В.А. Тобоев Т.В. Картузова М.С. Толстов

ПРИКЛАДНОЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ДЛЯ ИНЖЕНЕРОВ

Учебное пособие

**Чебоксары
2014**

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Чувашский государственный университет им. И.Н. Ульянова»

В.А. Тобоев Т.В. Картузова М.С. Толстов

ПРИКЛАДНОЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ДЛЯ ИНЖЕНЕРОВ

Учебное пособие

Чебоксары
2014

УДК 519.22 (076.5)
ББК В172я73

Р е ц е н з е н т ы:
доктор физ.-мат. наук, профессор *В.Н. Орлов*
канд. физ.-мат. наук, доцент *Е.А. Григорьев*

Тобоев В.А.
Т50 Прикладной статистический анализ для инженеров: учеб. пособие / В.А. Тобоев, Т.В. Картузова, М.С. Толстов. – Чебоксары: Изд-во Чуваш. ун-та, 2014. – 108 с.

ISBN 978-5-7677-1946-4

Содержит теоретический материал, задания по вариантам, подробный порядок выполнения лабораторных работ по основным разделам «Прикладного статистического анализа для инженеров», а также контрольные вопросы.

Для студентов II курса радиотехнического, машиностроительного, строительного факультетов и факультетов энергетики и электротехники, информатики и вычислительной техники

Ответственный редактор профессор В.Г. Агаков

Утверждено Учебно-методическим советом университета

ISBN 978-5-7677-1946-4

УДК 519.22 (076.5)
ББК В172я73
© Издательство Чувашского университета, 2014
© Тобоев В.А., Картузова Т.В., Толстов М.С., 2014

ПРЕДИСЛОВИЕ

При исследовании технических систем могут использоваться теоретические и эмпирические методы познания. Каждое из этих направлений обладает относительной самостоятельностью, имеет свои достоинства и недостатки. В общем случае теоретические методы в виде математических моделей позволяют описывать и объяснять взаимосвязи элементов изучаемой системы или объекта в относительно широких диапазонах изменения переменных величин. Однако при построении теоретических моделей неизбежно введение каких-либо ограничений, допущений, гипотез, поэтому возникает задача оценки достоверности полученной модели реальному процессу или объекту. Для этого проводится экспериментальная проверка разработанных теоретических моделей. Практика является решающей основой научного познания. В ряде случаев именно результаты экспериментальных исследований дают толчок к теоретическому обобщению изучаемого явления. Экспериментальное исследование дает более точное соответствие между изучаемыми параметрами. Таким образом, теоретические и экспериментальные исследования дополняют друг друга и являются составными элементами процесса познания окружающего нас мира.

Как правило, результаты экспериментальных исследований нуждаются в определенной математической обработке. В настоящее время процедура обработки экспериментальных данных достаточно хорошо формализована и исследователю необходимо только ее правильно использовать. Круг вопросов, решаемых при обработке результатов эксперимента, не так уж велик. Это – вопросы подбора эмпирических формул и оценка их параметров, вопросы оценки истинных значений измеряемых величин и точности измерений, вопросы исследования корреляционных зависимостей и некоторые другие.

Настоящее учебное пособие не претендует на оригинальность. Оно содержит некоторые результаты фундаментальных и прикладных работ в области обработки результатов экспериментальных исследований [1...21]. Пособие может служить практическим руководством по обработке результатов эксперимента как студентам, так и научным сотрудникам и инженерам.

1. ЛАБОРАТОРНАЯ РАБОТА №1. ОЦЕНКА СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК СЛУЧАЙНЫХ ДАННЫХ

1.1. ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

Основными задачами математической статистики являются следующие:

- оценка параметров распределения признака X в генеральной совокупности и определение точности этих оценок;
- проверка выдвинутых гипотез о виде принятого распределения и параметрах закона распределения исследуемой случайной величины X.

Первая задача разделяется на точечное и интервальное оценивание параметров распределения. Например, может возникнуть необходимость по экспериментальным данным получить точечные оценки параметров закона распределения. Такими параметрами, подлежащими вычислению, являются генеральная средняя $\bar{x}_0$, представляющая собой среднее значение признака X в генеральной совокупности, и генеральная дисперсия σ_0^2, характеризующая разброс значений признака вокруг $\bar{x}_0$.

Получение некоторого интервала, с той или иной степенью достоверности содержащего истинное значение параметра, составляет предмет интервального оценивания.

Вторая задача заключается в том, чтобы подтвердить или опровергнуть выдвинутое предположение о законе распределения или о параметре закона распределения случайной величины, используя выборочные (экспериментальные) данные.

Во многих случаях изучить всю генеральную совокупность невозможно, поэтому рассматривают небольшую часть генеральной совокупности называемой выборочной совокупностью или выборкой.

В ходе исследования выборки вычисляются ее основные параметры: выборочная средняя и выборочная дисперсия. Затем по определенному правилу полученные результаты переносятся на основные параметры генеральной совокупности.

Выборка называется случайной, если каждый элемент генеральной совокупности может попасть в выборку с одинаковой вероятностью. Выборка называется бесповторной, если после извлечения из генеральной совокупности и изучения по интересующему признаку элемент не возвращается в генеральную совокупность, и повторной, если возвращается. Ясно, что в бесповторной выборке каждый элемент, попавший в нее, будет изучен только один раз, в то время как в повторной выборке возможно неоднократное изучение элементов.

Объёмом выборки n называется количество содержащихся в ней элементов. Заметим, что объём выборки n значительно меньше объёма генеральной совокупности N. Поэтому оценки параметров выборки должны удовлетворять определённым требованиям. Кроме того, с помощью определенных правил необходима проверка полученных результатов на степень доверия к ним.

Если случайная выборка достаточно хорошо воспроизводит распределение исследуемого признака генеральной совокупности, то она называется *представительной* (*репрезентативной*). Поскольку исследуемые элементы генеральной совокупности попадают в выборку случайным образом, случайным будет и значение параметра, определенное с помощью этой выборки. Поэтому по выборке нельзя точно судить о значениях параметров генеральной совокупности. Численные значения параметров генеральной совокупности, полученные при изучении выборки, называются их *оценками*. Различают два вида оценок: *точечную* и *интервальную*.

Точечные оценки параметров случайной величины

Точечными оценками параметров называют такие оценки, которые выражаются каким-то одним числом (точкой).

Таким числом могут быть, например, параметры a и σ нормального распределения или параметр λ закона Пуассона. Не все переменные могут быть оценками.

Рассмотрим наиболее важные требования, которым должна удовлетворять оценка параметра генеральной совокупности.

1. *Несмещенность.* Оценка не должна содержать систематической ошибки. Это означает, что математическое ожидание оценки некоторого параметра, взятое по всем возможным выборкам, должно быть равно действительному значению параметра.

Если действительное значение оцениваемого параметра обозначить α_0, а его оценку α, то требование несмещенности запишется в виде $M(\alpha)=\alpha_0$. Если это требование не выполняется, то в среднем оценка α будет всегда давать значение α_0 с некоторым отклонением.

2. *Состоятельность.* Оценка α должна приближаться к α_0 по мере увеличения объема выборки. Но ввиду того, что оценка α является случайной величиной, об этом приближении можно говорить только в вероятностном смысле.

Для состоятельности оценки α, получаемой при выборке объема n, должно выполняться условие сходимости по вероятности α к α_0, т.е. $\alpha \xrightarrow[n\to\infty]{} \alpha_0$.

3. *Эффективность***.** Из всех несмещенных и состоятельных оценок следует предпочесть такую, для которой большие отклонения при использовании различных выборок встречались бы как можно реже. Оценкой эффективности несмещенной оценки является ее дисперсия.

Математически требование эффективности означает требование минимальной дисперсии оценки

$$D(\alpha)\to \min .$$

Рассмотрим несмещенные точечные оценки параметров распределения. Запишем в виде таблицы полученные в результате изучения выборки значения признака X в неубывающем порядке $x_0 \le x_1 \le x_2 \le ... \le x_{k-1}$. Такая последовательность называется вариационным рядом, а сами значения вариантами. Если среди вариантов есть одинаковые, то вариационный ряд записывают в виде таблицы.

Варианты, x_i	x_0	x_1	$\ldots$	x_{k-1}
Частоты, m_i	m_0	m_1	$\ldots$	m_{k-1}

При этом сумма частот равна объёму выборки, т.е.

$$m_0 + m_1 + m_2 + \ldots + m_{k-1} = n. \tag{1.1}$$

Величины $\frac{m_0}{n}, \frac{m_1}{n}, \ldots, \frac{m_{k-1}}{n}$ называются относительными частотами или частостями. Из (1.1) следует

$$\frac{m_0}{n} + \frac{m_1}{n} + \ldots + \frac{m_{k-1}}{n} = 1.$$

Часто вариационный ряд, записанный в виде такой таблицы, называют статистическим рядом распределения.

Если во второй строке статистического ряда распределения вместо частот записать частости, то полученная таблица будет иметь большое сходство с рядом распределения дискретной случайной величины.

Тогда для указанной выборки можно построить полигон распределения и выборочную (эмпирическую) функцию распределения аналогично тому, как это делалось для дискретной случайной величины X.

Если значения признака изменяются непрерывно в некотором промежутке, то строят интервальный ряд, имеющий вид:

Интервалы	$[x_0;\ x_1]$	$[x_1;\ x_2]$	…	$[x_{k-2};\ x_{k-1}]$
Частоты	m_0	m_1	…	m_{k-1}

Графическим представлением интервального ряда является гистограмма, пример которой изображен на рис. 1.1.

Поскольку точные значения параметров распределения признака X в генеральной совокупности $\bar{x}_0$ и σ_0^2 в большинстве случаев определить не представляется возможным, то эти параметры следует оценить, используя соответствующие параметры выборки.

Выборочная средняя $\bar{x}_0$ представляет собой среднее значение признака X в выборке и определяется по формуле

$$\bar{x} = \sum_{i=0}^{k-1} x_i \cdot \frac{m_i}{n} = \frac{1}{n}\sum_{i=0}^{k-1} x_i \cdot m_i .$$

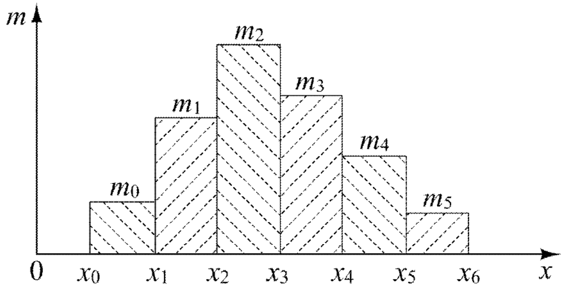

Рис. 1.1. Пример гистограммы интервального ряда для 6 интервалов

Если используется интервальный ряд, то за варианты принимаются, как правило, середины интервалов.

Выборочная дисперсия σ^2 определяется по формуле

$$\sigma^2 = \sum_{i=0}^{k-1} (x_i - \overline{x})^2 \cdot \frac{m_i}{n} = \frac{1}{n} \sum_{i=0}^{k-1} (x_i - \overline{x})^2 \cdot m_i \ ,$$

или по упрощенной формуле

$$\sigma^2 = \overline{x^2} - (\overline{x})^2 \ ,$$

где $\overline{x^2}$ определяется по формуле

$$\overline{x^2} = \sum_{i=0}^{k-1} x_i^2 \cdot \frac{m_i}{n} = \frac{1}{n} \sum_{i=0}^{k-1} x_i^2 \cdot m_i \ .$$

Выборочное среднее квадратическое отклонение σ определяется как квадратный корень из σ^2. В качестве точечной оценки параметров генеральной совокупности может приниматься соответствующий параметр выборки. Можно доказать следующее.

1. $\overline{x}$ является несмещенной точечной оценкой $\overline{x}_0$, т.е.

$$M(\overline{x}) = \overline{x}_0 \ .$$

2. Оценка σ^2 для генеральной дисперсии σ_0^2 является состоятельной, но не смещенной. Поэтому вводят величину

$$s^2 = \sigma^2 \cdot \frac{n}{n-1} \ ,$$

которая называется исправленной статистической выборочной дисперсией. Эта величина является несмещенной оценкой генеральной дисперсии σ_0^2, т.е.

$$M(s^2) = \sigma_0^2 .$$

Величина $s = \sqrt{s^2}$ называется исправленным выборочным средним квадратическим отклонением и является несмещенной точечной оценкой генерального среднего квадратического отклонения σ_0, т.е.

$$M(s) = \sigma_0 .$$

Для решения статистических задач используются специальные распределения случайных величин, сконструированных на основе нормального распределения. Рассмотрим частные случаи этого вида распределения.

Распределение «хи квадрат» χ^2

Важным частным случаем нормального распределения является так называемое распределение χ^2. Это распределение является двухпараметрическим и описывает случайные величины, распределенные на полуинтервале $[0, \infty)$.

Пусть имеется n независимых случайных величин $X_0, X_1, ..., X_{n-1}$, распределенных по нормальному закону с математическим ожиданием, равным нулю, и дисперсией, равной единице. Тогда случайная величина

$$\chi^2 = \sum_{i=0}^{n-1} X_i^2$$

распределена по закону Пирсона или χ^2 – «хи квадрат», с характеристиками $M(\chi^2) = n$, $D(\chi^2) = 2n$.

Распределением χ^2 с n степенями свободы называется распределение суммы квадратов n независимых случайных величин, каждая из которых подчинена нормальному закону с математическим ожиданием, равным нулю, и дисперсией, равной единице.

Очевидно, что χ^2 может принимать лишь неотрицательные значения. График плотности распределения случайной величины χ^2 представляет собой кривую, изображенную на рис. 1.2.

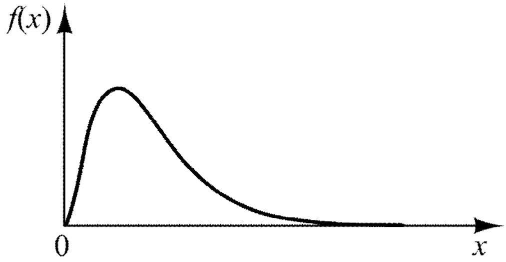

Рис. 1.2. Кривая плотности распределения χ^2

Для того чтобы определить вероятность попадания случайной величины χ^2 в какой-либо промежуток из множества положительных чисел, пользуются таблицей распределения χ^2 (прил. 2).

Обычно такая таблица позволяет по заданной вероятности α и по числу степеней свободы n определить так называемый квантиль $\chi^2_{кр}$ – критическое значение χ^2, если α и $\chi^2_{кр}$ связаны соотношением:

$$P\left(\chi^2 > \chi^2_{кр}\right) = \alpha .$$

Эта формула означает: вероятность того, что случайная величина χ^2 примет значение большее, чем табличное значение $\chi_{кр}$, равна α. Из таблицы (прил. 2) очевидно, что случайная величина χ^2 с 10 степенями свободы с вероятностью $\alpha = 0{,}95$ принимает значение, равное 3,94.

Распределение Стьюдента

Другим важным случаем нормального распределения является однопараметрическое распределение Стьюдента (рис. 1.3). Ему подчиняется случайная величина, распределенная на всей числовой оси. Пусть $X_0, X_1, ..., X_{n-1}$ – независимые стандартные нормальные случайные величины $X_i \in N(0; 1)$, $i = 0,1,2,...,n-1$. Тогда случайная величина

$$t_n = \frac{X}{\sqrt{\frac{1}{n}\sum_{i=0}^{n-1} X_i^2}}$$

имеет распределение Стьюдента с n степенями свободы.

Если число степеней свободы $n > 2$, то

$$M(t_n) = 0\,,\ D(t_n) = \frac{n}{n-2}\,.$$

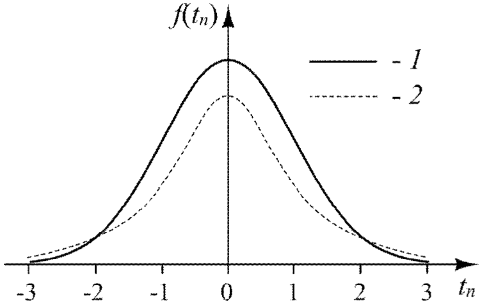

Рис. 1.3. Стандартное нормальное распределение (*1*)
и распределение Стьюдента (*2*)

При $n \to \infty$ распределение Стьюдента стремится к нормальному с математическим ожиданием 0 и дисперсией 1. Уже при $n > 30$ распределение Стьюдента можно приближенно заменить на стандартное нормальное.

Распределение Фишера

Пусть

$$X_i \in N(0;1),\ i = \overline{0, m-1};\ Y_j \in N(0;1),\ j = \overline{0, n-1}$$

независимые случайные величины, где $x = \sum_{i=0}^{m-1} X_i^2$ и $y = \sum_{j=0}^{n-1} Y_j^2$.

Тогда случайная величина

$$F_{m,n} = \frac{x/m}{y/n}$$

распределена по закону Фишера со степенями свободы m, n с математическим ожиданием (если $n > 2$)

$$M\left(F_{m,n}\right) = \frac{n}{n-2} \text{ и } D\left(F_{m,n}\right) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}, \text{ если } n > 4.$$

При заданных числах m и n, вероятности α по прил. 6 определяется значение $F_{m,n,\alpha}$ такое, что

$$P\left(F > F_{m,n,\alpha}\right) = \alpha.$$

Обычно таблицы составляются для значений α, равных 0,05 или 0,01, а иногда для обоих этих значений. В таблицах значений закона Фишера иногда вместо обозначений m и n используются k_1 и k_2 соответственно. Пример плотности распределения случайной величины $F_{m,n}$ приведён на рис. 1.4.

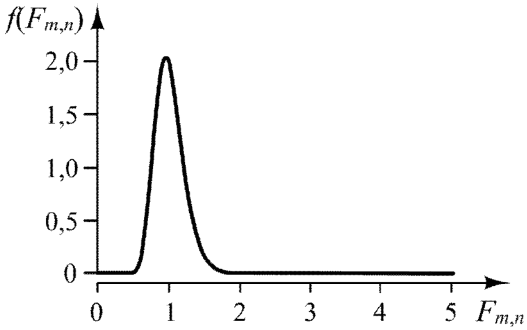

Рис. 1.4. Плотность распределения случайной величины $F_{m,n}$ при $m = 100$ и $n = 100$

Интервальные оценки параметров случайной величины

Пусть по результатам опыта построена точечная оценка α параметра α_0. Точечная оценка α параметра α_0 дает лишь некоторое приближенное значение α_0.

