперсия напряжения на входе в узел электроприбора не превосходит заданного значения σ_0 (табл. 5.6.3).
- 4.1 Была исследована зависимость случайной величины Y (показатель устойчивости работы прибора) от величины X (напряжение на входе в узел прибора). В результате проведения 10 измерений были получены следующие результаты (табл. 5.6.4). По этим данным построить диаграмму рассеяния.
- 4.2 Построить линейное уравнение регрессии.
- 4.3 Построить логарифмическое уравнение регрессии.
- 4.4 Для линейной модели проверить адекватность по F -критерию.
- 4.5 По адекватной модели вычислить прогнозируемое значение y^* при заданном значении x^* (табл. 5.6.5).
- 4.6 Вычислить выборочный линейный коэффициент корреляции.
- 4.7 При уровне значимости α проверить значимость коэффициента корреляции.

Таблица 5.6.1

N1	1	2	3	4	5	6
i	x_i					
1	2,80 + N2	3,10 + N2	1,13 + N2	0,07 + N2	0,99 + N2	0,23 + N2
2	3,54 + N2	3,18 + N2	2,89 + N2	0,72 + N2	1,08 + N2	1,14 + N2
3	3,83 + N2	3,99 + N2	5,10 + N2	0,92 + N2	1,42 + N2	1,43 + N2
4	4,50 + N2	5,41 + N2	5,45 + N2	1,15 + N2	1,50 + N2	1,66 + N2
5	4,74 + N2	5,46 + N2	6,04 + N2	1,31 + N2	1,59 + N2	2,10 + N2
6	5,68 + N2	5,49 + N2	6,14 + N2	1,37 + N2	1,68 + N2	2,16 + N2

Продолжение табл. 5.6.1

N1	1	2	3	4	5	6
i	x_i					
7	5,75 + N2	5,51 + N2	6,38 + N2	1,59 + N2	1,71 + N2	2,24 + N2
8	5,82 + N2	5,91 + N2	6,63 + N2	1,64 + N2	1,75 + N2	2,29 + N2
9	6,30 + N2	5,96 + N2	6,63 + N2	1,74 + N2	1,79 + N2	2,39 + N2
1	6,35 + N2	6,13 + N2	6,76 + N2	1,87 + N2	1,79 + N2	2,41 + N2
0	6,59 + N2	6,18 + N2	8,06 + N2	1,99 + N2	1,79 + N2	2,48 + N2
1	6,63 + N2	6,18 + N2	8,17 + N2	2,14 + N2	1,79 + N2	2,65 + N2
1	6,71 + N2	6,47 + N2	8,28 + N2	2,24 + N2	1,81 + N2	2,66 + N2
1	6,72 + N2	6,92 + N2	8,35 + N2	2,30 + N2	1,84 + N2	2,70 + N2
2	6,77 + N2	6,93 + N2	8,36 + N2	2,54 + N2	1,88 + N2	2,72 + N2
1	6,80 + N2	7,06 + N2	8,89 + N2	2,60 + N2	1,91 + N2	2,76 + N2
3	6,81 + N2	7,10 + N2	9,30 + N2	2,84 + N2	1,98 + N2	2,77 + N2
1	6,84 + N2	7,40 + N2	9,38 + N2	2,88 + N2	1,99 + N2	2,80 + N2
4	6,94 + N2	7,48 + N2	9,48 + N2	3,20 + N2	2,02 + N2	2,81 + N2
1						

5	6,97 + N2	7,76 + N2	9,49 + N2	3,28 + N2	2,07 + N2	2,92 + N2
1	7,29 + N2	7,85 + N2	9,95 + N2	3,65 + N2	2,07 + N2	2,97 + N2
6	7,56 + N2	8,00 + N2	10,07 +	3,69 + N2	2,10 + N2	2,98 + N2
1	7,56 + N2	8,18 + N2	N2	3,71 + N2	2,11 + N2	3,24 + N2
7	7,75 + N2	8,79 + N2	10,13 +	3,78 + N2	2,12 + N2	3,27 + N2
1	7,81 + N2	9,06 + N2	N2	3,81 + N2	2,12 + N2	3,42 + N2
8	7,84 + N2	9,15 + N2	10,47 +	3,84 + N2	2,15 + N2	3,44 + N2
1	7,85 + N2	9,16 + N2	N2	4,00 + N2	2,17 + N2	3,46 + N2
9	7,88 + N2	9,19 + N2	10,54 +	4,02 + N2	2,19 + N2	3,49 + N2
2	8,18 + N2	9,68 + N2	N2	4,27 + N2	2,22 + N2	3,60 + N2
0	8,56 + N2	9,89 + N2	10,61 +	4,55 + N2	2,23 + N2	3,65 + N2
2	8,80 + N2	9,92 + N2	N2	4,62 + N2	2,27 + N2	3,86 + N2
1	8,88 + N2	9,97 + N2	10,97 +	4,76 + N2	2,28 + N2	3,92 + N2
2	9,09 + N2	10,32 +	N2	4,83 + N2	2,29 + N2	3,96 + N2
2	9,25 + N2	N2	11,03 +	4,90 + N2	2,29 + N2	4,06 + N2
2	9,47 + N2	10,45 +	N2	4,90 + N2	2,29 + N2	4,17 + N2
3	9,61 + N2	N2	11,33 +	5,09 + N2	2,33 + N2	4,20 + N2
2	9,82 + N2	10,65 +	N2	5,31 + N2	2,34 + N2	4,46 + N2
4	10,13 +	N2	11,47 +	5,42 + N2	2,39 + N2	4,57 + N2
2	N2	11,14 +	N2	5,83 + N2	2,79 + N2	4,94 + N2
5	10,65 +	N2	11,72 +	5,85 + N2	2,97 + N2	5,23 + N2
2	N2	11,32 +	N2			
6	12,02 +	N2	12,86 +			
2	N2	11,92 +	N2			
7		N2	13,38 +			
2		12,41 +	N2			
8		N2	13,57 +			
2		13,86 +	N2			
9		N2	13,68 +			
3			N2			
0			13,90 +			
3			N2			
1			14,65 +			
3			N2			
2			14,71 +			
3			N2			
3			15,59 +			
3			N2			
4			17,03 +			
3			N2			
5						
3						
6						
3						
7						
3						
8						
3						
9						
4						
0						