Возникает вопрос: насколько эта оценка точна и достоверна? Чтобы получить представление о точности и надежности оценки, используют интервальную оценку параметра.

Интервальная оценка параметра α_0 ($\bar{x}_0$ или σ_0^2) признака X генеральной совокупности имеет следующий вид:

$$\alpha - \Delta < \alpha_0 < \alpha + \Delta \text{ или } \alpha_0 \in (\alpha - \Delta, \alpha + \Delta),$$

где Δ – наибольшее отклонение выборочного значения параметра от его истинного значения, или предельная ошибка выборки (точность оценки, отклонение).

Очевидно, что это неравенство, определяющее попадание α_0 в указанный интервал, при заданной точности оценки Δ верно лишь с какой-то вероятностью γ, которая называется доверительной вероятностью.

На практике обычно заранее задают доверительную вероятность γ, причем наиболее часто берут значения $\gamma = 0{,}95$; $\gamma = 0{,}99$ и $\gamma = 0{,}999$.

Интервал $(\alpha - \Delta, \alpha + \Delta)$ называется доверительным интервалом. В качестве α берут точечную несмещенную оценку α_0, полученную по результатам выборки. Следовательно, задача состоит в том, чтобы по заданной величине γ найти Δ.

Чтобы получить интервальную оценку генеральной средней $\bar{x}_0$, нужно найти такую величину Δ, для которой

$$P(\bar{x} - \Delta < \bar{x}_0 < \bar{x} + \Delta) = \gamma. \qquad (1.2)$$

Пусть известна генеральная дисперсия σ_0^2 признака X. Так как неравенство $\bar{x} - \Delta < \bar{x}_0$ эквивалентно неравенству $\bar{x} < \bar{x}_0 + \Delta$, а неравенство $\bar{x}_0 < \bar{x} + \Delta$ эквивалентно неравенству $\bar{x}_0 - \Delta < \bar{x}$, то выражение (1.2) эквивалентно выражению

$$P(\bar{x}_0 - \Delta < \bar{x} < \bar{x}_0 + \Delta) = \gamma.$$

Здесь $\bar{x}$ – случайная величина, и если признак X распределен нормально, то $\bar{x}$ тоже распределена нормально. Так как $\bar{x}$ – несмещенная оценка $\bar{x}_0$, то $M(\bar{x}) = \bar{x}_0$.

Можно доказать, что если дисперсия признака X равна σ_0^2, то дисперсия случайной величины $\bar{x}$, являющейся средним арифметическим n одинаково распределенных случайных величин с дисперсией σ_0^2, равна σ_0^2/n, а среднее квадратическое отклонение случайной величины $\bar{x}$, следовательно, равно $\sigma_0/\sqrt{n}$.

Тогда можно также показать, что

$$P(\bar{x}_0 - \Delta < \bar{x} < \bar{x}_0 + \Delta) = 2\Phi\left(\frac{\Delta \cdot \sqrt{n}}{\sigma_0}\right),$$

где $\Phi(x)$ – функция Лапласа (прил. 1).

Отсюда следует, что для нахождения Δ нужно решить уравнение $\Phi\left(\frac{\Delta \cdot \sqrt{n}}{\sigma_0}\right) = \frac{\gamma}{2}$.

Уравнение решается так. Обозначим $t = \frac{\Delta \cdot \sqrt{n}}{\sigma_0}$ и найдем по таблице значений функции Лапласа $\Phi(t)$ (прил. 1) такое t, что $\Phi(t) = \gamma/2$.

Вычислим окончательно Δ по следующей формуле

$$\Delta = \frac{t\sigma_0}{\sqrt{n}}.$$

Следовательно, доверительный интервал $\bar{x}_0$ для доверительной вероятности γ имеет вид

$$\bar{x}_0 \in \left(\bar{x} - \frac{t\sigma_0}{\sqrt{n}};\ \bar{x} + \frac{t\sigma_0}{\sqrt{n}}\right).$$

Такую оценку используют, когда объем выборки $n > 30$.

Если генеральная дисперсия σ_0^2 неизвестна, то она часто заменятся ее несмещенной оценкой s^2, хотя при небольших n это ведет к существенному уменьшению доверительного интервала.

Доверительный интервал получается точнее, если при этом вместо t взять параметр t_γ распределения Стьюдента, который можно найти в прил. 3 по известным значениям γ и n.

Тогда доверительный интервал для генеральной средней при неизвестной генеральной дисперсии имеет вид

$$\bar{x}_0 \in \left(\bar{x} - \frac{t_\gamma s}{\sqrt{n}}; \bar{x} + \frac{t_\gamma s}{\sqrt{n}}\right).$$

Интервальной оценкой (с надёжностью γ) среднего квадратического отклонения σ нормально распределенного количественного признака X «по исправленному» выборочному среднему квадратическому отклонению s служит доверительный интервал

$$s(1-q)<\sigma<s(1+q), \quad \text{при } q<1$$

$$0<\sigma<s(1+q), \quad \text{при } q>1,$$

где q находят из таблицы (прил. 4) по заданным n и γ.

Проверка статистических гипотез

Статистическая гипотеза – это предположение относительно вида закона распределения или параметров распределения в генеральной совокупности некоторого признака X, являющегося случайной величиной.

По своему содержанию статистические гипотезы можно подразделить на несколько основных типов.

1. Гипотезы о виде закона распределения исследуемой случайной величины.
2. О числовых значениях параметров случайной величины.
3. Об общем виде модели, описывающей статистическую зависимость между признаками.
4. О принадлежности некоторого признака к тому или иному классу величин.

Естественно, как любая гипотеза, статистическая гипотеза нуждается в проверке. Проверяемая гипотеза называется основной или нулевой и обозначается H_0.

Гипотеза H_0 является основной в том смысле, что было бы желательно убедиться в ее справедливости (например, H_0 – благополучная посадка самолета, выздоровление больного, успех в коммерческой деятельности и т.д.).

Если кроме нулевой выдвигаются и другие гипотезы, то они называются конкурирующими и обозначаются H_1, H_2 и т.д.

Правило, по которому принимается решение принять или отклонить H_0, называется статистическим критерием K. Проверка статистической гипотезы проводится путем рассмотрения некоторого параметра, определяемого по результатам изучения выборки. Этот параметр называется статистикой критерия.

Задача статистической проверки гипотез

Пусть по некоторым данным имеются основания выдвинуть предположения о законе распределения или о параметре закона распределения случайной величины (или генеральной совокупности, на множестве объектов которой определена эта случайная величина). Задача заключается в том, чтобы подтвердить или опровергнуть это предположение, используя выборочные (экспериментальные) данные.

В качестве примера критерия согласия рассмотрим статистический критерий χ^2 Пирсона. Этот критерий служит для проверки гипотез о виде закона распределения признака X в генеральной совокупности.

При проверке гипотез все множество возможных значений признака X разбивается на k непересекающихся областей $A_0, A_1, \ldots A_{k-1}$. Частоту попадания элементов выборки в область $A_i, i = 0,1,\ldots,k-1$ обозначим m_i. Ясно, что должно выполняться равенство

$$\sum_{i=0}^{k-1} m_i = n .$$

Теоретическую частоту попадания элементов выборки в область A_i обозначим m_i'. Эта частота определяется для каждой области по формуле

$$m_i' = np_i ,$$

где p_i – вероятность попадания признака X в A_i в соответствии с предполагаемым законом распределения.

Если неизвестны точные значения $\bar{x}$ и σ_0^2, необходимые для расчета вероятностей p_i, то в качестве параметров распределения берутся их несмещенные точечные оценки $\bar{x}$ и s^2.

Заметим, что критерий χ^2 Пирсона следует применять, если каждая частота m_i в рассматриваемой области имеет значение не менее $\alpha \cdot n$. В противном случае нужно укрупнять области разбиения исследуемого признака.

В качестве статистики критерия Пирсон предложил рассматривать величину

$$\chi^2 = \sum_{i=0}^{k-1} \frac{(m_i - m_i')^2}{m_i'} .$$

Сумма квадратов n независимых одинаково распределенных случайных величин подчиняется закону, получившему название распределения χ^2.

Число независимых слагаемых, входящих в сумму квадратов, называется числом степеней свободы и обозначается ν.

Так как сумма квадратов – величина всегда неотрицательная, то кривая распределения лежит в области от нуля до $+\infty$ (рис. 1.4). При $\nu > 2$ плотность распределения изображается скошенной кривой с максимумом в точке

$$\chi^2 = \nu - 2 .$$

Заметим, что вероятность превышения какого-либо значения χ^2 равна площади под кривой распределения от этого значения до $+\infty$. Поэтому за критическое значение статистики принимается значение, площадь под кривой после которого равна уровню значимости α. На рис. 1.4 критическая область критерия заштрихована. Значения $\chi^2_{\text{кр}}$ в зависимости от ν и α представлены в приложении.

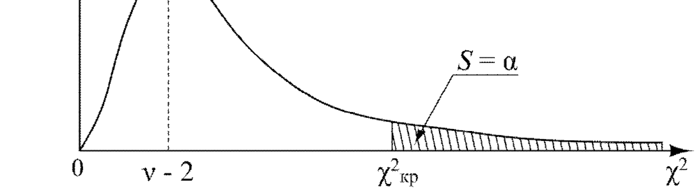

Рис. 1.4. Кривая распределения Пирсона

Используя рассчитанные по результатам изучения выборки оценки параметров предполагаемого распределения, вводят дополнительные ограничения. Поэтому число степеней свободы ν должно быть уменьшено по сравнению с числом слагаемых k , и определяется выражением

$$\nu = k - r - l,$$

где k – число рассматриваемых областей; r – число параметров в принятом распределении; l – число независимых линейных ограничивающих связей.

В рассматриваемом случае $l = 1$ из-за одной имеющейся связи: равенство суммы частот объему выборки.

При проверке гипотезы о нормальном распределении признака X с заранее неизвестными параметрами a и σ величина r равна 2. Тогда выражение для ν при исследовании нормально распределенной величины имеет вид

$$\nu = k - 3.$$

По значениям ν и α находят значение $\chi^2_{\text{кр}}$ по таблицам (прил. 2). Если $\chi^2 < \chi^2_{\text{кр}}$, то гипотеза H_0 принимается, а если $\chi^2 \geq \chi^2_{\text{кр}}$, то эта гипотеза отклоняется.

1.2. ЗАДАНИЯ ДЛЯ ЛАБОРАТОРНОЙ РАБОТЫ №1. ОЦЕНКА СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК СЛУЧАЙНЫХ ДАННЫХ

По статистическим данным, полученным в результате проведения опыта, требуется:

1. Произвести группировку, построить статистическое распределение относительных частот и изобразить его графически.

2. Найти эмпирическую функцию распределения и построить ее график.

3. Вычислить выборочную среднюю, выборочную дисперсию, выборочное среднее квадратическое отклонение, моду, медиану.

4. С надежностью 0,99 найти доверительный интервал для истинного значения рассматириваемой величины.

5. Построить теоретическую нормальную кривую.

6. Предполагая о нормальном распределении генеральной совокупности и пользуясь критерием χ^2 на уровне значимости 0,01, установить случайно или значимо расхождение между формой распределения выборки и генеральной совокупности.

Порядок выполнения задания

1. Провести группировку выборки, разбив весь интервал на 10 частичных интервалов одинаковой длины $h=(x_{\max}-x_{\min})/10$, подсчитать сколько значений признака попадает в каждый частичный интервал (значения, совпадающие с граничными, следует отнести к левому интервалу).

2. Построить эмпирическую функцию распределения по формуле

$$F(x)=n_x/n,$$

где n_x – число вариант меньших x.

3. Статистические оценки параметров распределения провести по формулам

$$\bar{x}_{\text{в}}=\sum_{i=0}^{k-1} x_i n_i/n,$$

$$D_{\text{в}}=\frac{\sum_{i=0}^{k-1}(x_i-\bar{x}_{\text{в}})^2 n_i}{n-1},$$

$$s=\sqrt{D_{\text{в}}}.$$

4. Найти доверительный интервал для математического ожидания m_x из неравенств

$$\bar{x}_{\text{в}}-t_\gamma s/\sqrt{n}<m_x<\bar{x}_{\text{в}}+t_\gamma s/\sqrt{n},$$

где $t_\gamma(v)$ – параметр, значение которого берется из таблицы в зависимости от объема выборки n и надежности γ. Число степеней свободы v определяется как объем выборки n минус количество параметров d в принятом распределении ($v=n-d$).

Доверительный интервал для среднего квадратического отклонения σ найти по формуле

$$s(1-q) < \sigma < s(1+q), \quad \text{при } q < 1$$

$$0 < \sigma < s(1+q), \quad \text{при } q > 1,$$

где q – параметр, значение которого зависит от объема выборки n и надежности γ берется из таблицы (прил. 4).

5. Построить нормальную кривую.
6. Вычислить значение критерия $\chi^2_{\text{расч}}$

$$\chi^2_{\text{расч}} = \sum_{i=0}^{k-1} \frac{\left(m_i^0 - m_i^t\right)^2}{m_i^t}$$

здесь m_i^0 – наблюдаемые частоты для каждой группы, m_i^t – теоретические частоты, вычисленные в предположении о нормальности распределения.

Теоретические частоты для каждой группы рассчитываются по формуле

$$f_i^t = p_i n, \quad i = \overline{0, k-1},$$

где p_i вероятность попадания в группу определяется как разность значений функции распределения между границами соответствующей группы. Для этого исходные границы групп нормируются и центрируются с использованием формулы

$$z_i = \frac{x_i - \overline{x}_{\text{в}}}{s}.$$

В заключении расчетная величина критерия $\chi^2_{\text{расч}}$ сравнивается с табличным значением $\chi^2_{\text{кр}}$ для заданного уровня значимости и числа степеней свободы $m = k - d - 1$, где k – количество групп после объединения, если она оказалась необходимой (в группах частоты должны быть больше 5), d – количество параметров, определенных по выборке, для нахождения теоретических частот (для нормального распределения $d = 2$).

Если $\chi^2_{\text{расч}} < \chi^2_{\text{кр}}$, то гипотеза о нормальном распределении принимается. А если $\chi^2_{\text{расч}} \ge \chi^2_{\text{кр}}$, то гипотеза отклоняется.

1.3. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Сформулируйте основные задачи математической статистики.
2. В чем разница между генеральной и выборочной совокупностями?
3. Перечислите основные точечные оценки параметров случайной величины.
4. Графическое представление данных выборки.
5. Перечислите основные статистические распределения.
6. Какие основные интервальные оценки параметров случайной величины вы знаете?
7. Сформулируйте основной принцип проверки статистических гипотез.
8. Основные типы статистических гипотез.
9. Что такое критические области? Способы отыскания критических областей.
10. Сформулируйте правило проверки гипотезы о нормальном распределении генеральной совокупности.

1.4. ВАРИАНТЫ ЗАДАНИЙ

Вариант №1

26,2	26,6	25,8	27,7	27,3	27,2	27,9	30,5	34,8	31,7
30,2	29.5	29,2	28,9	28,8	28,3	28,6	28,8	28,9	29,1
29,2	29,3	29,3	29,4	29,4	29,3	29,3	29,4	29,5	28,6
28,6	28,8	28,9	28,4	28,3	25,7	26,6	26,8	27,1	27,3
27,6	27,8	27,9	28,1	28.2	28,3	28,4	28,5	28,5	28,6
28,9	28,9	29,3	29,5	29,7	30,1	30,2	30,4	30,8	31,2
31,9	32,6	31,8	30,9	30,4	30,2	29,9	29,8	29,8	29,7
29,7	29,7	29,7	29,8	29,8	29,9	29,9	29,3	30,1	30,1
29,9	29,9	29,9	30,1	30,2	30,3	30,6	30,7	30,8	30,2
33,4	34,3	34,0	32,5	32,0	32,1	31,4	26,9	26,0	25,9

Вариант №2

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

Вариант №3

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

Вариант №4

13,2	11,9	11,9	13,4	13,4	13,3	11,9	12,1	12,6	13,9
10,7	12,3	10,6	10,4	10,6	11,0	11,0	10,8	10,8	10,6
10,9	11,9	11,6	11,9	11,3	11,9	11,4	11,3	11,0	10,8
10,9	10,9	11,0	11,2	11,8	11,8	12,7	12,9	12,4	14,2
14,8	14,8	15,4	14,6	14,1	13,3	12,6	11,1	11,2	11,6
11,6	11,5	11,2	12,9	14,1	14,0	12,5	12,5	10,8	10,9
11,3	11,4	11,0	11,9	12,4	12,9	13,0	13,5	14,3	14,6
14,8	14,5	14,0	13,2	12,2	12,1	12,2	11,8	11,2	10,9
10,6	10,8	10,8	10,9	10,5	10,7	12,4	12,8	13,6	14,0
14,2	14,1	14,2	13,9	13,6	13,9	12,9	13,0	12,5	12,6

Вариант №5

9,6	9,6	9,8	9,5	9,9	9,4	9,5	10,3	11,8	11,9
13,3	13,7	15,4	15,7	16,4	18,1	17,5	17,6	17,1	13,3
16,8	16,4	15,7	16,0	17,6	18,4	18,3	18,3	17,4	16,6
17,5	17,6	17,1	16,6	17,2	16,4	16,1	13,7	13,5	13,7
13,5	11,7	10,8	9,7	9,3	9,4	9,5	9,3	9,6	9,7
9,5	9,4	10,2	9,7	10,4	10,4	10,3	9,3	9,8	9,7
9,7	9,7	10,0	9,8	10,1	9,8	10,1	12,3	13,1	14,2
13,7	15,3	14,9	16,2	17,4	17,8	17,5	16,4	15,8	15,6
14,1	13,7	13,5	12,5	12,2	11,9	11,6	10,9	10,4	9,9
9,8	10,7	11,3	11,9	13,5	14,7	15,3	15,6	15,1	14,8

Вариант №6

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

Вариант №7

26,7	26,9	26,8	26,8	26,5	26,6	26,5	26,4	26,4	26,9
25,6	25,3	25,5	25,7	25,7	28,6	28,2	29,2	28,4	29,5
28,4	27,9	27,6	27,1	26,8	26,6	26,9	27,0	27,0	27,1
27,2	27,1	27,2	27,3	27,2	27,5	27,5	27,6	27,6	27,8
27,7	27,7	27,4	27,6	27,8	28,6	28,1	26,1	24,3	24,8
27,7	26,1	27,2	23,7	24,2	24,2	24,8	25,0	25,3	26,1
26,4	26,6	26,7	27,1	27,1	27,2	26,9	24,7	26,2	27,4
28,8	24,9	25,3	25,4	25,4	25,7	26,3	28,1	28,8	29,1
29,5	29,5	29,3	28,4	29,1	28,6	27,9	27,7	27,3	26,7
26,9	28,8	28,8	28,3	28,1	28,1	27,5	26,8	26,4	26,1