Таблица 5.6.2

N1	1	2	3	4	5	6
a	$7 + N2$	$8,5 + N2$	$11 + N2$	$3 + N2$	$2 + N2$	$3,5 + N2$

Таблица 5.6.3

N1	1	2	3	4	5	6
σ_0	2	2	3	1,5	0,2	1

Таблица 5.6.4

N1	1	2	3	4	5	6
i	x_i					
1	$5,60 + N2$	$4,50 + N2$	$4,15 + N2$	$0,18 + N2$	$1,87 + N2$	$1,83 + N2$
2	$5,75 + N2$	$6,02 + N2$	$6,82 + N2$	$0,21 + N2$	$1,91 + N2$	$2,08 + N2$
3	$5,91 + N2$	$6,93 + N2$	$7,70 + N2$	$0,79 + N2$	$1,91 + N2$	$2,16 + N2$
4	$6,04 + N2$	$7,66 + N2$	$8,28 + N2$	$1,11 + N2$	$1,98 + N2$	$2,53 + N2$
5	$6,13 + N2$	$8,18 + N2$	$9,46 + N2$	$1,93 + N2$	$2,06 + N2$	$2,71 + N2$
6	$6,71 + N2$	$8,25 + N2$	$11,27 + N2$	$2,95 + N2$	$2,12 + N2$	$2,87 + N2$
7	$7,50 + N2$	$9,48 + N2$	$N2$	$3,48 + N2$	$2,15 + N2$	$3,01 + N2$
8	$8,28 + N2$	$9,61 + N2$	$11,31 + N2$	$3,75 + N2$	$2,26 + N2$	$3,06 + N2$
9	$8,54 + N2$	$9,98 + N2$	$N2$	$4,02 + N2$	$2,34 + N2$	$3,44 + N2$
10	$9,11 + N2$	$13,04 + N2$	$11,97 + N2$	$5,85 + N2$	$2,41 + N2$	$4,72 + N2$
			$12,63 + N2$			
			$13,23 + N2$			
i	y_i					
1	5,98	3,37	5,87	7,73	6,60	3,72
2	5,73	3,58	4,38	6,68	6,84	3,54
3	5,88	2,60	4,53	6,91	6,01	3,89
4	6,03	2,33	3,62	7,74	4,62	3,69
5	6,06	2,64	3,11	8,36	5,31	3,95
6	5,94	2,06	2,08	8,60	4,77	3,82
7	6,97	2,29	2,10	8,83	5,05	3,31
8	7,12	1,04	1,34	7,86	5,01	3,56
9	6,99	1,07	2,12	8,36	4,41	3,20
10	8,15	0,69	1,36	8,92	4,97	2,98

Таблица 5.6.5

N1	1	2	3	4	5	6
x^*	$9 + N2$	$5 + N2$	$10 + N2$	$6 + N2$	$1 + N2$	$5 + N2$

Задача. Некоторый технологический процесс характеризуется выходным параметром, который может быть рассмотрен как случайная величина X , распределенная по нормальному закону. Было проведено 50 измерений этого параметра (табл. 6.1).

Таблица 6.1

i	x_I	i	x_I	i	x_I	i	x_I	i	x_I
1	19,7	11	23,21	21	24,33	31	25,5	41	26,9
	1						8		3
2	19,8	12	23,26	22	24,36	32	25,5	42	27,1
	3						9		8
3	20,6	13	23,43	23	24,49	33	25,6	43	27,5
	5						8		2
4	21,1	14	23,49	24	24,78	34	25,7	44	28,1
	4						5		3
5	21,1	15	23,59	25	24,78	35	26,4	45	28,2
	8						2		3
6	21,3	16	23,73	26	25,08	36	26,4	46	28,5
	3						3		1
7	22,2	17	23,8	27	25,09	37	26,4	47	29,4
	8						9		4
8	22,4	18	23,89	28	25,27	38	26,5	48	29,5
	5						2		6
9	22,4	19	24,16	29	25,52	39	26,6	49	29,6
	8						5		1
10	23,1	20	24,26	30	25,57	40	26,6	50	30,3
	5						8		9

1.1 Провести группировку данных, разбив варианты на 8 интервалов.

1.2 Для сгруппированного ряда построить гистограмму частот.

2.1 Найти выборочную среднюю, выборочную дисперсию, исправленную выборочную дисперсию, исправленное выборочное среднеквадратическое отклонение случайной величины X .

2.2 Построить доверительный интервал для генеральной средней и генерального среднеквадратического отклонения с заданным уровнем доверительной вероятности $\gamma = 0,95$.

3.1 При уровне значимости $\alpha = 0,05$ проверить утверждение, что среднее значение величины X соответствует проектному значению $a = 25$.

3.2 После переналадки оборудования, произведенной с целью повышения выходного параметра X , были проведены новые измерения, и получены следующие результаты (табл. 6.2). При уровне значимости $\alpha = 0,05$ проверить, является ли статистически обоснованным утверждение об увеличении среднего значения показателя X .

Таблица 6.2

I	1	2	3	4	5	6	7	8	9	10
x_i	24,86	25,3	25,97	26,65	27,99	28,85	28,9	29,01	29,88	31,42

4.1 Была исследована зависимость случайной величины Y (показатель качества выпускаемой продукции) от величины X (выходной параметр технологического процесса). Были получены следующие результаты (табл. 6.3). По этим данным построить диаграмму рассеяния.

Таблица 6.3

I	1	2	3	4	5	6	7	8	9	10	11	12
x_i	20,5	21,7	23,9	24,4	24,6	25,2	25,2	26,4	26,9	27,0	27,4	27,7
	8	4	5	2	4	2	5	9	7		6	9

y_i	4,88	5,66	7,41	8,95	9,17	9,75	10,0	12,7	14,2	13,0	15,0	15,8
							5	2	1	7	5	6

- 4.2 Построить линейное уравнение регрессии.
- 4.3 Построить показательное уравнение регрессии.
- 4.4 Для обеих моделей проверить адекватность по F -критерию на уровне значимости $\alpha = 0,05$.
- 4.5 По адекватной модели вычислить прогнозируемое значение y^* при заданном значении $x^* = 21$.
- 4.6 Вычислить выборочный линейный коэффициент корреляции.
- 4.7 Проверить значимость коэффициента корреляции.

Решение

1.1 Проведем группировку исходных данных, т.е. разобьем варианты на отдельные интервалы. Найдем разность между наибольшим и наименьшим значениями признака $x_{\max} - x_{\min} = 30,39 - 19,71 = 10,68$. Тогда при разбивке на 8 интервалов длина интервала составит $h = 10,68/8 = 1,335 \approx 1,4$ (округляем всегда с избытком). Для упрощения дальнейших расчетов выберем границы интервалов таким образом, чтобы середины интервалов являлись как можно более «круглыми» числами (с наименьшим числом знаков после запятой). Получим табл. 6.4.

Таблица 6.4

I	Интервал $x_i^{\text{нач}} - x_i^{\text{кон}}$	Середина интервала x_i	Частота n_i	Относительная частота $w_i = n_i/n$
1	19,70...21,09	20,4	3	0,06
2	21,10...22,49	21,8	6	0,12
3	22,50...23,89	23,2	9	0,18
4	23,90...25,29	24,6	10	0,2
5	25,30...26,69	26	12	0,24
6	26,70...28,09	27,4	3	0,06
7	28,10...29,49	28,8	4	0,08
8	29,50...30,49	30,2	3	0,06
Σ			50	1

1.2 Построим для сгруппированного ряда гистограмму частот.

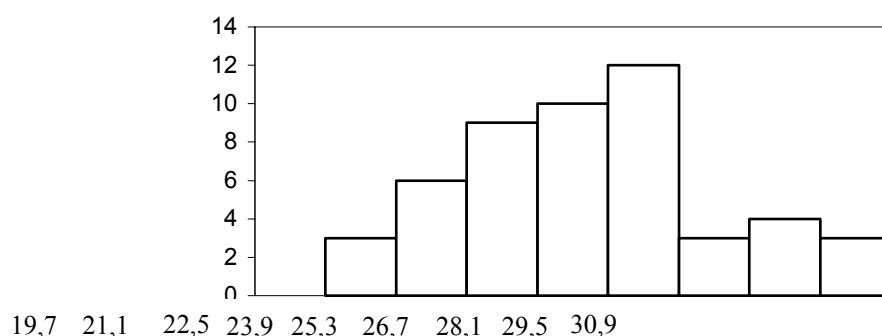


Рис. 6.1

2.1 Найдем выборочную среднюю \bar{x}_b по формуле (2.2) $\bar{x}_b = (n_1x_1 + n_2x_2 + \dots + n_kx_k) / n = (3 \cdot 20,4 + 6 \cdot 21,8 + 9 \cdot 23,2 + 10 \cdot 24,6 + 12 \cdot 26 + 3 \cdot 27,4 + 4 \cdot 28,8 + 3 \cdot 30,2) / 50 = 24,936 \approx 24,9$.

Найдем выборочную дисперсию D_b по формуле (2.4)

$$D_B = \frac{\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2}{n} =$$

$$= (3(20,4 - 24,9)^2 + 6(21,8 - 24,9)^2 + 9(23,2 - 24,9)^2 + 10(24,6 - 24,9)^2 +$$

$$+ 12(26 - 24,9)^2 + 3(27,4 - 24,9)^2 + 4(28,8 - 24,9)^2 + 3(30,2 - 24,9)^2) / 50 = 6,474.$$

Найдем исправленную выборочную дисперсию по формуле (2.5),

$$D_B^* = nD_B / (n - 1) = 50 \cdot 6,474 / 49 = 6,606.$$

Найдем исправленное выборочное среднеквадратическое отклонение случайной величины X

$$s = \sqrt{D_B^*} = \sqrt{6,606} \approx 2,57.$$

2.2 Построим доверительный интервал для генеральной средней с уровнем доверительной вероятности $\gamma = 0,95$. Так как значение генеральной дисперсии неизвестно, пользуемся формулой (2.10). Найдем значение $t_{1-\gamma, n-1} = t_{0,05;49}$ по таблице критических точек распределения Стьюдента при уровне вероятности $\alpha = 0,05$ и числе степеней свободы $k = n - 1 = 49$. Получаем $t_{0,05;49} = 2,01$. Далее находим точность оценки

$$\delta = \frac{t_{\gamma, n-1} s}{\sqrt{n}} = \frac{2,01 \cdot 2,57}{\sqrt{50}} \approx 0,7.$$

Согласно (2.8), доверительный интервал для генеральной средней имеет вид $\bar{x} \in (\bar{x}_B - \delta; \bar{x}_B + \delta)$. Подставляя значения, получаем, что с вероятностью 0,95 выполнено $\bar{x} \in (24,2; 25,6)$.