Вариант №8

18,2	14,9	14,1	13,4	18,1	12,2	11,6	10,8	10,6	12,9
17,8	14,7	14,1	13,5	18,1	12	11,5	10,8	10,3	13,4
17,3	14,8	13,9	13,6	18,3	11,9	11,3	10,7	10,2	13,7
16,8	14,7	13,8	13,8	19,1	11,7	11,2	10,7	10,1	13,5
16,3	14,4	13,8	13,9	18,5	11,3	11,2	10,8	10,2	13,2
15,9	14,4	13,7	14,2	18,2	11,5	11,2	10,8	10,2	13,7
15,6	14,4	13,6	14,6	18	11,4	11,1	10,6	10,9	13,6
15,3	14,4	13,5	15,2	14,7	11,6	10,9	10,4	11,4	14,3
15,2	14,2	13,4	16,7	12,7	11,7	10,9	10,5	11,7	14,3
14,9	14,1	13,4	18,1	12,2	11,6	10,8	10,6	12,9	14,9

Вариант №9

10,3	10,8	11,8	17,2	18,6	18,4	14,2	14,1	19,6	23,4
10,4	10,8	12,3	16,7	18,9	17,7	13,9	14,5	20,8	23,2
10,7	10,8	12,6	16,5	18,7	17,2	13,6	14,9	21,3	22,9
10,8	10,8	12,8	17,4	18,3	16,8	13,3	15,7	21,7	22,7
10,9	10,8	13,3	16,8	18,1	16,6	13,4	16,4	21,8	22,4
10,6	10,9	15,2	17,3	18,2	16,1	13,2	17,1	22,9	21,4
10,3	10,9	15,2	17,5	18,3	15,8	14,2	17,7	22,8	19,9
10,6	11,1	16,2	17,3	18,4	15,6	14,3	18,2	23,4	19,6
10,7	11,4	16,9	18	18,7	14,8	14,2	18,6	23,1	18,8
10,8	11,8	17,2	18,6	18,4	14,2	14,1	19,6	23,4	18,3

Вариант №10

14,7	13,7	18,8	23,2	19,7	17	15,7	15,7	18,9	15,9
14,3	13,4	19,8	23,6	19,2	16,4	15,7	16,5	18,2	16
14,2	13,1	20,5	22,6	18,3	16,3	15,3	17,1	17,8	14,4
13,9	13,2	21,6	22,4	18,2	16,1	15,1	17,4	17,4	13,8
14	13,3	22,3	22,5	17,8	16,2	14,8	18	17,1	13,6
14,5	14	22,3	22,1	17,5	15,9	14,7	18,7	16,7	13,3
14,4	15,1	22,3	21,6	17,4	15,9	14,7	18,3	16,6	13
14,4	15,8	22,3	21,2	16,9	15,8	14,9	18,9	16,3	12,8
14	17,3	22,7	20,2	17,2	15,8	15,3	18,7	16,1	12,6
13,7	18,8	23,2	19,7	17	15,7	15,7	18,9	15,9	12,4

Вариант №11

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

Вариант №12

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

Вариант №13

12,8	12,8	13,1	13,1	12,5	12,9	12,9	12,8	12,8	13,1
12,5	12,8	13,1	12,5	12,8	12,7	12,9	12,7	13,6	12,4
12,6	12,7	13,1	11,9	12,8	12,9	12,8	12,7	16,7	13
12,5	13	13	12	12,9	12,7	12,8	12,7	16	13,4
12,7	13	13,2	12,1	12,9	12,8	13	12,6	12,8	14
12,6	13	13,2	12,9	12,9	12,9	13	12,6	12,4	14,2
12,6	12,9	13,4	12,8	12,8	12,6	12,9	12,6	12,2	14,2
12,8	12,9	13,1	13,4	13	12,8	12,9	12,6	12,7	14,2
12,8	13,1	13,1	13,3	13	12,7	12,8	12,6	13,2	13,9
12,8	13,1	13,1	12,5	12,9	12,9	12,8	12,8	13,1	13,8

Вариант №14

16	13,1	10,3	9,8	7,1	6,7	7,6	7,4	6,6	5,6
15,8	12,7	10,3	9,8	7	6,7	7,7	7	6,5	5,3
15,8	11,9	10,2	9,6	6,5	6,8	7,6	6,9	6,6	5,2
15,7	11,8	10	9	6,6	7,2	7,6	7	6,3	5,3
15,7	11,6	10	8,9	5,8	7,2	7,6	6,7	6,6	5,3
15,6	11,2	9,9	8,5	6	7,4	7,6	6,6	6,1	5,4
15,5	11	9,8	8,1	6,3	7,6	7,6	6,5	6,2	5,6
14,9	10,8	9,9	8	6,7	7,4	7,1	6,6	6	5,7
13,7	10,6	9,8	7,4	6,5	7,8	7	6,8	5,6	5,7
13,1	10,3	9,8	7,1	6,7	7,6	7,4	6,6	5,6	5,7

Вариант №15

-11,4	-5	-1,7	-7,7	-10,9	-12,5	-11,7	-10,7	-11,1	-12,2
-11,2	-4,8	-1,7	-8,1	-11,3	-12,4	-11,5	-10,6	-11	-12,1
-10,6	-4,2	-3,1	-8	-11,3	-12,3	-11,1	-10,5	-11,5	-12,9
-10	-3,7	-4,1	-8,7	-11,6	-12,5	-11,3	-10,4	-11,4	-13,8
-9,1	-3,4	-4,5	-8,9	-11,7	-12	-11	-10,6	-11,4	-13,9
-8,4	-2,9	-5,2	-9,3	-11,8	-12,1	-10,9	-10,7	-11,3	-14,3
-7,5	-2,4	-6,1	-9,6	-11,8	-12,3	-10,7	-10,8	-11,3	-14,3
-6,5	-2,1	-6,3	-9,9	-11,9	-12	-10,6	-10,9	-11,8	-14,2
-5,9	-1,9	-7	-10,2	-12	-11,8	-10,8	-11	-11,9	-14,7
-5	-1,7	-7,7	-10,9	-12,5	-11,7	-10,7	-11,1	-12,2	-14,8

Вариант №16

83,5	95,9	100	84,6	80,7	75,4	80,1	97,1	96,1	99,1
84,8	100	100	80,8	79,2	71,2	90,5	100	96,9	99,7
82,3	100	100	88,9	81	79,8	84,8	100	96,1	100
82	100	100	79,8	79,7	81,9	90,9	100	96,9	98,8
85,7	100	100	84,6	81,3	73,4	94,7	100	96,8	99,3
87,6	100	100	87,4	79,5	86,7	100	100	94	99,4
87,7	100	100	84,6	80	73,2	99	100	100	99
85,5	100	100	81,7	83,6	75,6	97,1	99,4	100	99,3
84,2	100	94,4	76,5	81,3	69,6	96,5	100	99,6	100
95,9	100	84,6	80,7	75,4	80,1	97,1	96,1	99,1	100

Вариант №17

3,4	3	3	3	3,4	3,2	3,4	3,6	3,8	4,1
3,4	3	3	3,3	3,4	3,4	3,4	3,6	3,9	4,1
3,2	3	3	3,3	3,4	3,2	3,4	3,6	3,9	4,4
3,2	3	3	3,4	3,4	3,2	3,4	3,6	4,1	4,4
3,2	3	3	3,8	3,2	3,2	3,4	3,6	3,9	4,4
3,2	3	3	3	3,2	3,2	3,4	3,8	4,1	4,4
3	3,1	3	4	3,2	3,2	3,4	3,8	4,1	4,6
3	3,1	3	4	3,2	3,4	3,6	3,8	4,1	4,6
3	3,1	3	3,7	3,2	3,4	3,6	3,8	4,1	4,6
3	3	3	3,4	3,2	3,4	3,6	3,8	4,1	4,6

Вариант №18

9,9	10,6	11,1	11,3	11,1	9,7	11,1	11,6	11,4	10,6
9,9	10,6	11,1	11,3	11	9,9	11,1	11,6	11,4	9,9
10	10,8	11,1	11,3	11	10,3	11,3	11,4	11,4	10
10	10,8	11,1	11,3	11,1	10,3	11,3	11,4	11,4	10,2
10	10,8	11,1	11,3	11,1	10,5	11,4	11,4	11,4	9,9
10,2	10,8	11,1	11,3	11,3	10,5	11,4	11,4	11,4	9,8
10,4	11	11,1	11,3	11	10,7	11,4	11,4	11,4	9,8
10,4	11	11,1	11,3	10,3	10,8	11,4	11,4	11,3	9,8
10,4	11	11,3	11,1	10,1	11	11,6	11,4	10,8	9,8
10,6	11,1	11,3	11,1	9,7	11,1	11,6	11,4	10,6	9,9

Вариант №19

10,8	10,7	11,1	11,1	11,4	13,5	12,5	12	11,3	12,1
10,8	10,6	11,1	11,1	12,1	13	12,2	11,9	11,3	11,1
10,9	10,8	11	11,1	12,3	12,9	12,2	11,7	11,3	11
10,8	10,8	11	11,1	12,6	12,6	12,1	11,6	11,2	11,8
10,6	10,8	11	11,1	12,6	12,7	12,2	11,6	11,2	13
10,7	10,6	10,9	11,3	12,7	12,6	12,1	11,5	11,2	12,4
10,8	10,9	10,9	11,3	13,2	12,6	12	11,5	11,2	12,4
10,7	10,9	10,9	11,5	13,3	12,5	11,9	11,4	11,5	12,4
10,5	10,8	11,1	11,4	13,6	12,5	12,1	11,4	11,8	12,3
10,7	11,1	11,1	11,4	13,5	12,5	12	11,3	12,1	13

Вариант №20

25,3	16,2	14,4	13,2	14,3	20,4	19,7	14,8	12,9	15,5
19,1	15,6	14,3	13,1	14,4	21,9	19	14,2	12,8	18,3
18,6	15,3	14,2	13,1	15,3	21,2	19,5	14,2	12,6	15,8
18,3	15,2	14	13	16,7	21,8	19,1	14,4	12,7	16,6
18,4	14,9	13,9	13,2	20,6	22,1	17,5	14,3	12,5	18,2
18,3	14,6	13,9	13,8	23	22,7	16,8	13,6	12,2	16,7
16,8	14,8	13,8	14,2	20,4	23,1	16,3	13,6	12,3	16,9
16,6	14,9	13,4	15,2	23,2	22,2	16	13,4	12,9	20,6
16	14,6	13,3	14,3	23,5	21,2	15,3	12,9	13,9	21,2
16,2	14,4	13,2	14,3	20,4	19,7	14,8	12,9	15,5	21,2

Вариант №21

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

Вариант №22

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

Вариант №23

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

Вариант №24

-16,8	-7,9	-3,2	-9,7	-13,9	-19,7	-19,7	-17,8	-13,6	-14,6
-16,3	-6,8	-4,6	-9,4	-15,3	-19,2	-19,5	-18	-14	-16,2
-15,4	-6,8	-7	-10,2	-16,1	-18,4	-19,7	-17,4	-12,9	-17,8
-14,5	-6,5	-7,4	-9,7	-17,3	-18,1	-19,6	-16,8	-12,9	-17,3
-13,3	-5,5	-8,6	-9,1	-17,8	-19,8	-18,9	-17,2	-11,5	-17,8
-10,8	-4,8	-8,8	-9,5	-18,1	-20,3	-18,7	-17,7	-11,6	-17,1
-10,1	-4,6	-7,9	-9,6	-18,2	-20	-17,8	-17,2	-13,3	-18,3
-9,4	-4,3	-8,9	-12,1	-19,6	-19,9	-18,5	-16,3	-14,4	-19,2
-8,2	-4,3	-9,3	-13,4	-19,4	-19,5	-18,1	-15,6	-14,6	-19,3
-7,9	-3,2	-9,7	-13,9	-19,7	-19,7	-17,8	-13,6	-14,6	-17,8

Вариант №25

9,7	5,6	3,2	4	4	12,9	16,9	12,9	3,2	3,2
10,5	4,8	3,2	3,2	5,6	12,1	16,9	13,7	3,2	4
9,7	5,6	4,8	3,2	6,4	9,7	17,7	12,1	2,4	4,8
9,7	5,6	4,8	2,4	8,1	10,5	17,7	9,7	2,4	4
8,9	4,8	5,6	1,6	8,9	16,1	16,1	10,5	1,6	4
5,6	4,8	4,8	1,6	9,7	17,7	16,1	11,3	1,6	3,2
6,4	4,8	3,2	1,6	9,7	16,9	12,9	9,7	2,4	4,8
6,4	4,8	4	3,2	13,7	16,9	15,3	7,2	3,2	5,6
5,6	4,8	4	4	12,1	16,9	13,7	5,6	3,2	5,6
5,6	3,2	4	4	12,9	16,9	12,9	3,2	3,2	4,8

2. ЛАБОРАТОРНАЯ РАБОТА №2. КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ

2.1. ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

Основные положения корреляционного анализа. Двумерная модель

Корреляционный анализ применяют, когда данные наблюдений или эксперимента считаются случайными и выбранными из нормальной совокупности.

На практике часто имеет место генеральная совокупность, которая исследуется не по одному, а по двум признакам или переменным X и Y, являющимся случайными величинами. При этом требуется установить наличие и степень связи между X и Y.

Поскольку исследовать всю генеральную совокупность по этим признакам в большинстве случаев не представляется возможным, то проводится случайная выборка.

В результате изучения выборки исследователь получает пары чисел (точки):

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_{n-1}, y_{n-1}),$$

где n – объём выборки.

Если нанести эти точки на координатную плоскость, то можно получить возможные картины распределения точек выборки (рис. 2.1).

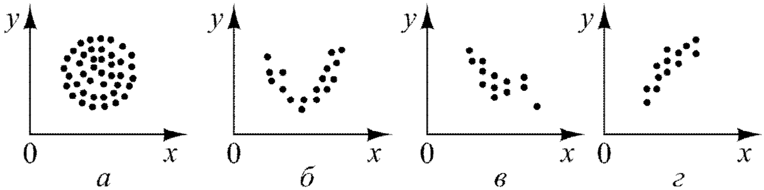

Рис. 2.1. Характер распределения точек выборки

На рис. 2.1, *а* X и Y скорее независимы: зная, какое значение приняла величина X, ничего однозначно нельзя сказать о значении Y. На остальных рисунках зависимость очевидна, причем на рис. 2.1, *б* она близка к квадратичной, а на рис. 2.1, *в* и *г* она близка к линейной. Такие величины называются коррелированными.

В случае, когда было известно двумерное распределение случайной величины (X, Y) в генеральной совокупности, можно найти генеральный коэффициент корреляции и по его величине судить о степени связи X и Y. Поскольку исследователь располагает только результатами изучения выборки, то в качестве оценки генерального коэффициента корреляции ρ может быть использован выборочный (статистический) коэффициент корреляции r.

Алгоритм нахождения выборочного коэффициента корреляции r следующий:

а) вычисляются выборочные средние переменных X, Y и XY по формулам

$$\bar{x} = \frac{1}{n}\sum_{i=0}^{n-1} x_i; \ \ \bar{y} = \frac{1}{n}\sum_{i=0}^{n-1} y_i; \ \ \overline{xy} = \frac{1}{n}\sum_{i=0}^{n-1} x_i \cdot y_i;$$

б) вычисляются выборочные средние квадратические отклонения переменных X и Y

$$S_X = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1} x_i^2 - (\bar{x})^2}, \ S_Y = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1} y_i^2 - (\bar{y})^2} \ ;$$

в) вычисляется выборочный корреляционный момент (выборочная ковариация)

$$K_{XY} = \overline{xy} - \bar{x} \cdot \bar{y} \ ;$$

г) вычисляется выборочный коэффициент корреляции

$$r = \frac{K_{XY}}{S_X \cdot S_Y}.$$

Основные свойства выборочного коэффициента корреляции при достаточно большом объеме выборки аналогичны свойствам коэффициента корреляции двух случайных величин.

1. Коэффициент корреляции принимает значения на отрезке $[-1; \ \ 1]$, т.е.

$$1 \le r \le 1.$$

В зависимости от того, насколько $|r|$ приближается к единице, различают слабую, умеренную и сильную связь: чем ближе $|r|$ к единице, тем теснее связь между признаками X и Y.

2. Если переменные X и Y умножить на одно и то же число, то коэффициент корреляции не изменится.

3. Если $r = \pm 1$, то корреляционная связь между X и Y представляет собой линейную зависимость.

Выборочный коэффициент корреляции r является оценкой генерального коэффициента корреляции ρ, тем более точной, чем больше объем выборки n.

Так как r вычисляется по данным выборки, то, в отличие от генерального коэффициента корреляции ρ, r является величиной случайной. Если $r \neq 0$, то возникает вопрос: объясняется ли это действительно существующей линейной связью между X и Y или вызвано случайными факторами?

Для выяснения значимости коэффициента корреляции проверяется гипотеза H_0 об отсутствии линейной корреляционной связи между переменными X и Y, т.е. $H_0:\ \rho = 0$.

При справедливости этой гипотезы статистика

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

имеет распределение Стьюдента с $k = n-2$ степенями свободы. Поэтому гипотеза H_0 отвергается, если

$$|t| > t_{\text{кр}},$$

где $t_{\text{кр}}$ – табличное значение распределения Стьюдента (прил. 5), определенное на уровне значимости α при числе степеней свободы $k = n-2$.

Пользуясь таблицей приложения, следует вместо n подставлять вычисленное значение k и вероятность $\gamma = 1-\alpha$.

Рассмотренный коэффициент корреляции является показателем тесноты связи лишь в случае линейной зависимости.

Разброс значений величины Y вокруг математического ожидания $M(Y)$ измеряется общей дисперсией $D(Y)$, или $S^2_{\text{общ}}$. По определению

$$S^2_{\text{общ}} = D(Y) = M(Y - M(Y))^2 .$$

Этот разброс может быть вызван:

– наличием случайных факторов, действующих на Y через переменную X (межгрупповая факторная дисперсия);

– зависимостью Y от случайных факторов (остаточных), влияющих только на Y и не влияющих на X (остаточная дисперсия).