Построим доверительный интервал для генерального среднеквадратического отклонения с заданным уровнем доверительной вероятности $\gamma = 0,95$. Найдем значение χ_1^2 по таблице критических точек распределения χ^2 при уровне вероятности $(1 + \gamma) / 2 = 0,975$ и числе степеней свободы $k = n - 1 = 49$. Получаем $\chi_1^2 = 31,55$, следовательно, $\chi_1 = 5,62$. Найдем значение χ_2^2 по таблице критических точек распределения χ^2 при уровне вероятности $(1 - \gamma) / 2 = 0,025$ и числе степеней свободы $k = n - 1 = 49$. Получаем $\chi_2^2 = 70,22$, следовательно, $\chi_2 = 8,38$. Согласно (2.11), доверительный интервал для генерального

среднеквадратического отклонения имеет вид $\sigma \in \left(\frac{s\sqrt{n-1}}{\chi_2}; \frac{s\sqrt{n-1}}{\chi_1} \right)$. Подставляя значения, получаем, что с вероятностью 0,95 выполнено $\sigma \in \left(\frac{2,57\sqrt{49}}{8,38}; \frac{2,57\sqrt{49}}{5,62} \right)$ или $\sigma \in (2,15; 3,20)$.

3.1 При уровне значимости $\alpha = 0,05$ проверим утверждение, что среднее значение величины X соответствует проектному значению $a = 25$. Так как выборка имеет большой объем ($n = 50 > 30$), то для проверки нулевой гипотезы $H_0: \bar{x} = a$ в качестве критерия проверки можно принять случайную величину U , определенную по формуле (3.1). При этом в качестве генерального среднеквадратического отклонения σ можно принять выборочное значение s .

Вычислим наблюдаемое значение критерия

$$U_{\text{набл}} = \frac{(\bar{x}_B - a)\sqrt{n}}{s} = \frac{(24,9 - 25)\sqrt{50}}{2,57} = 0,28.$$

Конкурирующей является гипотеза $H_1: \bar{x} \neq a$, поэтому критическую точку $U_{\text{кр}}$ находим по таблице функции Лапласа из условия $\Phi(U_{\text{кр}}) = (1 - \alpha) / 2 = 0,475$. Получаем $U_{\text{кр}} = 1,96$. Так как $|U_{\text{набл}}| < U_{\text{кр}}$, то нет оснований отвергнуть нулевую гипотезу. Следовательно, утверждение, что среднее значение выходного параметра X соответствует проектному значению, является статистически обоснованным.

3.2 Для первой выборки объема $n_1 = 50$ были получены значения $\bar{x}_{B1} = 24,9$; $s_1 = 2,57$. Найдем выборочную среднюю и выборочное среднеквадратическое отклонение для второй выборки объема $n_2 = 10$ (табл. 6.2). Так как все варианты встречаются в выборке по одному разу, для нахождения выборочной средней пользуемся формулой (2.1). Получаем $\bar{x}_{B2} = (x_1 + x_2 + \dots + x_{n2}) / n_2 = (24,86 + 25,3 + 25,97 + 26,65 + 27,99 + 28,85 + 28,9 + 29,01 + 29,88 + 31,42) / 10 = 27,883 \approx 27,9$.

Для нахождения выборочной дисперсии пользуемся формулой (2.3). Получаем

$$D_{B2} = \frac{\sum_{i=1}^{n_2} (x_i - \bar{x}_{B2})^2}{n_2} =$$

$$= ((24,86 - 27,9)^2 + (25,3 - 27,9)^2 + (25,97 - 27,9)^2 + (26,65 - 27,9)^2 +$$

$$+ (27,99 - 27,9)^2 + (28,85 - 27,9)^2 + (28,9 - 27,9)^2 + (29,01 - 27,9)^2 +$$

$$+ (29,88 - 27,9)^2 + (31,42 - 27,9)^2) / 10 = 4,074.$$

Найдем исправленную выборочную дисперсию по формуле (2.5)

$$D_{B2}^* = n_2 D_{B2} / (n_2 - 1) = 10 \cdot 4,074 / 9 = 4,527.$$

Найдем исправленное выборочное среднеквадратическое отклонение случайной величины X

$$s_2 = \sqrt{D_{B2}^*} = \sqrt{4,527} \approx 2,13.$$

Проверяем теперь гипотезу о равенстве двух генеральных средних, т.е. гипотезу $H_0: \bar{x}_1 = \bar{x}_2$. Предполагаем, что дисперсии двух выборок равны. Поскольку мы проверяем утверждение, что среднее значение выходного параметра увеличилось, то в качестве конкурирующей выбираем гипотезу $H_1: \bar{x}_1 < \bar{x}_2$. Так как генеральные дисперсии неизвестны, а объем одной из выборок мал, используем критерий (3.9). Наблюдаемое значение критерия

$$T_{\text{набл}} =$$

$$= \frac{\bar{x}_{B1} - \bar{x}_{B2}}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{24,9 - 27,9}{\sqrt{\frac{49 \cdot 2,57^2 + 9 \cdot 2,13^2}{50 + 10 - 2} \left(\frac{1}{50} + \frac{1}{10} \right)}} = -3,46.$$

Критическую точку $T_{\text{кр}}(2\alpha; n_1 + n_2 - 2) = T_{\text{кр}}(0,1; 58) = 1,67$ левосторонней критической области находим по таблице критических точек распределения Стьюдента при $n_1 + n_2 - 2$ степенях свободы и вероятности 2α . Так как $T_{\text{набл}} < -T_{\text{кр}}$, то отвергаем нулевую гипотезу в пользу альтернативной гипотезы, т.е. утверждение об увеличении среднего значения показателя X является статистически обоснованным.

4.1 Построим диаграмму рассеяния по данным табл. 6.3.

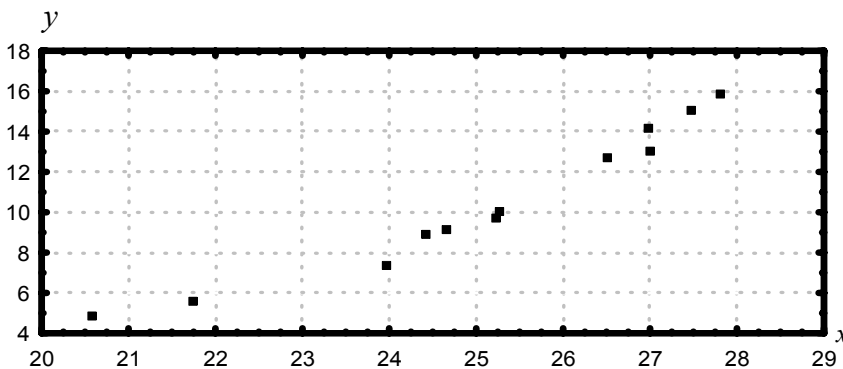


Рис. 6.2

4.2 Построим линейное уравнение регрессии, т.е. модель вида $\hat{y} = a_0 + a_1x$. Согласно методу наименьших квадратов, параметры регрессии a_0, a_1 находим из системы уравнений

$$\begin{cases} na_0 + a_1 \sum x = \sum y; \\ a_0 \sum x + a_1 \sum x^2 = \sum xy. \end{cases}$$

В табл. 6.4 внесем данные, необходимые для расчета коэффициентов линейного уравнения регрессии.

Таблица 6.4

i	x	y	x^2	$x y$
1	20,58	4,88	423,54	100,43
2	21,74	5,66	472,63	123,05
3	23,95	7,41	573,60	177,47
4	24,42	8,95	596,34	218,56
5	24,64	9,17	607,13	225,95
6	25,22	9,75	636,05	245,90
7	25,25	10,05	637,56	253,76
8	26,49	12,72	701,72	336,95
9	26,97	14,21	727,38	383,24
10	27	13,07	729,00	352,89
11	27,46	15,05	754,05	413,27
12	27,79	15,86	772,28	440,75
Сумма	301,51	126,78	7631,28	3272,22

Следовательно, получаем линейное уравнение

$$\begin{cases} 12a_0 + 301,51a_1 = 126,78; \\ 301,51a_0 + 7631,28a_1 = 3272,22. \end{cases}$$

Решая это уравнение, получаем $a_0 = -28,65$, $a_1 = 1,561$. Таким образом, уравнение линейной регрессии имеет вид $\hat{y} = -28,65 + 1,561x$.

4.3 Построим показательное уравнение регрессии, т.е. модель вида $\hat{y} = a_0 a_1^x$. Прологарифмируем это уравнение $\ln \hat{y} = \ln a_0 + x \ln a_1$. Проведем замену переменных $\hat{y}' = \ln \hat{y}$, $a'_0 = \ln a_0$, $a'_1 = \ln a_1$. Таким образом, получаем линейное уравнение регрессии $\hat{y}' = a'_0 + a'_1 x$. Согласно методу наименьших квадратов, параметры регрессии a'_0 , a'_1 находим из системы уравнений

$$\begin{cases} na'_0 + a'_1 \sum x = \sum y'; \\ a'_0 \sum x + a'_1 \sum x^2 = \sum xy'. \end{cases}$$

В табл. 6.5 внесем данные, необходимые для расчета коэффициентов полученного уравнения регрессии.

Таблица 6.5

I	x	y	$\ln y$	x^2	$x \ln y$
1	20,58	4,88	1,59	423,54	32,62
2	21,74	5,66	1,73	472,63	37,68
3	23,95	7,41	2,00	573,60	47,97
4	24,42	8,95	2,19	596,34	53,52
5	24,64	9,17	2,22	607,13	54,60
6	25,22	9,75	2,28	636,05	57,43
7	25,25	10,05	2,31	637,56	58,27
8	26,49	12,72	2,54	701,72	67,37
9	26,97	14,21	2,65	727,38	71,58
10	27	13,07	2,57	729,00	69,40
11	27,46	15,05	2,71	754,05	74,45
12	27,79	15,86	2,76	772,28	76,81
Сумма	301,51		27,56	7631,28	701,70

Таким образом, получаем систему линейных уравнений

$$\begin{cases} 12a_0' + 301,51a_1' = 27,56; \\ 301,51a_0' + 7631,28a_1' = 701,70. \end{cases}$$

Решая эту систему, получим $a_0' = -1,876$, $a_1' = 0,1661$. Далее находим коэффициенты исходного уравнения $a_0 = e^{a_0'} = 0,153$, $a_1 = e^{a_1'} = 1,18$. Таким образом, показательное уравнение регрессии имеет вид $\hat{y} = 0,153 \cdot 1,18^x$.