Запишем правило сложения дисперсий в форме

$$S^2_{\text{общ}} = S^2_{\text{межгруп}} + S^2_{\text{внутригруп}}.$$

$S^2_{\text{общ}} = \dfrac{1}{n}\sum\limits_j \left(y_j - \bar{y}\right)^2 m_j$ – общая дисперсия переменной Y (дисперсия, вычисленная для всей совокупности),

$\bar{y} = \dfrac{1}{n}\sum\limits_j y_j m_j$ – среднее значение для всей совокупности.

Разделим значения y на K групп $y^{<i>}$, $i = 0,1,2,...,K-1$, тогда $\bar{y}_i = \dfrac{1}{n_i}\sum\limits_j y_j^{<i>} m_{i,j}$ – групповые средние, n_i – объём i-й группы, $m_{i,j}$ – j-е частоты для $y_j^{<i>}$ в группе i. Аналогично для групповых дисперсий: $\bar{s}_i^2 = \dfrac{1}{n_i}\sum\limits_j \left(y_j^{<i>} - \bar{y}_i\right)^2 m_{i,j}$.

Получаем: $S^2_{\text{межгруп}} = \dfrac{1}{n}\sum\limits_{i=0}^{K-1} (\bar{y}_i - \bar{y})^2 n_i$ – межгрупповая дисперсия; $S^2_{\text{внутригруп}} = \dfrac{1}{n}\sum\limits_{i=0}^{K-1} \bar{s}_i^2 \cdot n_i$ – внутригрупповая дисперсия.

Величина

$$\eta_{YX} = \frac{S_{\text{межгруп}}}{S_{\text{общ}}}$$

называется *генеральным корреляционным отношением.*

Чем теснее связь, тем большее влияние на вариацию Y оказывает изменчивость X, тем больше η_{YX}.

Величина $\bar{d} = \eta_{YX}^2$ является генеральным коэффициентом детерминации. Он показывает, какую долю дисперсии величины Y составляет дисперсия условных математических ожиданий, или, иначе говоря, какая доля дисперсии $D(Y)$ объясняется корреляционной зависимостью Y от X.

Чем ближе $\bar{d}$ к единице, тем теснее наблюдения примыкают к линии регрессии, тем лучше регрессия описывает зависимость переменных. Можно показать, что для линейной модели $\bar{d} = r^2$.

Свойства корреляционного отношения.

1. η_{YX} принимает только положительные значения и диапазоне

$$0 \le \eta_{YX} \le 1 .$$

2. Если $\eta_{YX} = 0$, то корреляционная зависимость между Y и X отсутствует.

3. Условие $\eta_{YX} = 1$ является необходимым и достаточным для функциональной зависимости величины Y от X. Это основное свойство корреляционного отношения.

Проиллюстрируем сформулированные свойства диаграммами рассеивания (рис. 2.2).

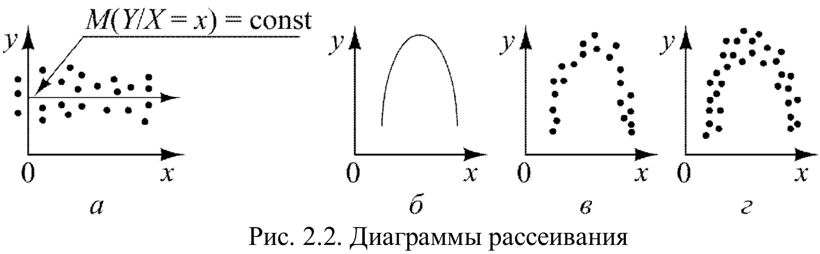

Рис. 2.2. Диаграммы рассеивания

Пояснения к рисунку: *а* – корреляционная зависимость отсутствует; *б* – зависимость Y от X функциональная: $\eta_{YX} = 1$; *в* – стохастическая зависимость довольно сильная: точки сконцентрированы в относительно узкой «параболической полосе»; *г* – стохастическая зависимость по сравнению с зависимостью, изображенной на рис. 2.2, *в*, менее сильная: точки концентрируются в более широкой «параболической полосе».

Основные положения регрессионного анализа. Парная регрессионная модель

При изучении стохастических зависимостей в задачах техники, экономики и т.д. одним из главных моментов является установление вида зависимости Y от $X_0, X_1, ..., X_{n-1}$, т.е. вида уравнения регрессии. Это связано в первую очередь с необходимостью прогнозирования исследуемых процессов.

Установление формы зависимости, оценка функции регрессии и ее параметров являются задачами *регрессионного анализа.*

В регрессионном анализе изучаются модели вида

$$Y = \varphi(X) + \varepsilon, \tag{2.1}$$

где X – неслучайная независимая переменная, называемая фактором*;* Y – случайная зависимая переменная (результирующий признак); ε – случайная переменная, характеризующая отклонение фактора X от линии регрессии (остаточная переменная).

Основными предпосылками регрессионного анализа являются следующие.

1. Для каждого из наблюдений каждая случайная переменная ε_i распределена по нормальному закону. Причем

$$\sigma^2 = D(Y_i) = D(\varepsilon_i) = \text{const}.$$

2. Для разных наблюдений переменные ε_i и ε_j некоррелированы.

Уравнение регрессии (2.1) записывается в виде

$$y_x = \varphi(x, b_0, b_1, ..., b_p), \tag{2.2}$$

где x – значения величины X;

$$y_x = M_x Y$$

$b_0, b_1, ..., b_p$ – параметры функции регрессии.

Таким образом, задача регрессионного анализа состоит в определении функции φ, ее параметров $b_0, b_1, ..., b_p$ и дальнейшем статистическом исследовании уравнения (2.2).

Рассмотрим понятие корреляционного поля. Пусть зависимость между X и Y на координатной плоскости (x, y) имеет вид, представленный на рис. 2.3. Графическое изображение такой зависимости называется корреляционным полем.

Для каждого значения $\bar{x}_i$ вычислим групповые средние $\bar{y}_i$. Соединив полученные точки ломаной линией, получим график эмпирической линии регрессии Y по X. По виду эмпирической линии регрессии (см. рис. 2.3) можно предположить, что связь между величинами Y и X является линейной.

Рассмотрим парную линейную регрессионную модель взаимосвязи двух переменных, для которой функция регрессии $\varphi(X)$ линейна относительно переменной Y. Обозначим через y_x условную среднюю признака Y в генеральной совокупности при фиксированном значении x переменной X. Тогда уравнение регрессии будет иметь вид

$$y_x = ax + b .$$

Коэффициент a в этом уравнении называется коэффициентом регрессии. Он показывает, на сколько единиц в среднем изменяется переменная Y при изменении переменной X на одну единицу.

Поскольку коэффициенты a и b для генеральной совокупности неизвестны, проводится их оценка с использованием выборки. С этой целью обычно используют метод наименьших квадратов.

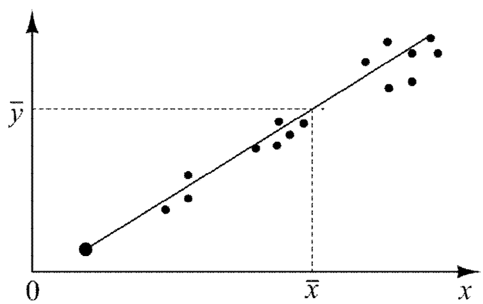

Рис. 2.3. Корреляционное поле

Как указывалось ранее, в распоряжении исследователя имеются только результаты изучения выборки в виде пар чисел (точек)

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_{n-1}, y_{n-1}),$$

где n – объём выборки.

Применяя к этим точкам метод наименьших квадратов, получаем

$$\begin{cases} a\sum_{i=0}^{n-1} x_i^2 + b\sum_{i=0}^{n-1} x_i = \sum_{i=0}^{n-1} x_i y_i \\ a\sum_{i=0}^{n-1} x_i + b \cdot n = \sum_{i=0}^{n-1} y_i \end{cases}$$

разделив оба уравнения на n и используя обозначения

$$\overline{x} = \frac{1}{n}\sum_{i=0}^{n-1} x_i; \quad \overline{x^2} = \frac{1}{n}\sum_{i=0}^{n-1} x_i^2; \quad \overline{y} = \frac{1}{n}\sum_{i=0}^{n-1} y_i; \quad \overline{xy} = \frac{1}{n}\sum_{i=0}^{n-1} x_i \cdot y_i,$$

получаем систему уравнений в виде

$$\begin{cases} a \cdot \overline{x^2} + b \cdot \overline{x} = \overline{xy} \\ a \cdot \overline{x} + b = \overline{y} \end{cases}.$$

Решаем эту систему по правилу Крамера

$$\Delta = \begin{vmatrix} \overline{x^2} & \overline{x} \\ \overline{x} & 1 \end{vmatrix} = \overline{x^2} - (\overline{x})^2 = s_X^2,$$

$$\Delta_1 = \begin{vmatrix} \overline{xy} & \overline{x} \\ \overline{y} & 1 \end{vmatrix} = \overline{xy} - \overline{x} \cdot \overline{y} = K_{XY}$$

отсюда

$$a = \frac{\Delta_1}{\Delta} = \frac{K_{XY}}{s_X^2}.$$

Свободный член b проще вычислить по формуле

$$b = \overline{y} - a \cdot \overline{x}$$

Вычислив значения a и b, найдем прямую, наилучшим образом выражающую статистическую связь между величинами X и Y. Эта прямая, определяемая уравнением регрессии $y_x = a \cdot x + b$, называется прямой регрессии Y на X.

Коэффициент регрессии можно рассчитать по другой формуле:

$$a = r\frac{S_y}{S_x}.$$

Линейное уравнение регрессии обычно записывается в принятой математической статистике форме:

$$y_x - \overline{y} = a \cdot (x - \overline{x})$$

или

$$y_x - \overline{y} = r\frac{S_y}{S_x}(x - \overline{x}).$$

Аналогично, уравнение регрессии X на Y имеет вид

$$x_y - \overline{x} = r\frac{S_x}{S_y}(y - \overline{y}).$$

Статистический анализ уравнения регрессии

Для того чтобы установить, соответствует ли выбранная регрессионная модель экспериментальным данным, используют основное уравнение дисперсионного анализа

$$Q_{\text{общ}} = Q_{\text{фактор}} + Q_{\text{ост}},$$

где $Q_{\text{общ}} = \sum_{j=0}^{n-1}(y_j - \overline{y})^2$ – общая сумма квадратов отклонений от среднего;

$Q_{\text{фактор}} = \sum_{j=0}^{n-1}(y_{x,j} - \overline{y})^2$ – сумма квадратов от среднего, обусловленная регрессией y_x;

$Q_{\text{ост}} = \sum_{j}^{n-1}(y_j - y_{x,j})^2$ – остаточная сумма квадратов.

Для заданного уровня значимости α по таблицам находим критическое значение $F_{\text{кр}}$ распределения Фишера при $k_1 = l - 1$, $k_2 = n - l$ степенях свободы, где n – число наблюдений; l – число групп в корреляционной таблице (прил. 6) или число оцениваемых параметров. Если расчетное значение критерия Фишера

(2.3) будет меньше, чем табличное значение, то нет оснований считать, что независимый фактор оказывает влияние на разброс средних значений, в противном случае, независимый фактор оказывает существенное влияние на разброс средних значений.

$$t = \frac{Q_{\text{фактор}}(n-l)}{Q_{\text{ост}}(l-1)}. \qquad (2.3)$$

Воздействие неучтенных случайных факторов в линейной модели определяется остаточной дисперсией $S^2_{\text{ост}}$. Оценкой этой дисперсии является выборочная остаточная дисперсия:

$$S^2_{\text{ост}} = \frac{1}{n-l} Q_{\text{ост}}.$$

Интервальная оценка и проверка значимости уравнения регрессии

Интервальная оценка уравнения регрессии заключается в вычислении доверительных интервалов коэффициентов a и b в уравнении регрессии с доверительной вероятностью γ. Поскольку коэффициенты a и b определены по результатам изучения выборки, то они являются лишь точечными оценками истинных коэффициентов уравнения регрессии в генеральной совокупности $a_{\text{ист}}$ и $b_{\text{ист}}$.

Доверительными интервалами коэффициентов $a_{\text{ист}}$ и $b_{\text{ист}}$ с доверительной вероятностью γ называются интервалы с границами

$$a_{\text{ист}} \in \left(a - t_\gamma s_a;\ a + t_\gamma s_a\right)$$
$$b_{\text{ист}} \in \left(b - t_\gamma s_b;\ b + t_\gamma s_b\right),$$

где s_a и s_b – средние квадратические отклонения коэффициентов a и b; t_γ – параметр распределения Стьюдента с $v = n-2$ степенями свободы (n – объём выборки). Этот коэффициент определяется в прил. 5, в котором вместо n берётся значение v.

Параметры s_a и s_b вычисляются по следующим формулам:

$$s_a = \sqrt{\frac{\sum\limits_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{(n-2)\sum\limits_{i=0}^{n-1}(x_i - \bar{x})^2}},\ s_b = \sqrt{\frac{\sum\limits_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{(n-2)n}}\ .$$

Проверить значимость уравнения регрессии – значит установить, соответствует ли математическая модель, выражающая зависимость между переменными, экспериментальным данным. Проверку значимости уравнения регрессии обычно проводят по критерию Фишера – Снедекора, который часто называют просто критерием Фишера.

Согласно этому критерию уравнение регрессии

$$y_x = ax + b$$

является значимым при уровне значимости α, если фактически наблюдаемое значение статистики критерия Фишера

$$F = \frac{s_R^2}{s^2}$$

больше, чем табличное значение F_{α,k_1,k_2}, определенное для уровня значимости α при

$$k_1 = l - 1 \text{ и } k_2 = n - l\ ,$$

где l – число оцениваемых параметров уравнения регрессии. В рассматриваемом случае их два: a и b; n – объём выборки.

Значения F_{α,k_1,k_2} приведены в прил. 6. Величины s_R^2 и s^2 определяются по формулам

$$s_R^2 = \frac{\sum\limits_{i=0}^{n-1}(y_{x,i} - y)^2}{l-1} = \frac{\sum\limits_{i=0}^{n-1}\left[(ax_i + b) - \bar{y}\right]^2}{l-1},$$

$$s^2 = \frac{\sum\limits_{i=0}^{n-1}(y_i - y_{x,i})^2}{n-2} = \frac{\sum\limits_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{n-2}\ .$$

2.2. ЗАДАНИЯ ДЛЯ ЛАБОРАТОРНОЙ РАБОТЫ №2. КОРРЕЛЯЦИОННЫЙ И РЕГРЕССИОННЫЙ АНАЛИЗ

По данным, полученным в результате выборочного наблюдения (первая строка таблицы – дни, вторая строка – Y; третья строка таблицы – X), требуется:

1. Найти уравнение линейной регрессии Y на X и X на Y.
2. Оценить тесноту связи.
3. Построить графики регрессий.
4. Найти интервальную оценку коэффициентов k и b с доверительной вероятностью $\gamma = 0{,}95$ и прверить значимость уравнения регрессии Y на X по критерию Фишера – Снедекера при уровне значимости $\alpha = 0{,}05$.

Порядок выполнения задания

1. Для получения уравнения линейной регрессии $y = ax + b$ необходимо решить систему уравнений относительно коэффициентов a и b

$$\begin{cases} a\sum_{i=0}^{n-1} x_i^2 + b\sum_{i=0}^{n-1} x_i = \sum_{i=0}^{n-1} x_i y_i \\ a\sum_{i=0}^{n-1} x_i + nb = \sum_{i=0}^{n-1} y_i \end{cases}.$$

Замечание: при отыскании уравнения регрессии X на Y в системе x и y нужно поменять местами.

2. Теснота связи оценивается с помощью выборочного коэффициента корреляции

$$r = \frac{n\sum_{i=0}^{n-1} x_i y_i - \left(\sum_{i=0}^{n-1} x_i\right)\left(\sum_{i=0}^{n-1} y_i\right)}{\sqrt{\left(n\sum_{i=0}^{n-1} x_i^2 - \left(\sum_{i=0}^{n-1} x_i\right)^2\right)\cdot\left(n\sum_{i=0}^{n-1} y_i^2 - \left(\sum_{i=0}^{n-1} y_i\right)^2\right)}}.$$

3. Построить уравнения линейной регрессии Y на X и X на Y.

4. Доверительные интервалы коэффициентов a и b с доверительной вероятностью γ имеют вид

$$a - t_\gamma s_k < a_{\text{ист}} < a + t_\gamma s_k$$

$$b - t_\gamma s_k < b_{\text{ист}} < b + t_\gamma s_k,$$

где t_γ – параметр распределения Стьюдента с $\nu = n - 2$ степенями свободы; s_k и s_b – средние квадратические отклонения коэффициентов a и b, вычисляемые соответственно по формулам

$$s_a = \sqrt{\frac{\sum_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{(n-2)\sum_{i=0}^{n-1}(x_i - \bar{x})^2}}, \quad s_b = \sqrt{\frac{\sum_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{(n-2)n}}.$$

Для проверки значимости уравнения регрессии вычислить значение критерия Фишера $F_{\text{расч}} = s_R^2 / s^2$, где

$$s_R^2 = \frac{\sum_{i=0}^{n-1}\left(y_{x,i} - \bar{y}\right)^2}{l-1} = \frac{\sum_{i=0}^{n-1}\left[(ax_i + b) - \bar{y}\right]^2}{l-1},$$

$$s^2 = \frac{\sum_{i=0}^{n-1}\left(y_i - y_{x,i}\right)^2}{n-2} = \frac{\sum_{i=0}^{n-1}\left[y_i - (ax_i + b)\right]^2}{n-2}$$

и сравнить с табличным значением $F_{\text{кр}}$ для данного уровня значимости α при $k_1 = l - 1$ и $k_2 = n - l$ (l – число оцениваемых параметров уравнения регрессии).

Если $F_{\text{расч}} < F_{\text{кр}}$, то уравнение регрессии не является значимым с уровнем значимости α. А если $F_{\text{расч}} \geq F_{\text{кр}}$, то подтверждается значимость уравнения регрессии.

2.3. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. В чем разница между функциональной, стохастической и корреляционными зависимостями?
2. Выведите линейное уравнение регрессии и сформулируйте понятие выборочного коэффициента корреляции.
3. Объясните связь между уравнением линейной регрессии и метода наименьших квадратов.
4. Объясните смысл выборочного коэффициента корреляции и его значимость.
5. Алгоритм нахождения выборочного коэффициента корреляции.
6. Приведите пример нелинейной функции регрессии.
7. Что является основными задачами регрессионого анализа?
8. Перечислите основные предпосылки регрессионного анализа.
9. Как проводится статистичекий анализ уровня регрессии?
10. Сформулируйте понятие интервальной оценки и правило проверки значимости уравнения регрессии.

2.4. ВАРИАНТЫ ЗАДАНИЙ

Вариант №1

1	2	3	4	5	6	7	8	9	10
28,3	28,4	28,5	30,5	28,5	28,6	30,6	28,8	28,8	29,3
26,1	27,2	26,3	26,2	27,3	26,5	27,7	26,9	26,9	26,0

11	12	13	14	15	16	17	18	19	20
31,8	31,1	31,4	30,1	31,2	31,9	37,3	35,8	34,3	32,3
27,4	26,8	27,0	27,6	28,6	28,9	33,5	32,2	31,0	31,2

21	22	23	24	25	26	27	28	29	30
30,9	30,4	30,2	29,9	29,4	29,0	28,9	29,7	29,3	29,5
28,5	28,1	27,9	27,8	27,5	27,0	27,3	26,7	26,6	26,4

Вариант №2

1	2	3	4	5	6	7	8	9	10
9,6	11,9	10,2	11,2	9,6	9,7	13,3	10,7	9,5	12,2
-2,6	-2,2	-2,4	-1,7	-2,1	-2,5	-2,3	-2,8	-2,9	-2,7

11	12	13	14	15	16	17	18	19	20
10,3	11,5	13,2	12,4	10,9	13,7	12,5	10,6	13,8	14,1
-1,9	-2,0	-3,0	-3,3	-3,2	-3,7	-3,9	-3,5	-3,1	-3,8

21	22	23	24	25	26	27	28	29	30
9,2	8,8	8,5	9,2	7,8	13,4	13,9	14,3	14,5	13,0
-4,3	-4,1	-4,5	-4,7	-4,9	-3,4	-3,6	-4,2	-4,6	-1,8

Вариант №3

1	2	3	4	5	6	7	8	9	10
12,6	12,5	12,9	13,1	13,7	13,5	13,7	12,8	13,1	12,9
24,0	23,8	23,6	23,4	23,3	23,2	23,1	22,8	22,6	22,3

11	12	13	14	15	16	17	18	19	20
13,1	12,9	12,9	12,7	12,5	12,7	12,8	12,7	12,7	12,9
22,1	22,0	21,9	21,8	21,8	21,7	21,7	21,7	21,7	21,6

21	22	23	24	25	26	27	28	29	30
12,9	12,9	13,0	12,9	13,0	13,0	12,9	13	12,9	12,8
21,6	21,6	21,6	21,6	21,5	21,5	21,5	21,5	21,4	21,4

Вариант №4

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

11	12	13	14	15	16	17	18	19	20
1,8	2,2	2,2	2,7	2,2	1,8	2,2	1,8	1,8	1,8
9,7	9,9	9,9	10,1	10,1	10,3	10,3	10,3	10,5	10,5

21	22	23	24	25	26	27	28	29	30
1,8	1,8	2,2	1,8	1,3	0,9	1,3	0,4	0,9	0,9
10,5	10,5	10,5	10,5	10,7	10,7	10,7	10,7	10,7	10,9

Вариант №5

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

11	12	13	14	15	16	17	18	19	20
5,5	5,6	6,0	6,1	6,2	6,4	6,9	7,0	7,2	7,1
4,5	5,4	4,9	5,4	6,3	6,7	6,7	7,2	6,7	5,8

21	22	23	24	25	26	27	28	29	30
7,1	7,6	8,0	8,2	8,3	8,6	8,7	8,8	8,7	8,7
5,8	5,4	5,8	4,0	3,6	3,1	2,7	2,2	2,7	2,7

Вариант №6

1	2	3	4	5	6	7	8	9	10
10,6	10,9	11,1	11,0	10,9	11,1	11,3	11,4	11,4	11,5
10,7	10,9	11,2	11,2	11,0	11,6	11,4	11,5	11,5	11,6

11	12	13	14	15	16	17	18	19	20
11,5	16,4	11,9	12,1	12,1	10,8	10,6	10,4	10,1	10,2
11,6	11,6	11,4	15,1	11,2	11,1	10,2	10,6	10,4	10,2

21	22	23	24	25	26	27	28	29	30
10,2	10,1	10,2	10,4	10,7	11,1	11,4	11,3	12,1	12,2
10,2	10,2	10,3	10,4	10,7	11,1	11,4	11,4	12,1	12,3

Вариант №7

1	2	3	4	5	6	7	8	9	10
3,2	3,2	2,4	2,4	2,4	2,4	3,2	2,4	2,4	3,2
6,3	5,4	5,8	5,4	4,5	5,4	4,9	5,4	4,9	5,8

11	12	13	14	15	16	17	18	19	20
3,2	3,2	4,0	4,0	4,0	4,8	4,8	4,8	4,0	4,8
5,8	6,3	7,2	7,2	7,6	8,5	7,2	8,9	8,9	7,6

21	22	23	24	25	26	27	28	29	30
4,8	4,8	4,8	4,8	4,8	4	4	3,2	3,2	2,4
7,6	9,4	8	8,5	7,6	6,7	7,2	7,2	5,8	5,8

Вариант №8

1	2	3	4	5	6	7	8	9	10
7,3	7,1	7,4	7,4	7,5	7,3	7,2	7,1	6,8	6,9
18,1	18,2	17,9	16,9	16,7	17,2	17,7	17,7	17,8	17,8

11	12	13	14	15	16	17	18	19	20
6,9	6,8	6,8	6,8	6,8	6,7	6,6	6,6	6,6	6,2
17,8	18,1	17,9	17,9	17,9	17,8	17,8	17,8	17,8	17,8

21	22	23	24	25	26	27	28	29	30
6,2	6,1	6,1	6	6	6	5,9	5,9	5,8	5,5
17,8	17,8	17,8	17,8	17,8	17,8	17,8	17,8	17,7	17,7

Вариант №9

1	2	3	4	5	6	7	8	9	10
8,1	8	7,8	7,6	7,7	7,7	7	7,3	7,1	6,6
16,8	16,8	16,7	16,7	16,6	16,6	16,6	16,6	16,6	16,4

11	12	13	14	15	16	17	18	19	20
6,6	6,2	6,3	6,3	6,1	5,9	5,9	5,9	6,2	6,3
16,4	16,4	16,3	16,3	16,3	16,3	16,3	16,4	16,6	16,7

21	22	23	24	25	26	27	28	29	30
6,3	7,6	6,2	7	6,6	6,9	8	6,2	6,9	6,9
16,7	17,8	19,1	20,5	20,7	20,9	21,2	20,9	21,2	20,6

Вариант №10

1	2	3	4	5	6	7	8	9	10
-6,8	-8,1	-6,9	-5,7	-7	-7	-6,7	-6,6	-5,4	-5,5
28,2	28,2	28,2	28,2	28,2	28,2	28,2	28,2	28,2	28,3

11	12	13	14	15	16	17	18	19	20
-5,5	-7	-7,1	-7,1	-7,3	-7,2	-6,1	-6	-7,2	-7,3
28,3	28,3	28,3	28,3	28,3	28,4	28,4	28,4	28,5	28,5

21	22	23	24	25	26	27	28	29	30
-7,3	-7,3	-5,8	-5,5	-5,8	-6,2	-5,9	-6	-5,9	-6,1
28,5	28,5	28,5	28,6	28,6	28,6	28,7	28,7	28,7	28,7

Вариант №11

1	2	3	4	5	6	7	8	9	10
15,1	14,9	14,6	14,2	14	14	14	14,2	14,4	14,9
10,3	10,3	10,4	10,2	10,2	10,2	10,2	10,4	10,6	10,7

11	12	13	14	15	16	17	18	19	20
14,9	15,7	15,9	16,2	16,8	16,8	17,1	18,5	19,6	19,7
10,7	11,2	11,3	11,6	12,1	12,1	12,4	13,1	12,1	13

21	22	23	24	25	26	27	28	29	30
19,7	20,3	21,1	22,4	23,6	24,7	25,1	25,6	25,6	25,2
13	13	13,5	14,5	14,3	13,3	13	12,4	11,7	11,7

Вариант №12

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

11	12	13	14	15	16	17	18	19	20
3,9	3,6	2,5	2,2	2,1	1,9	1,8	1,4	1	0,6
12,6	12,6	12,9	12,7	12,7	12,5	12,6	12,5	12,4	12,3

21	22	23	24	25	26	27	28	29	30
0,8	0,9	2,2	1,6	1,8	1,9	2,6	1,1	0,1	0
12,2	12,2	12,7	12,7	12,5	12,3	12,3	12,3	12,2	12,7

Вариант №13

1	2	3	4	5	6	7	8	9	10
10,9	11,2	11,4	11,6	12,4	12,7	13,1	12,9	13,2	12,7
0	0,3	0,5	0,4	1,4	1,5	1,5	1,5	1,2	2,5

11	12	13	14	15	16	17	18	19	20
12,7	11,8	12,6	12,6	12,9	12,7	12,7	12,5	12,6	12,5
2,5	3,7	3,9	3,6	2,5	2,2	2,1	1,9	1,8	1,4

21	22	23	24	25	26	27	28	29	30
12,5	5	5,2	5,5	4,8	5,1	5,5	5,6	5,5	5,3
1,4	1,1	1,1	1	1,2	1,2	2,4	1,9	1,4	1,4

Вариант №14

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

11	12	13	14	15	16	17	18	19	20
0,9	0,9	0,9	1,3	1,8	2,7	2,2	2,7	2,2	1,8
2,7	2,5	1,9	1,8	2,1	1,3	1,1	0,9	0,7	0,7

21	22	23	24	25	26	27	28	29	30
1,8	1,8	1,3	1,8	0,9	1,3	1,3	0,9	0,9	0,9
2,3	2,7	2,9	3	3,1	3,1	2,9	2,7	2,4	2

Вариант №15

1	2	3	4	5	6	7	8	9	10
18,6	19,6	20,8	21,3	21,7	21,8	22,9	22,8	23,4	23,1
18,2	20,9	21,7	22,7	22,6	23,1	22,9	21,9	20,5	18,4

11	12	13	14	15	16	17	18	19	20
14,9	15,3	15,6	16,5	17,1	17,4	18	18,7	18,3	18,9
3	5,5	6,6	8,2	7,5	8,5	9	13,3	6,5	8,3

21	22	23	24	25	26	27	28	29	30
17,5	17,9	18	17,7	18,2	18,3	18,3	18,8	18,5	18,8
17,9	19,4	20,7	22	21,9	20	21	20,9	18,2	15,4

Вариант №16

1	2	3	4	5	6	7
-12,3	-13,3	-15,4	-15,4	-15	-14,3	-12,8
1,3	2,8	5,4	5,8	7,3	10	12

8	9	10	11	12	13	14
-11,8	-11,1	-10,5	-10,5	-9,5	-8,4	-9,1
9,1	13,6	13,3	13,3	11,7	10,2	10,7

15	16	17	18	19	20	21	22
-10,1	-9,7	-9,7	-8,7	-7,4	-9,2	-18,8	-18,8
9,2	6,9	4,3	2,7	1	0,3	0,2	0,8

23	24	25	26	27	28	29	30
-17,7	-16,9	-17,4	-16,1	-15,6	-15,3	-14,4	-13,1
1,9	4,7	7,2	9,3	10,9	12,1	12,9	13

Вариант №17

1	2	3	4	5	6	7	8	9	10
-6,9	-5,5	-5,5	-5,5	-5,6	-6	-6	-5,9	-5,8	-5,8
29,3	29,3	29,3	29,3	29,2	28,6	28,6	28,7	28,8	28,9

11	12	13	14	15	16	17	18	19	20
-8,5	-8,7	-13	-7,7	-7,2	-6,9	-6,8	-6,8	-6,8	-6,9
24,6	27,1	25,4	28,2	28,9	29,3	29,4	29,4	29,4	29,3

21	22	23	24	25	26	27	28	29	30
-6,8	-6,5	-9,3	-9,4	-9,8	-8	-9,9	-9,6	-9,7	-9,5
29,4	29,8	20	17,4	19,9	27,8	25,3	25,9	27	24,6

Вариант №18

1	2	3	4	5	6	7	8	9	10
9	9	8	11	12	9	6	7	6	7
29,4	29,8	20	17,4	19,9	27,8	25,3	25,9	27	24,6

11	12	13	14	15	16	17	18	19	20
-3,1	-3,7	-3,2	-3,4	-4,1	-4,3	-4,3	-4,5	-4,5	-5,2
25,5	26	26,1	26,2	26,3	26,3	26,4	26,5	26,5	26,6

21	22	23	24	25	26	27	28	29	30
5,2	4,3	4,1	4,6	4,4	5,1	6,8	7	5,4	6,7
19,3	19,2	19,2	19,1	19,1	19,2	19,4	19,6	19,6	19,8

Вариант № 19

1	2	3	4	5	6	7	8	9	10
6,6	5,8	5,5	6	5,6	5,7	6,5	6,2	6,3	7,1
53	53	53	52	52	52	51	51	51	50

11	12	13	14	15	16	17	18	19	20
7,1	7,1	7,4	7,9	7,8	8,2	8,6	8,1	9,2	8,7
50	49	46	47	48	48	49	49	49	47

21	22	23	24	25	26	27	28	29	30
8,7	9,1	7,8	7,8	8,1	7,8	7,4	7,3	7,1	6,6
47	49	50	51	51	51	52	52	53	53

Вариант №20

1	2	3	4	5	6	7	8	9	10
6,8	6,6	6,6	6,7	6,8	6,8	6,8	7	7	7,1
10,1	9,8	10,1	9,8	9,8	9,8	9,8	9,8	9,8	9,8

11	12	13	14	15	16	17	18	19	20
7,1	7,1	7	7	6,8	6,8	6,8	6,7	6,7	6,7
9,8	9,8	9,8	9,8	9,8	9,8	9,8	9,7	9,7	9,7

21	22	23	24	25	26	27	28	29	30
6,7	6,7	6,7	6,7	6,6	6,6	6,5	6,5	6,5	6,6
9,7	9,7	9,7	9,7	9,7	9,7	9,7	9,6	9,6	9,6

Вариант №21

1	2	3	4	5	6	7	8	9	10
7,3	7,3	7,3	7,3	7,2	7,2	7,1	7,1	7,1	7,1
5,9	5,9	5,9	5,3	5,4	5,9	5,3	4,8	4,7	5,3

11	12	13	14	15	16	17	18	19	20
7,1	7,1	7,2	7,1	7,1	7,1	6,9	6,9	6,9	6,8
5,3	5,3	5,3	4,8	5,3	5,3	4,6	4,2	4,1	4,4

21	22	23	24	25	26	27	28	29	30
6,8	6,7	6,6	6,6	6,4	6,3	6,4	6,4	6,4	6,3
4,4	3,8	3,8	3,7	3,4	3,3	2,9	3,4	3,4	3,2

Вариант №22

1	2	3	4	5	6	7	8	9	10
6,9	6,8	6,7	6,6	6,6	6,4	6,4	6,5	6,5	6,4
5,8	5,2	5,2	5,1	4,8	4,8	4,5	4,9	4,9	4,7

11	12	13	14	15	16	17	18	19	20
6,4	6,3	6,3	6,2	6,3	6,4	6,3	6,3	6,1	6,1
4,7	4,7	4,9	4,4	5,1	5,6	5	4,1	4,2	4,3

21	22	23	24	25	26	27	28	29	30
6,1	6,2	6,3	7,2	7,4	8,4	8,6	8,7	8,3	7,7
4,3	4,4	5	6,1	6,4	7,8	7,4	7,1	6,3	6,2

Вариант №23

1	2	3	4	5	6	7	8	9	10
5,9	6,1	6,1	6,3	7,1	7,1	7,9	8,3	7,7	7
8,5	8,6	8,6	8,2	8,3	8,4	8,4	8	7,9	7,8

11	12	13	14	15	16	17	18	19	20
7	6,4	6	5,6	5,4	5,4	5,7	5,8	5,5	5,2
7,8	7,7	7,6	7,6	7,5	7,4	7,4	7,4	7,4	7,3

21	22	23	24	25	26	27	28	29	30
5,2	4,9	4,7	4,4	4,1	3,9	3,8	3,7	3,6	3,4
7,3	7,3	7,3	7,3	7,3	7,2	6,8	6,7	6,7	6,7

Вариант №24

1	2	3	4	5	6	7	8	9	10
5,2	4,9	4,7	4,4	4,1	3,9	3,8	3,7	3,6	3,4
23,1	23,1	23,1	23,1	23,1	22,9	22,9	22,7	22,7	22,7

11	12	13	14	15	16	17	18	19	20
3,4	3,3	3,2	2,9	2,7	2,6	2,8	2,8	2,6	2,3
22,7	22,7	22,7	22,7	22,7	22,7	22,7	22,7	22,7	22,7

21	22	23	24	25	26	27	28	29	30
2,3	2,8	3,1	4,5	5,9	6,6	7,1	8,1	8,4	10,1
22,7	22,8	23,3	24,2	24,8	23,4	22,3	20,9	20,9	21,2

Вариант №25

1	2	3	4	5	6	7
-10,6	-11	-11,2	-11,2	-10,3	-9,9	-10
-4,3	-10,2	-11,4	-12,4	-12,3	-12,9	-12

8	9	10	11	12	13	14
-9,4	-9,5	-9,7	-9,7	-9,5	-9,5	-9,9
-12,8	-8,5	-8	-8	-5,6	-3,7	-1,8

15	16	17	18	19	20	21	22
-10	-9,4	-9,4	-9,5	-9,7	-10,6	-10,6	-10,3
-1,3	-4,4	-10,7	-14,9	-16,3	-17,1	-17,1	-15,4

23	24	25	26	27	28	29	30
-10	-10,5	-10,4	-10,4	-9,8	-9,8	-10,5	-10
-15,9	-16,3	-16,6	-17	-13	-11,4	-10,4	-9,9

3. ЛАБОРАТОРНАЯ РАБОТА №3. АППРОКСИМАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

3.1. ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

Основная задача аппроксимации – построение приближенной (аппроксимирующей) функции $y = \varphi(x)$ наиболее близко проходящей около экспериментально полученных точек (x_i, y_i), $i = 0,...,n-1$ или заданной непрерывной функции $y = f(x)$. В общем смысле термином «аппроксимация» называют замену одних математических объектов другими, например, сложных функций более простой функцией или набором простых функций, в том или ином смысле близкими к исходным. Для количественной характеристики близости вводится соответствующая мера расстояния, определяемая через разности $\varepsilon_i = f(x_i) - \varphi(x_i)$. При этом весьма желательно, чтобы аппроксимация была как можно более точной и вычислялась за малое время.