4.4 Для обеих моделей проверим адекватность по F -критерию. Имеем линейное уравнение $\hat{y}_i = -28,65 + 1,561x_i$ и показательное уравнение $\hat{y}_i = 0,153 \cdot 1,18^{x_i}$. В табл. 6.6 внесем данные, необходимые проверки адекватности моделей. Предварительно найдем среднее значение

$$\bar{y} = \sum y_i / n = 126,78 / 12 = 10,56.$$

Таблица 6.6

I	y	$y_{\text{лин}}$	$y_{\text{показ}}$	$(y_{\text{лин}} - y)^2$	$(y_{\text{показ}} - y)^2$	$(y_{\text{лин}} - \bar{y})^2$	$(y_{\text{показ}} - \bar{y})^2$
1	4,88	3,48	4,61	1,973	0,071	50,26	35,42
2	5,66	5,29	5,59	0,140	0,005	27,87	24,75
3	7,41	8,74	8,06	1,758	0,421	3,35	6,28
4	8,95	9,47	8,71	0,270	0,057	1,20	3,44
5	9,17	9,81	9,03	0,414	0,019	0,565	2,34
6	9,75	10,72	9,94	0,938	0,038	0,02	0,39
7	10,05	10,77	9,99	0,512	0,003	0,04	0,33
8	12,72	12,70	12,27	0,000	0,202	4,56	2,91
9	14,21	13,45	13,28	0,577	0,856	8,32	7,40
10	13,07	13,50	13,35	0,182	0,079	8,60	7,76
11	15,05	14,22	14,41	0,697	0,414	13,32	14,76
12	15,86	14,73	15,22	1,276	0,415	17,35	21,63
Сум- ма				8,73	2,58	135,4	127,40

Проверяем адекватность линейного уравнения регрессии на уровне значимости $\alpha = 0,05$. Вычисляем остаточную дисперсию

$$\sigma_{\text{ост}}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n-1} = \frac{8,73}{11} = 0,79.$$

Вычисляем дисперсию уравнения регрессии

$$\sigma_y^2 = \frac{\sum (\bar{y} - \hat{y}_i)^2}{n-1} = \frac{135,46}{11} = 12,3.$$

Находим наблюдаемое значение критерия

$$F_{\text{набл}} = \frac{\sigma_y^2(n-2)}{\sigma_{\text{ост}}^2} = \frac{12,3 \cdot 10}{0,79} = 155.$$

Находим критическое значение критерия $F(\alpha, 1, n-2) = F(0,05, 1, 10) = 4,96$ по таблице критических точек распределения Фишера при $k_1 = 1$, $k_2 = n - 2 = 10$ степенях свободы и уровне значимости 0,05. Так как $F_{\text{набл}} \gg F(\alpha, 1, n-2)$, то уравнение регрессии является значимым, т.е. статистически подтверждается наличие линейной связи между факторным и результативным признаком.

Проверяем адекватность показательного уравнения регрессии на уровне значимости $\alpha = 0,05$. Вычисляем остаточную дисперсию

$$\sigma_{\text{ост}}^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n-1} = \frac{2,58}{11} = 0,23.$$

Вычисляем дисперсию уравнения регрессии

$$\sigma_y^2 = \frac{\sum (\bar{y} - \hat{y}_i)^2}{n-1} = \frac{127,4}{11} = 11,6.$$

Находим наблюдаемое значение критерия

$$F_{\text{набл}} = \frac{\sigma_y^2 (n-2)}{\sigma_{\text{ост}}^2} = \frac{11,6 \cdot 10}{0,23} = 504.$$

Находим критическое значение критерия $F(\alpha, 1, n-2) = 4,96$. Так как $F_{\text{набл}} > F(\alpha, 1, n-2)$, то уравнение показательной регрессии является значимым.

Таким образом, как линейное, так и показательное уравнения регрессии адекватно описывают экспериментальные данные. Однако остаточная сумма квадратов линейного уравнения регрессии $S_{\text{лин}} = 8,73$ существенно больше остаточной суммы квадратов показательного уравнения регрессии $S_{\text{показ}} = 2,58$. Следовательно, показательное уравнение регрессии является более предпочтительным.

4.5 Вычислим прогнозируемое значение y^* при заданном значении $x^* = 21$. Так как обе модели, как линейная, так и показательная, согласно F -критерию, оказались адекватными, то вычислим прогнозируемое значение по обеим моделям.

По линейной модели $y_{\text{лин}}^* = -28,65 + 1,561x^* = -28,65 + 1,561 \cdot 21 = 4,13$.

По показательной модели $y_{\text{показ}}^* = 0,153 \cdot 1,18^x = 0,153 \cdot 1,18^{21} = 4,95$.

4.6 Вычислим выборочный линейный коэффициент корреляции, используя формулу (4.6). Для вычисления среднеквадратических отклонений вычислим сначала средние значения, используя данные табл. 6.4.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{301,51}{12} = 25,13.$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{126,78}{12} = 10,57.$$

Далее вычисляем выборочные среднеквадратические отклонения по формуле (2.6).

$$s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{(20,58 - 25,13)^2 + (21,74 - 25,13)^2 + \dots + (27,79 - 25,13)^2}{11} = 5,063;$$

$$s_x = \sqrt{5,063} = 2,25;$$

$$s_y^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{(4,88 - 10,57)^2 + (5,66 - 10,57)^2 + \dots + (15,86 - 10,57)^2}{11} = 13,104;$$

$$s_y = \sqrt{13,104} = 3,62.$$

Далее вычисляем коэффициент корреляционной связи, используя значение коэффициента a_1 в уравнении линейной регрессии

$$r = a_1 \frac{s_x}{s_y} = 1,561 \frac{2,25}{3,62} = 0,97.$$

Так как выборочный линейный коэффициент корреляции близок к 1, между переменными существует сильная прямая зависимость.