В простейшем случае задача аппроксимации экспериментальных данных выглядит следующим образом: через заданные точки на плоскости (x, y) или вблизи этих точек провести кривую $y = \varphi(x)$, которая наилучшим образом воспроизводила бы график исходной экспериментальной закономерности, но в то же время была бы нечувствительна к случайным отклонениям измеряемой величины (рис. 3.1).

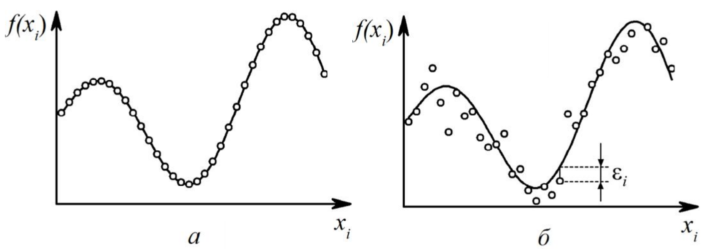

Рис. 3.1. Иллюстрация задачи аппроксимации: *а* - аппроксимирующая кривая проходит через заданные точки; *б* - приближенное описание экспериментальной зависимости (точки не ложатся в точности на кривую)

Принято рассматривать эту задачу в двух несколько различных широко распространенных постановках. В первом случае зависимость между величинами x и y задается заранее известной функцией $y = \varphi(x, a_0, a_1, ..., a_m)$ с точностью до значений параметров a_k, $(k = 0,1,...,m)$. Задача состоит в том, чтобы как можно точнее оценить эти параметры. Важно отметить, что параметры аппроксимации не могут быть непосредственно найдены из эксперимента.

Во второй постановке требуется подобрать функцию $y = \varphi(x, a_0, a_1, ..., a_m)$ соответствующую наилучшему приближению к фактической экспериментальной зависимости $y = f(x_i), i = 0,...,n-1$. В этом случае непрерывную функцию $\varphi(x)$ для аппроксимации дискретной зависимости $y = f(x_i)$ выбирают из условия наилучшего квадратичного приближения (стандартный метод наименьших квадратов)

$$S(a_0, a_1, ..., a_m) = \sum_{i=0}^{n-1} \left(y_i - \varphi(x_i, a_0, a_1, ..., a_m)\right)^2 \to \min. \qquad (3.1)$$

Задачу на поиск минимума $S(a_0, a_1, ..., a_m)$ можно свести к задаче поиска корня системы уравнений с неизвестными параметрами a_k, $(k = 0,1,...,m)$, $S(a_0, a_1, ..., a_m)$

$$\frac{\partial S}{\partial a_k} = 0, \;\; k = 0,1,...m. \qquad (3.2)$$

Подстановка (3.1) в (3.2) приведет к системе $(m+1)$ алгебраических уравнений (3.3) относительно неизвестных параметров.

Решая эту систему уравнений, найдем значения параметров $a_0, a_1, ..., a_m$ и, следовательно, искомую аппроксимирующую функцию $\varphi(x, a_0, a_1, ..., a_m)$, удовлетворяющую сформулированным выше требованиям.

$$\begin{cases} \sum\limits_{i=0}^{n-1}\left(y_i-\varphi(x_i,a_0,a_1,\ldots,a_m)\right)\dfrac{\partial}{\partial a_0}\varphi(x_i,a_0,a_1,\ldots,a_m)=0; \\ \sum\limits_{i=0}^{n-1}\left(y_i-\varphi(x_i,a_0,a_1,\ldots,a_m)\right)\dfrac{\partial}{\partial a_1}\varphi(x_i,a_0,a_1,\ldots,a_m)=0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum\limits_{i=0}^{n-1}\left(y_i-\varphi(x_i,a_0,a_1,\ldots,a_m)\right)\dfrac{\partial}{\partial a_m}\varphi(x_i,a_0,a_1,\ldots,a_m)=0. \end{cases} \tag{3.3}$$

Для аппроксимации могут использоваться любые функции, но задача значительно упрощается, когда параметры $a_k(k=0,1,...,m)$ в аппроксимирующую функцию входят линейно

$$\varphi(x,a_0,a_1,...,a_m)=a_0\varphi_0(x)+a_1\varphi_1(x)+...+a_m\varphi_m(x),\ m<N\ , \tag{3.4}$$

где $\varphi_k(x)$ произвольные базисные функции. Общих рекомендаций по выбору базисных функций нет. Выбор класса и задание свойств этих функций должны осуществляться с учетом свойств исследуемых экспериментальных данных. Иначе аппроксимация будет трудоемка и осуществлена недостаточно эффективно. Как правило, разные базисные функции имеют свои преимущества и недостатки, но наилучшей считается аппроксимирующая функция, в которой используется небольшое число базисных функций.

Если $\sqrt{S}>\varepsilon$, то число коэффициентов аппроксимации недостаточно для правильного воспроизведения графика экспериментальной зависимости, а в случае $\sqrt{S}<\varepsilon$ значительно увеличиваются ошибка оценки и время вычисления параметров, а также усложняется их интерпретация (физический смысл).

Подстановка (3.4) в (3.1), а затем вычисление производных (3.2) приведет в итоге к следующей системе линейных алгебраических уравнений:

$$\begin{cases} \sum\limits_{i=0}^{n-1}\left(a_0\varphi_0(x_i)+a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-f(x_i)\right)\cdot\varphi_0(x_i)=0; \\ \sum\limits_{i=0}^{n-1}\left(a_0\varphi_0(x_i)+a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-f(x_i)\right)\cdot\varphi_1(x_i)=0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum\limits_{i=0}^{n-1}\left(a_0\varphi_0(x_i)+a_1\varphi_1(x_i)+\ldots+a_m\varphi_m(x_i)-f(x_i)\right)\cdot\varphi_m(x_i)=0. \end{cases} \tag{3.5}$$

Для нахождения неизвестных параметров $a_0, a_1, \ldots, a_m$ обычно составляется расширенная матрица, которую называют матрицей Грамма. Элементами этой матрицы являются скалярные произведения базисных функций и столбец свободных коэффициентов:

$$G(\varphi_0,\varphi_1,\ldots,\varphi_m)=\begin{pmatrix} (\varphi_0,\varphi_0) & (\varphi_0,\varphi_1) & \ldots & (\varphi_0,\varphi_m) & (\varphi_0,f) \\ (\varphi_1,\varphi_0) & (\varphi_1,\varphi_1) & \ldots & (\varphi_1,\varphi_m) & (\varphi_1,f) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (\varphi_m,\varphi_0) & (\varphi_m,\varphi_1) & \ldots & (\varphi_m,\varphi_m) & (\varphi_m,f) \end{pmatrix},$$

где

$$(\varphi_j,\varphi_m)=\sum_{i=0}^{n-1}\varphi_j(x_i)\varphi_k(x_i),\quad (\varphi_j,f)=\sum_{i=0}^{n-1}\varphi_j(x_i)f(x_i),\ j,k=0,1,\ldots,m\,.$$

Основное преимущество рассмотрения матрицы Грамма состоит в том, что для нахождения решения системы (3.5) необходимо сосчитать только элементы первой строки и двух последних столбцов: остальные элементы заполняются сдвигом предшествующей строки (за исключением двух последних столбцов) на одну позицию влево. А в случае ортогональности базисных функций $\varphi_k(x)$

$$\int_a^b \varphi_j(x)\varphi_m(x)dx=0,\quad j\neq m,$$

все недиагональные элементы матрицы Грамма равны нулю, что увеличивает точность вычисления параметров. Хорошо известными системами ортогональных базисных функций являются полиномы Лежандра тригонометрические ряды Фурье, вейвлет-функции.

Рассмотрим некоторые примеры аппроксимирующих функций:

1. *Полиномиальная аппроксимация.*

Аппроксимирующая функция представляет собой полином степени m вида

$$\varphi(x, a_0, a_1, \ldots, a_m) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m, \quad m < n,$$

здесь базисные функции представлены в виде последовательности степеней аргумента x:

$$\varphi_0(x) = x^0 = 1, \varphi_1(x) = x^1 = 1, \ldots, \varphi_m = x^m.$$

Система уравнений линейных алгебраических уравнений для параметров выглядит следующим образом

$$\begin{cases} a_0 n + a_1 \sum\limits_{i=0}^{n-1} x_i + a_2 \sum\limits_{i=0}^{n-1} x_i^2 + \cdots + a_m \sum\limits_{i=0}^{n-1} x_i^m = \sum\limits_{i=0}^{n-1} y_i; \\ a_0 \sum\limits_{i=0}^{n-1} x_i + a_1 \sum\limits_{i=0}^{n-1} x_i^2 + a_2 \sum\limits_{i=0}^{n-1} x_i^3 + \cdots + a_m \sum\limits_{i=0}^{n-1} x_i^{m+1} = \sum\limits_{i=0}^{n-1} x_i y_i; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_0 \sum\limits_{i=0}^{n-1} x_i^m + a_1 \sum\limits_{i=0}^{n-1} x_i^{m+1} + a_2 \sum\limits_{i=0}^{n-1} x_i^{m+2} + \cdots + a_m \sum\limits_{i=0}^{n-1} x_i^{2m} = \sum\limits_{i=0}^{n-1} x_i^m y_i. \end{cases}$$

Расширенная матрица Грамма для степенного базиса имеет вид:

$$G(1, x, \ldots, x^m) = \begin{pmatrix} n & \sum\limits_{i=0}^{n-1} x_i & \ldots & \sum\limits_{i=0}^{n-1} x_i^m & \sum\limits_{i=0}^{n-1} y_i \\ \sum\limits_{i=0}^{n-1} x_i & \sum\limits_{i=0}^{n-1} x_i^2 & \ldots & \sum\limits_{i=0}^{n-1} x_i^{m+1} & \sum\limits_{i=0}^{n-1} x_i y_i \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sum\limits_{i=0}^{n-1} x_i^m & \sum\limits_{i=0}^{n-1} x_i^{m+1} & \ldots & \sum\limits_{i=0}^{n-1} x_i^{2m} & \sum\limits_{i=0}^{n-1} x_i^m y_i \end{pmatrix}.$$

Вычисляя все входящие в матрицу суммы по экспериментальным данным, параметры $a_k, k = 0,1,\ldots,m$ могут быть найдены, например, методом Гаусса.

В частности, если аппроксимирующая функция является линейной

$$\varphi(x,a_0)=a_0+a_1x$$

с базисными функциями $\varphi_0(x)=1$, $\varphi_1(x)=x$, для которых скалярные произведения соответственно равны

$$(\varphi_0,\varphi_0)=1;\ (\varphi_0,\varphi_1)=\sum_{i=0}^{n-1}x_i;\ (\varphi_1,\varphi_0)=\sum_{i=0}^{n-1}x_i;$$

$$(\varphi_1,\varphi_1)=\sum_{i=1}^{n}x_i^2;\ (\varphi_0,f)=\sum_{i=0}^{n-1}y_i;\ (\varphi_1,f)=\sum_{i=0}^{n-1}x_iy_i,$$

расширенная матрица примет вид

$$G(1,x)=\begin{pmatrix} n & \sum\limits_{i=0}^{n-1}x_i & \sum\limits_{i=0}^{n-1}y_i \\ \sum\limits_{i=0}^{n-1}x_i & \sum\limits_{i=0}^{n-1}x_i^2 & \sum\limits_{i=0}^{n-1}x_iy_i \end{pmatrix}.$$

Преобразуя матрицу $G(1,x)$ к диагональному виду, получим следующие параметры линейной аппроксимации:

$$a_0=\frac{\sum\limits_{i=0}^{n-1}x_i^2\sum\limits_{i=0}^{n-1}y_i-\sum\limits_{i=0}^{n-1}x_i\sum\limits_{i=0}^{n-1}x_iy_i}{n\sum\limits_{i=0}^{n-1}x_i^2-\left(\sum\limits_{i=0}^{n-1}x_i\right)^2};\ a_1=\frac{n\sum\limits_{i=0}^{n-1}x_iy_i-\sum\limits_{i=1}^{n-1}x_i\sum\limits_{i=0}^{n-1}y_i}{n\sum\limits_{i=0}^{n-1}x_i^2-\left(\sum\limits_{i=0}^{n-1}x_i\right)^2}.$$

Аналогично для параболической аппроксимации

$$\varphi(x,a_0,a_1,a_2)=a_0+a_1x+a_2x^2$$

расширенную матрицу можно записать в виде

$$G(1,x,x^2)=\begin{pmatrix} n & \sum\limits_{i=0}^{n-1}x_i & \sum\limits_{i=0}^{n-1}x_i^2 & \sum\limits_{i=0}^{n-1}y_i \\ \sum\limits_{i=0}^{n-1}x_i & \sum\limits_{i=0}^{n-1}x_i^2 & \sum\limits_{i=0}^{n-1}x_i^3 & \sum\limits_{i=0}^{n-1}x_iy_i \\ \sum\limits_{i=0}^{n-1}x_i^2 & \sum\limits_{i=0}^{n-1}x_i^3 & \sum\limits_{i=0}^{n-1}x_i^4 & \sum\limits_{i=0}^{n-1}x_i^2y_i \end{pmatrix}.$$

Вычисляя из экспериментальных данных необходимые суммы и диагонализируя расширенную матрицу, находим параметры аппроксимации a_0,a_1,a_2.

Коэффициенты $a_0, a_1, \ldots, a_m$ вычислены путем математической обработки величин x_i, y_i, полученных с погрешностью ε. Поэтому аппроксимирующая функция также будет получена с погрешностью.

Подставив найденные коэффициенты $a_0, a_1, \ldots, a_m$ в соотношение (3.1), найдем значение суммы S – квадратов отклонений табличных значений y_i относительно полученной аппроксимирующей функции $\varphi(x, a_0, a_1, \ldots, a_m)$.

Степень аппроксимирующего полинома m выбирают так, чтобы обеспечить по возможности точное приближение аппроксимируемой $y = f(x)$ и аппроксимирующей $\varphi(x)$ функций на всем интервале $[x_0, x_{n-1}]$. При обработке экспериментальных данных, определенных с погрешностью $\varepsilon_{\text{заданое}}$, когда аналитическое выражение для $f(x)$ неизвестно, обычно начинают с аппроксимации полиномом степени $m = 1$. После определения коэффициентов $a_0, a_1, \ldots, a_m$ вычисляют величину ε. Если получится, что $\varepsilon_{\text{заданое}} > \varepsilon$, то необходимо увеличить степень полинома m на единицу и повторить вычисления. Увеличение m и повторные вычисления необходимо осуществлять до тех пор, пока не выполнится условие $\varepsilon_{\text{заданое}} \le \varepsilon$.

Для контроля целесообразно на одном чертеже построить исходные точки x_i, y_i и график аппроксимирующего полинома. При правильном выборе значения m график полинома усредняет, сглаживает случайные погрешности ε_i, на всем интервале $[x_0, x_{n-1}]$.

Многие нелинейные задачи аппроксимации могут быть сведены к линейным путем замены переменных и значений функции. При этом данные таблицы пересчитывают и представляют в виде удобном для линейной аппроксимации. Рассмотрим несколько примеров.

2. *Степенная аппроксимация.* Аппроксимирующая функция представляет собой нелинейную функцию параметров a и b

$$\varphi(x, a, b) = ax^b.$$

В этом случае, логарифмируя аппроксимирующую функцию, задачу можно свести к линейной аппроксимации

$$\ln|\varphi(x,a,b)| = \ln|a| + b\ln|x|.$$

Покажем, как получить для степенной функции формулы из табл. 3.1.

Сумма квадратов отклонений имеет вид:

$$S = \sum_{i=0}^{n-1}(\ln a + b\ln x_i - \ln y_i)^2 .$$

Частные производные по основаниям b и $\ln a$:

$$\frac{\partial S}{\partial b} = 2\sum_{i=0}^{n-1}(\ln a + b\ln x_i - \ln y_i)\ln x_i ,$$

$$\frac{\partial S}{\partial(\ln a)} = 2\sum_{i=0}^{n-1}(\ln a + b\ln x_i - \ln y_i) .$$

Приравняв частные производные нулю, имеем

$$\begin{cases} b\sum_{i=0}^{n-1}(\ln x_i)^2 + \ln a\sum_{i=0}^{n-1}\ln x_i = \sum_{i=0}^{n-1}(\ln x_i \ln y_i) \\ b\sum_{i=0}^{n-1}\ln x_i + n\ln a = \sum_{i=0}^{n-1}\ln y_i \end{cases} .$$

Можно поступить другим способом. Введем замену переменных $\tilde{x} = \ln x$, а также новый параметр $\tilde{a} = \ln a$. Тогда

$$\ln\varphi(x,a,b) = \tilde{a} + b\tilde{x} .$$

Таким образом, задача сведена к линейной. Для практической реализации решения задачи аппроксимации необходимо:

- прологарифмировать данные исходной таблицы;
- по новой таблице, решая линейную задачу аппроксимации, найти параметры $\tilde{a}, b$;
- вычислить $a = e^{\tilde{a}}$.

3. *Экспоненциальная аппроксимация* $\varphi(x,a,b) = ae^{bx}$.

Прологарифмируем исходную функцию $\varphi(x,a,b)$. Тогда

$$\ln\varphi(x,a,b) = \ln a + bx .$$

Обозначим: $\tilde{a} = \ln a$. В результате получим

$$\ln \varphi(x,a,b) = \tilde{a} + bx \ .$$

Задача сведена к линейной. Практически для решения данной задачи аппроксимации необходимо выполнить следующие шаги:

- прологарифмировать значения функции исходной таблицы;
- по новой таблице найти параметры $\tilde{a}, b$;
- вычислить $a = e^{\tilde{a}}$.

Покажем, как получить для данной функции формулы из табл. 3.1.

Сумма квадратов отклонений имеет вид

$$S = \sum_{i=0}^{n-1} (\ln a + bx_i - \ln y_i)^2 \ .$$

Частные производные по основаниям b и $\ln a$:

$$\frac{\partial S}{\partial b} = 2\sum_{i=0}^{n-1} (\ln a + bx_i - \ln y_i)x_i \ ,$$

$$\frac{\partial S}{\partial (\ln a)} = 2\sum_{i=0}^{n-1} (\ln a + bx_i - \ln y_i) \ .$$

Приравняв частные производные нулю, имеем

$$\begin{cases} b\sum_{i=0}^{n-1} x_i^2 + \ln a \sum_{i=0}^{n-1} x_i = \sum_{i=0}^{n-1} \left(x_i \ln y_i\right) \\ b\sum_{i=0}^{n-1} x_i + n\ln a = \sum_{i=0}^{n-1} \ln y_i \end{cases} .$$

В результате решения нормальной системы можно найти b и $\ln a$, а по таблице антилогарифмов найти a.