4.7 Проверим значимость коэффициента корреляции, т.е. проверяем статистическую обоснованность нулевой гипотезы о равенстве нулю генерального коэффициента корреляции $H_0: \rho = 0$ при альтернативной гипотезе $H_1: \rho \neq 0$. Вычисляем наблюдаемое значение критерия Стьюдента по формуле

(4.7) $T_{\text{набл}} = \frac{|r|}{\sqrt{1-r^2}} \sqrt{n-2} = 12,6$. В таблице критических точек распределения Стьюдента для числа степеней свободы $n - 2 = 10$ и вероятности $\alpha = 0,05$ находим значение $T(\alpha; n - 2) = T(0,05; 10) = 2,23$. Так как $T_{\text{набл}} > T(\alpha, n - 2)$, то гипотеза H_0 отвергается, т.е. коэффициент корреляции признается существенно отличающимся от нуля.

приложения

П 1 Значения функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dx$

x	1,28	1,64	1,96	2,33	2,58
$\Phi(x)$	0,4	0,45	0,475	0,49	0,495

П 2 Критические точки $T(\alpha; k)$ распределения Стьюдента

Число степеней свободы k	Вероятность α				
	0,2	0,1	0,05	0,02	0,01
8	1,40	1,86	2,31	2,90	3,36
9	1,38	1,83	2,26	2,82	3,25
10	1,37	1,81	2,23	2,76	3,17
11	1,36	1,80	2,20	2,72	3,11
12	1,36	1,78	2,18	2,68	3,05
13	1,35	1,77	2,16	2,65	3,01
14	1,35	1,76	2,14	2,62	2,98
15	1,34	1,75	2,13	2,60	2,95
16	1,34	1,75	2,12	2,58	2,92
17	1,33	1,74	2,11	2,57	2,90
18	1,33	1,73	2,10	2,55	2,88
19	1,33	1,73	2,09	2,54	2,86
20	1,33	1,72	2,09	2,53	2,85
21	1,32	1,72	2,08	2,52	2,83
22	1,32	1,72	2,07	2,51	2,82
23	1,32	1,71	2,07	2,50	2,81
24	1,32	1,71	2,06	2,49	2,80
25	1,32	1,71	2,06	2,49	2,79
26	1,31	1,71	2,06	2,48	2,78
27	1,31	1,70	2,05	2,47	2,77
28	1,31	1,70	2,05	2,47	2,76
29	1,31	1,70	2,05	2,46	2,76
30	1,31	1,70	2,04	2,46	2,75
32	1,31	1,69	2,04	2,45	2,74
34	1,31	1,69	2,03	2,44	2,73
36	1,31	1,69	2,03	2,43	2,72
38	1,30	1,69	2,02	2,43	2,71
40	1,30	1,68	2,02	2,42	2,70
45	1,30	1,68	2,01	2,41	2,69
50	1,30	1,68	2,01	2,40	2,68
55	1,30	1,67	2,00	2,40	2,67

60	1,30	1,67	2,00	2,39	2,66
65	1,29	1,67	2,00	2,39	2,65
70	1,29	1,67	1,99	2,38	2,65

П 3 Критические точки χ^2 (α ; k) распределения Пирсона

Число степеней свободы	Вероятность α									
	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995
1	7,9	6,6	5,0	3,8	2,7	0,02	0,00	0,00	0,00	0,0000
2	10,6	9,2	7,4	6,0	4,6	0,2	4	1	0,2	4
3	12,8	11,3	9,3	7,8	6,3	0,6	0,1	0,1	0,02	0,01
4	14,9	13,3	11,1	9,5	7,8	1,1	0,4	0,2	0,1	0,1
5	16,7	15,1	12,8	11,1	9,2	1,6	0,7	0,5	0,3	0,2
6	18,5	16,8	14,4	12,6	10,6	2,2	1,1	0,8	0,6	0,4
7	20,3	18,5	16,0	14,1	12,0	2,8	1,6	1,2	0,9	0,7
8	22,0	20,1	17,5	15,5	13,4	3,5	2,2	1,7	1,2	1,0
9	23,6	21,7	19,0	16,9	14,7	4,2	2,7	2,2	1,6	1,3
10	25,2	23,2	20,5	18,3	16,0	4,9	3,3	2,7	2,1	1,7
11	26,8	24,7	21,9	19,7	17,3	5,6	3,9	3,2	2,6	2,2
12	28,3	26,2	23,3	21,0	18,5	6,3	4,6	3,8	3,1	2,6
13	29,8	27,7	24,7	22,4	19,8	7,0	5,2	4,4	3,6	3,1
14	31,3	29,1	26,1	23,7	21,1	7,8	5,9	5,0	4,1	3,6
15	32,8	30,6	27,5	25,0	22,3	8,5	6,6	5,6	4,7	4,1
16	34,3	32,0	28,8	26,3	23,5	9,3	7,3	6,3	5,2	4,6
17	35,7	33,4	30,2	27,6	24,8	10,1	8,0	6,9	5,8	5,1
18	37,2	34,8	31,5	28,9	26,0	10,9	8,7	7,6	6,4	5,7
19	38,6	36,2	32,9	30,1	27,2	11,7	9,4	8,2	7,0	6,3
20	40,0	37,6	34,2	31,4	28,4	12,4	10,1	8,9	7,6	6,8
21	41,4	38,9	35,5	32,7	29,6	13,2	10,9	9,6	8,3	7,4
22	42,8	40,3	36,8	33,9	30,8	14,0	11,6	10,3	8,9	8,0
23	44,2	41,6	38,1	35,2	32,0	14,8	12,3	11,0	9,5	8,6
24	45,6	43,0	39,4	36,4	33,2	15,7	13,1	11,7	10,2	9,3
25	46,9	44,3	40,6	37,7	34,4	16,5	13,8	12,4	10,9	9,9
26	48,3	45,6	41,9	38,9	35,6	17,3	14,6	13,1	11,5	10,5
27	49,6	47,0	43,2	40,1	36,7	18,1	15,4	13,8	12,2	11,2
28	51,0	48,3	44,5	41,3	37,9	18,9	16,2	14,6	12,9	11,8
29	52,3	49,6	45,7	42,6	39,1	19,8	16,9	15,3	13,6	12,5
30	53,7	50,9	47,0	43,8	40,3	20,6	17,7	16,0	14,3	13,1
31	55,0	52,2	48,2	45,0	41,4	21,4	18,5	16,8	15,0	13,8
32	56,3	53,5	49,5	46,2	42,6	22,3	19,3	17,5	15,7	14,5
33	57,6	54,8	50,7	47,4	43,7	23,1	20,1	18,3	16,4	15,1
34	59,0	56,1	52,0	48,6	44,9	24,0	20,9	19,0	17,1	15,8
35	60,3	57,3	53,2	49,8	46,1	24,8	21,7	19,8	17,8	16,5
36	61,6	58,6	54,4	51,0	47,2	25,6	22,5	20,6	18,5	17,2
37	62,9	59,9	55,7	52,2	48,4	26,5	23,3	21,3	19,2	17,9
38	64,2	61,2	56,9	53,4	49,5	27,3	24,1	22,1	20,0	18,6
39	65,5	62,4	58,1	54,6	50,7	28,2	24,9	22,9	20,7	19,3
40	66,8	63,7	59,3	55,8	51,8	29,1	25,7	23,7	21,4	20,0
							26,5	24,4	22,2	20,7

Продолжение табл. П 3

Число степеней свободы	Вероятность α									
	0,005	0,01	0,025	0,05	0,1	0,9	0,95	0,975	0,99	0,995

41	68,1	64,9	60,6	56,9	52,9	29,9	27,3	25,2	22,9	21,4
42	69,3	66,2	61,8	58,1	54,1	30,8	28,1	26,0	23,7	22,1
43	70,6	67,5	63,0	59,3	55,2	31,6	29,0	26,8	24,4	22,9
44	71,9	68,7	64,2	60,5	56,4	32,5	29,8	27,6	25,1	23,6
45	73,2	70,0	65,4	61,7	57,5	33,4	30,6	28,4	25,9	24,3
46	74,4	71,2	66,6	62,8	58,6	34,2	31,4	29,2	26,7	25,0
47	75,7	72,4	67,8	64,0	59,8	35,1	32,3	30,0	27,4	25,8
48	77,0	73,7	69,0	65,2	60,9	35,9	33,1	30,8	28,2	26,5
49	78,2	74,9	70,2	66,3	62,0	36,8	33,9	31,6	28,9	27,2
50	79,5	76,2	71,4	67,5	63,2	37,7	34,8	32,4	29,7	28,0

**П 4 Критические точки $F(\alpha; k_1; k_2)$ распределения
Фишера–Снедекора**

Уровень значимости $\alpha = 0,05$										
k_2	k_1									
	1	9	10	11	12	14	20	30	40	50
1	16	24	24	24	24	24	24	250	251	25
9	1	1	2	3	4	5	8	2,86	2,83	2
10	5,1	3,1	3,1	3,1	3,0	3,0	2,9	2,70	2,66	2,8
11	2	8	4	0	7	3	4	2,57	2,53	0
12	4,9	3,0	2,9	2,9	2,9	2,8	2,7	2,47	2,43	2,6
13	6	2	8	4	1	6	7	2,38	2,34	4
14	4,8	2,9	2,8	2,8	2,7	2,7	2,6	2,31	2,27	2,5
15	4	0	5	2	9	4	5	2,25	2,20	1
20	4,7	2,8	2,7	2,7	2,6	2,6	2,5	2,04	1,99	2,4
30	5	0	5	2	9	4	4	1,84	1,79	0
40	4,6	2,7	2,6	2,6	2,6	2,5	2,4	1,74	1,69	2,3
50	7	1	7	3	0	5	6	1,69	1,63	1
	4,6	2,6	2,6	2,5	2,5	2,4	2,3			2,2
	0	5	0	7	3	8	9			4
	4,5	2,5	2,5	2,5	2,4	2,4	2,3			2,1
	4	9	4	1	8	2	3			8
	4,3	2,3	2,3	2,3	2,2	2,2	2,1			1,9
	5	9	5	1	8	2	2			7
	4,1	2,2	2,1	2,1	2,0	2,0	1,9			1,7
	7	1	6	3	9	4	3			6
	4,0	2,1	2,0	2,0	2,0	1,9	1,8			1,6
	8	2	8	4	0	5	4			6
	4,0	2,0	2,0	1,9	1,9	1,8	1,7			1,6
	3	7	3	9	5	9	8			0

Примечание. Число степеней свободы большей дисперсии – k_1 , меньшей дисперсии – k_2 .