4. *Гиперболическая аппроксимация* $\varphi(x,a,b) = \dfrac{1}{ax+b}$.

Введем новую функцию $\dfrac{1}{y} = ax + b$. Задача сведена к линейной. Фактически необходимо построить новую таблицу, в которой значения аргумента остаются неизменными, а значения функции заменены обратными к начальным и, используя новую таблицу, найти параметры a, b.

5. *Гиперболическая аппроксимация* $\varphi(x,a,b)=\dfrac{x}{ax+b}$.

Рассмотрим обратную функцию

$$\frac{1}{y}=\frac{ax+b}{x}=x+\frac{b}{x}$$

Введем замену переменных $\widetilde{x}=1\backslash x$. Тогда

$$\frac{1}{y}=a+b\widetilde{x}.$$

Таким образом, для сведения данной нелинейной задачи к линейной достаточно заменить данные исходной таблицы обратными величинами.

6. *Гиперболическая аппроксимация* $\varphi(x,a,b)=\dfrac{a}{x}+b$.

Сумма квадратов отклонений имеет вид

$$S=\sum_{i=0}^{n-1}\left(\frac{a}{x_i}+b-y_i\right)^2.$$

Частные производные по a и b равны

$$\frac{\partial S}{\partial a}=2\sum_{i=0}^{n-1}\left(\frac{a}{x_i}+b-y_i\right)\frac{1}{x_i},$$

$$\frac{\partial S}{\partial b}=2\sum_{i=0}^{n-1}\left(\frac{a}{x_i}+b-y_i\right).$$

Приравнивая к нулю, имеем

$$\begin{cases} bn+a\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i}=\sum\limits_{i=0}^{n-1}y_i \\ b\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i}+a\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i^2}=\sum\limits_{i=1}^{n-1}\dfrac{y_i}{x_i} \end{cases}.$$

Если нелинейную зависимость аппроксимирующей функции от параметров нельзя представить в виде суммы (3.4), то задачу минимизации решают, используя численные итерационные методы (градиентный спуск, методы Ньютона и др.)

В табл. 3.1 приведены системы нормальных уравнений для некоторых исходных функций.

Таблица 3.1. Системы нормальных уравнений

Исходное уравнение	Система нормальных уравнений
$y = ax^b$	$\begin{cases} b\sum_{i=0}^{n-1}(\ln x_i)^2 + a\sum_{i=0}^{n-1}\ln x_i - \sum_{i=0}^{n-1}(\ln x_i \ln y_i) = 0 \\ b\sum_{i=0}^{n-1}\ln x_i + na - \sum_{i=0}^{n-1}\ln y_i = 0 \end{cases}$
$y = a\ln x + b$	$\begin{cases} a\sum_{i=0}^{n-1}(\ln x_i)^2 + b\sum_{i=0}^{n-1}\ln x_i - \sum_{i=0}^{n-1}(y_i \ln x_i) = 0 \\ a\sum_{i=0}^{n-1}\ln x_i + bn - \sum_{i=0}^{n-1} y_i = 0 \end{cases}$
$y = e^{ax+b}$	$\begin{cases} a\sum_{i=0}^{n-1}x_i^2 + b\sum_{i=0}^{n-1}x_i - \sum_{i=0}^{n-1}(x_i \ln y_i) = 0 \\ a\sum_{i=0}^{n-1}x_i + nb - \sum_{i=0}^{n-1}\ln y_i = 0 \end{cases}$
$y = ae^{bx}$	$\begin{cases} b\sum_{i=0}^{n-1}x_i^2 + \ln a\sum_{i=0}^{n-1}x_i - \sum_{i=0}^{n-1}(x_i \ln y_i) = 0 \\ b\sum_{i=0}^{n-1}x_i + n\ln a - \sum_{i=0}^{n-1}\ln y_i = 0 \end{cases}$
$y = \frac{a}{x} + b$	$\begin{cases} a\sum_{i=0}^{n-1}\frac{1}{x_i^2} + b\sum_{i=0}^{n-1}\frac{1}{x_i} - \sum_{i=0}^{n-1}\frac{y_i}{x_i} = 0 \\ a\sum_{i=0}^{n-1}\frac{1}{x_i} + bn - \sum_{i=0}^{n-1} y_i = 0 \end{cases}$

Окончание табл. 3.1.

$y=\dfrac{1}{ax+b}$	$\begin{cases} a\sum\limits_{i=0}^{n-1}x_i^2+b\sum\limits_{i=0}^{n-1}x_i-\sum\limits_{i=0}^{n-1}\dfrac{x_i}{y_i}=0 \\ a\sum\limits_{i=0}^{n-1}x_i+bn-\sum\limits_{i=0}^{n-1}\dfrac{1}{y_i}=0 \end{cases}$
$y=\dfrac{x}{ax+b}$	$\begin{cases} a\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i^2}+b\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i}-\sum\limits_{i=0}^{n-1}\dfrac{1}{x_iy_i}=0 \\ a\sum\limits_{i=0}^{n-1}\dfrac{1}{x_i}+bn-\sum\limits_{i=0}^{n-1}\dfrac{1}{y_i}=0 \end{cases}$

3.2. ЗАДАНИЯ ДЛЯ ЛАБОРАТОРНОЙ РАБОТЫ № 3. АППРОКСИМАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

1. Для таблично заданной функции найти параметры следующих аппроксимирующих функций: линейной, степенной, полиномиальной и экспоненциальной.
2. Оценить величину достоверности для каждого случая.
3. Построить на одном рисунке исходные данные и графики аппроксимирующих функций.
4. Построить нормальность распределения остатков наилучшей аппроксимации.

Порядок выполнения задания

1. Ввод исходных данных. Исходными данными является таблица значений узлов аппроксимации и соответствующих им значений аппроксимируемой функции. В первой строке таблицы находятся номера узлов, во второй строке значения x_j, в третьей значения y_j. Количество узлов равно $n=20$ (если аппроксимирующая функция $\varphi(x)$ нелинейна относительно коэффициентов аппроксимации, то предварительно следует линеаризовать ее путем подходящей замены переменных).

2. Расчет значений аппроксимирующей функции и среднего квадратического отклонения.

Расчет значений аппроксимирующей функции состоит из двух этапов:

- решение системы уравнений;
- вычисление значений полинома в узловых точках x_i.

3. Пересчитать коэффициенты, если это необходимо для перехода к исходной аппроксимирующей функции $\varphi(x)$, и напечатать их.

4. Вывести в графическом виде построенную функцию $\varphi(x)$.

5. Определить остатки наилучшей аппроксимации и проверить их нормальность в соответствии с заданиями № 5, 6 к лабораторной работе № 1.

3.3. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Постановка задачи аппроксимации функций.
2. Чем отличается интерполяция от аппроксимации функций?
3. Что является исходными данными к задаче аппроксимации функций?
4. Постановка задачи аппроксимации функций методом наименьших квадратов.
5. Сущность аппроксимации МНК.
6. Когда применяется метод наименьших квадратов?
7. Определение коэффициента достоверности аппроксимации.
8. Какие действия необходимо выполнить на диаграмме для построения линии аппроксимации?
9. Каким образом задается тип аппроксимации?
10. Что необходимо сделать, чтобы вывести уравнение и величину достоверности аппроксимации?

3.4. ВАРИАНТЫ ЗАДАНИЙ

Вариант №1

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	3,9
0,2	0,3	0,3	0,3	0,4	0,5	0,4	0,7	0,8	1,0

Вариант №2

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	3,9
2,9	4,1	4,9	4,6	6,2	6,9	5,1	7,9	7,6	8,3

Вариант №3

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,7	5,0
6,2	6,2	7,6	8,4	8,0	9,2	8,9	10,4	10,5	13,0

Вариант №4

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,3	3,5	3,7	4,0	4,3	4,5	4,7	5,0
-5,4	-5,5	-6,0	-6,8	-7,2	-7,5	-8,5	-8,5	-9,0	-9,3

Вариант №5

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,7	5,0
3,1	3,2	4,6	5,4	5,0	6,3	5,9	7,4	7,5	9,9

Вариант №6

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,2	3,5	3,7	4,0	4,3	4,5	4,8	5,0
1,7	1,9	2,1	2,0	2,4	2,5	2,1	2,7	2,5	2,7

Вариант №7

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,7	5,0
4,4	4,5	6,0	7,0	6,9	8,4	8,3	10,4	11,1	14,3

Вариант №8

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,2	3,5	3,7	4,0	4,3	4,5	4,8	5,0
1,8	2,1	2,2	2,1	2,4	2,5	2,1	2,6	2,5	2,6

Вариант №9

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	4,0
2,8	3,5	3,9	3,8	4,7	5,1	4,3	5,9	5,9	6,4

Вариант №10

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,7	5,0
11,4	12,2	17,1	20,4	20,1	24,7	24,3	30,2	31,5	40,0

Вариант №11

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

11	12	13	14	15	16	17	18	19	20
1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	1,9
2,7	3,3	3,7	3,8	4,5	5,3	4,8	6,2	6,6	6,6

Вариант №12

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	3,9
0,8	0,9	1,0	1,0	1,1	1,2	1,1	1,4	1,5	1,6

Вариант №13

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,2	3,5	3,7	4,0	4,3	4,5	4,8	4,9
0,4	0,5	0,5	0,4	0,5	0,4	0,2	0,4	0,3	0,3

Вариант №14

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,2	3,4	3,8	4,0	4,2	4,5	4,8	5,0
0,2	0,3	0,2	0,2	0,2	0,5	0,3	0,1	0,4	0,3

Вариант №15

1	2	3	4	5	6	7	8	9	10
0,2	0,4	0,6	0,8	1,1	1,2	1,4	1,6	1,8	2,0
6,4	3,2	7,4	5,8	7,1	6,1	9,8	6,3	9,1	12,9

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,3	3,6	3,8	4,0
15,3	15,9	21,9	25,1	24,2	29,4	26,9	34,5	34,6	44,8

Вариант №16

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	4,0
1,7	2,0	2,2	2,2	2,6	2,7	2,3	3,0	3,0	3,2

Вариант №17

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,3	3,5	3,7	4,0	4,3	4,5	4,7	5,0
3,2	2,8	3,3	3,6	3,4	3,8	4,0	4,2	4,3	4,5

Вариант №18

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

11	12	13	14	15	16	17	18	19	20
2,8	3,0	3,3	3,5	3,7	4,0	4,3	4,5	4,7	5,0
1,7	2,2	2,5	2,2	2,4	3,1	3,7	4,3	3,9	4,7

Вариант №19

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,7	5,0
5,1	5,2	7,3	8,6	8,0	9,8	9,2	11,5	11,6	15,1

Вариант №20

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,3	3,6	3,8	4,0
3,5	3,2	5,5	6,6	5,2	7,0	5,5	8,0	7,4	11,4

Вариант №21

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,7	4,0	4,2	4,5	4,7	5,0
4,5	6,6	7,3	8,6	8,7	7,4	9,1	10,1	11,8	14,0

Вариант №22

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

11	12	13	14	15	16	17	18	19	20
2,7	3,0	3,3	3,5	3,8	4,0	4,2	4,5	4,8	5,0
7,0	9,1	8,1	9,3	10,5	12,6	12,1	13,5	13,6	13,5

Вариант №23

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	3,9
5,1	7,0	8,4	8,4	10,9	12,5	10,2	14,8	14,9	16,0

Вариант №24

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

11	12	13	14	15	16	17	18	19	20
2,2	2,4	2,6	2,8	3,0	3,2	3,4	3,6	3,8	3,9
2,2	3,1	3,7	3,5	4,7	5,3	3,9	6,1	5,9	6,4

Вариант №25

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

11	12	13	14	15	16	17	18	19	20
2,1	2,4	2,6	2,7	3,0	3,2	3,4	3,6	3,8	4,0
0,2	0,3	0,2	0,2	0,2	0,4	0,2	0,1	0,3	0,2

4. ЛАБОРАТОРНАЯ РАБОТА №4. ОЦЕНИВАНИЕ ХАРАКТЕРИСТИК СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

4.1. ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

На практике часто приходится иметь дело со случайными величинами, непрерывно изменяющимися в процессе опыта. В отличие от обычных случайных величин будем называть их случайными функциями.

Известно, что за последнее время в технике все большее распространение получают системы с автоматизированным управлением. Соответственно все большие требования предъявляются к теоретической базе этого вида техники – к теории автоматического управления. Развитие этой теории невозможно без анализа ошибок, неизбежно сопровождающих процессы управления, которые всегда протекают в условиях непрерывно воздействующих случайных возмущений (так называемых «помех»). Эти возмущения по своей природе являются случайными функциями. Для того чтобы рационально выбрать конструктивные параметры системы управления, необходимо изучить ее реакцию на непрерывно воздействующие случайные возмущения, а единственным аппаратом, пригодным для такого исследования, является аппарат теории случайных функций.

Первым из основных понятий, с которыми нам придется иметь дело, является само понятие случайной функции:

Случайной функцией называется функция, которая в результате опыта может принять тот или иной конкретный вид и неизвестно заранее – какой именно.

Конкретный вид, принимаемый случайной функцией в результате опыта, называется *реализацией случайной функции*. Если над случайной функцией произвести группу опытов, то мы получим группу или «семейство» реализаций этой функции.

Примерами случайных функций в энергетике, связанными с метеорологическими условиями, могут быть: изменение располагаемой мощности и энергии гидростанций, зависящие от приточности рек; изменения суммарного спроса мощности и энергии в энергосистемах, зависящие как от изменения температуры наружного воздуха, так и от других факторов; отклонение напряжений в узлах нагрузок, зависящие от различных потребителей, так и от режима системы и времени.

Будем рассматривать только случайные функции одного аргумента. Так как этим аргументом чаще всего является время, будем обозначать его буквой t. Кроме того, условимся, как правило, обозначать случайные функции большими буквами $(X(t), Y(t), ...)$ в отличие от неслучайных функций.

Характеристики случайных функций

Большое значение в теории вероятностей имеют основные числовые характеристики случайных величин, математическое ожидание и дисперсия – для одной случайной величины, математические ожидания и корреляционная матрица – для системы случайных величин.

Для случайных функций также вводятся простейшие основные характеристики, аналогичные числовым характеристикам случайных величин, и устанавливаются правила действий с этими характеристиками.

В отличие от числовых характеристик случайных величин, представляющих собой определенные *числа*, характеристики случайных функций представляют собой в общем случае не числа, а *функции*.

Математическое ожидание случайной функции $X(t)$ определяется следующим образом. Рассмотрим сечение случайной функции $X(t)$ при фиксированном t. В этом сечении мы имеем обычную случайную величину; определим ее математическое ожидание. Очевидно, в общем случае оно зависит от t, т.е. представляет собой некоторую функцию t:

$$m_x(t) = M[X(t)].$$

Таким образом, *математическим ожиданием случайной функции* $X(t)$ *называется неслучайная функция* $m_x(t)$, *которая при каждом значении аргумента* t *равна математическому ожиданию соответствующего сечения случайной функции*.

По смыслу математическое ожидание случайной функции есть некоторая *средняя функция*, около которой различным образом варьируются конкретные реализации случайной функции.

Аналогичным образом определяется дисперсия случайной функции. *Дисперсией случайной функции* $X(t)$ *называется неслучайная функция* $D_x(t)$, *значение которой для каждого* t *равно дисперсии соответствующего сечения случайной функции*:

$$D_x(t) = D[X(t)].$$

Дисперсия случайной функции при каждом t характеризует разброс возможных реализаций случайной функции относительно среднего, т.е. «степень случайности» случайной функции.

Очевидно, что $D_x(t)$ есть неотрицательная функция. Извлекая из нее квадратный корень, получим функцию $\sigma_x(t)$ – среднее квадратическое отклонение случайной функции:

$$\sigma_x(t) = \sqrt{D_x(t)}\ .$$

Математическое ожидание и дисперсия представляют собой весьма важные характеристики случайной функции; однако для описания основных особенностей случайной функции этих характеристик недостаточно. У случайных функций $X_1(t)$ и $X_2(t)$ могут быть примерно одинаковые математические ожидания и дисперсии; однако характер этих случайных функций резко различен. Очевидно, внутренняя структура обоих случайных процессов при этом окажется совершенно различной, но это различие не улавливается ни математическим ожиданием, ни дисперсией; для его описания необходимо ввести специальную характеристику. Эта характеристика называется *корреляционной функцией* (иначе – *автокорреляционной* функцией). Корреляционная функция характеризует степень зависимости между сечениями случайной функции, относящимися к различным t.

Пусть имеется случайная функция $X(t)$. Рассмотрим два ее сечения, относящихся к различным моментам: t и t', т.е. две случайные величины $X(t)$ и $X(t')$. Очевидно, что при близких значениях t и t' величины $X(t)$ и $X(t')$ связаны тесной зависимостью: если величина $X(t)$ приняла какое-то значение, то и величина $X(t')$ с большой вероятностью примет значение, близкое к нему (см. рис. 4.1). Очевидно также, что при увеличении интервала между сечениями t и t_1' зависимость величин $X(t)$ и $X(t'_1)$ вообще должна убывать.

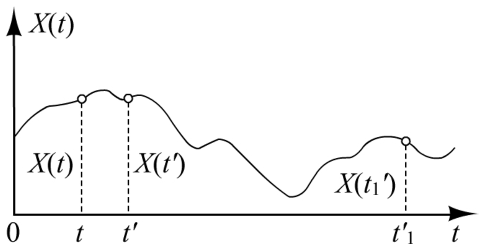

Рис. 4.1. Три сечения t, t' и t_1' случайной функции $X(t)$

Степень зависимости величин $X(t)$ и $X(t')$ может быть в значительной мере охарактеризована их *корреляционным моментом*.

Очевидно, он является функцией двух аргументов t и t'. Эта функция и называется корреляционной функцией.

Таким образом, *корреляционной функцией случайной функции* $X(t)$ *называется неслучайная функция двух аргументов* $K_x(t,t')$, *которая при каждой паре значений* t, t' *равна корреляционному моменту соответствующих сечений случайной функции*:

$$K_x(t,t') = M\left[\overset{\circ}{X}(t)\,\overset{\circ}{X}(t')\right],$$

где $\overset{\circ}{X}(t) = X(t) - m_x(t)$, $\overset{\circ}{X}(t') = X(t') - m_x(t')$.

Две случайные функции $X_1(t)$ и $X_2(t)$ при одинаковых математических ожиданиях и дисперсиях могут иметь совершенно различные корреляционные функции. Полагая $t' = t$, имеем

$$K_x(t,t') = M\left[(\overset{\circ}{X}(t))^2\right] = D_x(t),$$

т.е. *при* $t' = t$ *корреляционная функция обращается в дисперсию случайной функции*.

Таким образом, необходимость в дисперсии как отдельной характеристике случайной функции отпадает: в качестве основных характеристик случайной функции достаточно рассматривать ее математическое ожидание и корреляционную функцию.

Так как корреляционный момент двух случайных величин $X(t)$ и $X(t')$ не зависит от последовательности, в которой эти величины рассматриваются, то корреляционная функция симметрична относительно своих аргументов, т.е. не меняется при перемене аргументов местами:

$$K_x(t,t') = K_x(t',t).$$

Вместо корреляционной функции $K_x(t,t')$ можно пользоваться нормированной корреляционной функцией:

$$r_x(t,t') = \frac{K_x(t,t')}{\sigma_x(t)\sigma_x(t')},$$

которая представляет собой коэффициент корреляции величин $X(t)$, $X(t')$.

При $t' = t$ нормированная корреляционная функция равна единице.

$$r_x(t, t') = \frac{K_x(t, t')}{[\sigma_x(t)]^2} = \frac{D_x(t)}{[\sigma_x(t)]^2} = 1 .$$

Стационарный случайный процесс

На практике очень часто встречаются случайные процессы, протекающие во времени приблизительно однородно и имеющие вид непрерывных случайных колебаний вокруг некоторого среднего значения, причем ни средняя амплитуда, ни характер этих колебаний не обнаруживают существенных изменений с течением времени. Такие случайные процессы называются стационарными.

В качестве примеров стационарных случайных процессов можно привести следующие: 1) колебания самолета на установившемся режиме горизонтального полета; 2) колебания напряжения в электрической осветительной сети; 3) случайные шумы в радиоприемнике; 4) процесс качки корабля и т.д.

Каждый стационарный процесс можно рассматривать как продолжающийся во времени неопределенно долго; при исследовании стационарного процесса в качестве начала отсчета можно выбрать любой момент времени. Исследуя стационарный процесс на любом участке времени мы должны получить одни и те же его характеристики. Образно выражаясь, стационарный процесс «не имеет ни начала, ни конца».

Примером стационарного случайного процесса может служить изменение высоты центра тяжести самолета на установившемся режиме горизонтального полета (рис. 4.2).

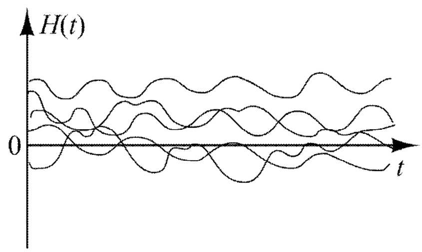

Рис. 4.2. Пример реализации стационарного случайного процесса

Стационарные случайные процессы очень часто встречаются в физических и технических задачах.

Случайная функция $X(t)$ называется *стационарной*, если все ее вероятностные характеристики не зависят от t (точнее, не меняются при любом сдвиге аргументов, от которых они зависят, по оси t).

Так как изменение стационарной случайной функции должно протекать однородно по времени, то естественно потребовать, чтобы для стационарной случайной функции математическое ожидание было постоянным:

$$m_x(t) = m_x = \text{const} . \tag{4.1}$$

Заметим, однако, что это требование не является существенным: мы знаем, что от случайной функции $X(t)$ всегда можно перейти к центрированной случайной функции $\overset{\circ}{X}(t)$, для которой математическое ожидание тождественно равно нулю и, следовательно, удовлетворяет условию (4.1). Таким образом, если случайный процесс нестационарен только за счет переменного математического ожидания, это не помешает нам изучать его как стационарный процесс.

Второе условие, которому должна удовлетворять стационарная случайная функция, это условие постоянства дисперсии:

$$D_x(t) = D_x = \text{const} . \tag{4.2}$$

Установим, какому условию должна удовлетворять корреляционная функция стационарной случайной функции. Рассмотрим случайную функцию $X(t)$ (рис. 4.3). Положим в выражении $K_x(t, t')$ $t' = t + \tau$ и рассмотрим $K_x(t,\ t + \tau)$ – корреляционный момент двух сечений случайной функции, разделенных интервалом времени τ. Очевидно, если случайный процесс $X(t)$ действительно стационарен, то этот корреляционный момент не должен зависеть от того, где именно на оси $0t$ мы взяли участок τ, а должен зависеть только от длины этого участка. Например, для участков I и II на рис. 4.3, имеющих одну и ту же длину τ, значения корреляционной функции $K_x(t, t + \tau)$ и $K_x(t_1, t_1 + \tau)$ должны быть одинаковыми.

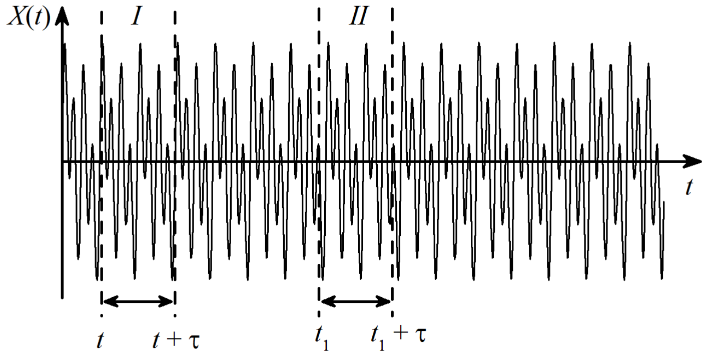

Рис. 4.3. Выделенные участки *I* и *II*
для случайной функции $X(t)$

Вообще, корреляционная функция стационарного случайного процесса должна зависеть не от положения t первого аргумента на оси абсцисс, а только от промежутка τ между первым и вторым аргументами:

$$K_x(t, t+\tau) = k_x(\tau). \tag{4.3}$$

Следовательно, корреляционная функция стационарного случайного процесса есть функция не двух, а всего одного аргумента. Это обстоятельство в ряде случаев сильно упрощает операции над стационарными случайными функциями.

Заметим, что условие (4.2), требующее от стационарной случайной функции постоянства дисперсии, является частным случаем условия (4.3). Действительно, полагая в формуле (4.3) $t+\tau = t$ $(\tau = 0)$, имеем $D_x(t) = K_x(t, t) = k_x(0) = \text{const}$.

Таким образом, условие (4.3) есть единственное существенное условие, которому должна удовлетворять стационарная случайная функция.

Поэтому в дальнейшем под стационарной случайной функцией мы будем понимать такую случайную функцию, корреляционная функция которой зависит не от обоих своих аргументов t и t', а только от разности τ между ними. Чтобы не накладывать специальных условий на математическое ожидание, мы будем рассматривать только центрированные случайные функции.

Мы знаем, что корреляционная функция любой случайной функции обладает свойством симметрии:

$$K_x(t,t') = K_x(t',t).$$

Отсюда для стационарного процесса, полагая $t' - t = \tau$, имеем:

$$k_x(\tau) = k(-\tau),$$

т.е. корреляционная функция $k_x(\tau)$ есть четная функция своего аргумента. Поэтому обычно корреляционную функцию определяют только для положительных значений аргумента (рис. 4.4).

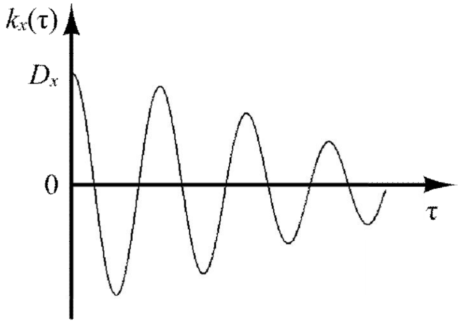

Рис. 4.4. График корреляционной функции

На практике вместо корреляционной функции $k_x(\tau)$ часто пользуются *нормированной корреляционной функцией*

$$\rho_x(\tau) = \frac{k_x(\tau)}{D_x},$$

где $D_x = k_x(0)$ – постоянная дисперсия стационарного процесса. Функция $\rho_x(\tau)$ есть не что иное, как коэффициент кореляции между сечениями случайной функции, разделенными интервалом τ по времени. Очевидно, что $\rho_x(0) = 1$.

Эргодическое свойство стационарных случайных функций

Рассмотрим некоторую стационарную случайную функцию $X(t)$, для которой требуется оценить ее характеристики: математическое ожидание m_x и корреляционную функцию $k_x(\tau)$. Для этого нужно располагать известным числом реализаций случайной функции $X(t)$. Обрабатывая эти реализации, можно найти оценки для математического ожидания $\widetilde{m}_x(t)$ и корреляционной функции $\widetilde{K}_x(t,t')$.

В связи с ограниченностью числа наблюдений функция $\widetilde{m}_x(t)$ не будет строго постоянной; ее придется осреднить и заменить некоторым постоянным $\widetilde{m}_x$; аналогично, осредняя значения $\widetilde{K}_x(t,t')$ для разных $\tau = t' - t$, получим корреляционную функцию $\widetilde{k}_x(\tau)$.

Этот метод обработки является довольно сложным и громоздким. Возникает вопрос: нельзя ли для стационарной случайной функции этот сложный процесс обработки заменить более простым, который базируется на предположении, что математическое ожидание не зависит от времени, а корреляционная функция – от начала отсчета?

Поскольку случайный процесс является стационарным и протекает однородно по времени, естественно предположить, что единственная реализация достаточной продолжительности может служить достаточным опытным материалом для получения характеристик случайной функции.

При более подробном рассмотрении этого вопроса оказывается, что такая возможность существует не для всех случайных процессов; не всегда одна реализация достаточной продолжительности оказывается эквивалентной множеству отдельных реализаций. Для примера рассмотрим две случайные стационарные функции $X_1(t)$ и $X_2(t)$, представленные совокупностью своих реализаций на рис. 4.5 и 4.6.

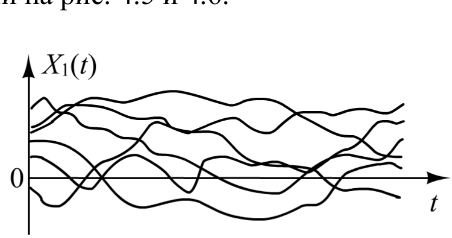

Рис. 4.5. Реализации случайной стационарной функции $X_1(t)$

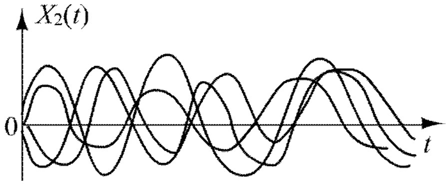

Рис. 4.6. Реализации случайной стационарной функции $X_2(t)$, которая обладает эргодическим свойством

Каждая реализация случайной функции $X_2(t)$ обладает одними и теми же характерными признаками: средним значением, вокруг которого происходят колебания, и средним размахом этих колебаний. Выберем произвольно одну из таких реализаций и продолжим мысленно опыт, в результате которого она получена, на некоторый участок времени T. Очевидно, при достаточно большом T эта одна реализация сможет дать нам достаточно хорошее представление о свойствах случайной функции в целом. В частности, осредняя значения этой реализации вдоль оси абсцисс – по времени, мы должны получить приближенное значение математического ожидания случайной функции; осредняя квадраты отклонений от этого среднего, мы должны получить приближенное значение дисперсии и т.д.

Про такую случайную функцию говорят, что она обладает *эргодическим свойством*. Эргодическое свойство состоит в том, что каждая отдельная реализация случайной функции является как бы «полномочным представителем» всей совокупности возможных реализаций; одна реализация достаточной продолжительности может заменить при обработке множество реализаций той же общей продолжительности.

Теперь рассмотрим случайную функцию $X_1(t)$. Выберем произвольно одну из ее реализаций, продолжим ее мысленно на достаточно большой участок времени и вычислим ее среднее значение по времени на всем участке наблюдения. Очевидно, это среднее значение для каждой реализации будет свое и может существенно отличаться от математического ожидания случайной функции, построенного как среднее из множества реализаций. Про такую случайную функцию говорят, что она не обладает эргодическим свойством.

Если случайная функция $X(t)$ обладает эргодическим свойством, то для нее *среднее по времени* (*на достаточно большом участке наблюдения*) *приближенно равно среднему по множеству наблюдений*. То же будет верно и для $X^2(t)$, $X(t)$, $X(t+\tau)$ и т.д. Следовательно, все характеристики случайной функции (математическое ожидание, дисперсию, корреляционную функцию) можно будет приближенно определять по одной достаточно длинной реализации.

Определение характеристик эргодической стационарной случайной функции по одной реализации

Рассмотрим стационарную случайную функцию $X(t)$, обладающую эргодическим свойством, и предположим, что в нашем распоряжении имеется всего одна реализация этой случайной функции, но на достаточно большом участке времени *T*. Для эргодической стационарной случайной функции одна реализация достаточно большой продолжительности практически эквивалентна (в смысле объема сведений о случайной функции) множеству реализаций той же общей продолжительности; характеристики случайной функции могут быть приближенно определены не как средние по множеству наблюдений, а *как средние по времени* t. В частности, при достаточно большом T математическое ожидание m_x может быть приближенно вычислено по формуле

$$m_x \approx \frac{1}{T}\int_0^T x(t)dt. \tag{4.4}$$

Аналогично может быть приближенно найдена корреляционная функция $k_x(\tau)$ при любом τ. Действительно, корреляционная функция, по определению, представляет собой не что иное, как *математическое ожидание* случайной функции $\overset{\circ}{X}(t)\overset{\circ}{X}(t+\tau)$:

$$k_x(\tau) = M\left[\overset{\circ}{X}(t)\overset{\circ}{X}(t+\tau)\right].$$

Это математическое ожидание также, очевидно, может быть приближенно вычислено как среднее по времени.

Фиксируем некоторое значение τ и вычислим указанным способом корреляционную функцию $k_x(\tau)$. Для этого удобно предварительно «центрировать» данную реализацию $x(t)$, т.е. вычесть из нее математическое ожидание (4.5):

$$\overset{\circ}{x}(t) = x(t) - m_x . \qquad (4.5)$$

Вычислим при заданном τ математическое ожидание случайной функции $\overset{\circ}{X}(t)\overset{\circ}{X}(t+\tau)$ как среднее по времени. При этом, очевидно, нам придется учитывать не весь участок времени от 0 до T, а несколько меньший, так как второй сомножитель $\overset{\circ}{X}(t+\tau)$ известен нам не для всех t, а только для тех, для которых $t+\tau \le T$. Вычисляя среднее по времени указанным выше способом, получим:

$$k(\tau) \approx \frac{1}{T-\tau}\int_0^{T-\tau} \overset{\circ}{x}(t)\overset{\circ}{x}(t+\tau)dt . \qquad (4.6)$$

Вычислив интеграл (4.6) для ряда значений τ, можно приближенно воспроизвести по точкам весь ход корреляционной функции. На практике обычно интегралы (4.4) и (4.6) заменяют конечными суммами. Покажем, как это делается. Разобьем интервал записи случайной функции на n равных частей длиной Δt и обозначим середины полученных участков $t_0, t_1, t_2, ..., t_{n-1}$ (рис. 4.7).

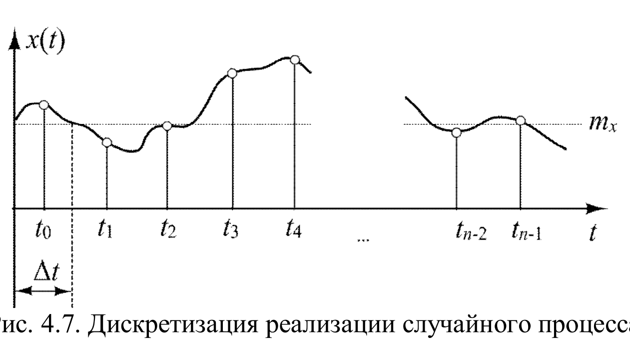

Рис. 4.7. Дискретизация реализации случайного процесса

Представим интеграл (4.4) как сумму интегралов по элементарным участкам Δt и на каждом из них вынесем функцию $x(t)$ из-под знака интеграла средним значением, соответствующим центру интервала $x(t_i)$. Получим приближенно:

$$m_x = \frac{1}{T} \cdot \frac{T}{n} \sum_{i=0}^{n-1} x(t_i),$$

или

$$m_x = \frac{1}{n} \sum_{i=0}^{n-1} x(t_i) .$$

Аналогично можно вычислить корреляционную функцию для значений τ, равных $0, \Delta t, 2\Delta t, \ldots$. Придадим, например, величине τ значение

$$\tau = m\Delta t = \frac{mT}{n}$$

и вычислим интеграл (4.6), деля интервал интегрирования

$$T - \tau = T - \frac{mT}{n} = \frac{n-m}{n} T$$

на $n-m$ равных участков длиной Δt и вынося на каждом из них функцию $\overset{\circ}{x}(t)\overset{\circ}{x}(t+\tau)$ за знак интеграла средним значением.

Получим:

$$k_x\left(\frac{mT}{n}\right) = \frac{n}{(n-m)T} \cdot \frac{T}{n} \sum_{i=0}^{n-m-1} \overset{\circ}{x}(t_i)\overset{\circ}{x}(t_{i+m}),$$

или окончательно

$$k_x\left(\frac{mT}{n}\right) = \frac{1}{(n-m)} \sum_{i=0}^{n-m-1} \overset{\circ}{x}(t_i)\overset{\circ}{x}(t_{i+m}) . \qquad (4.7)$$

Вычисление корреляционной функции по формуле (4.7) производят для $m = 0, 1, 2, \ldots$ последовательно вплоть до таких значений m, при которых корреляционная функция становится практически равной нулю или начинает совершать небольшие нерегулярные колебания около нуля. Общий ход функции $k_x(\tau)$ воспроизводится по отдельным точкам (рис. 4.8).

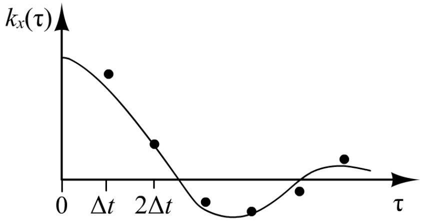

Рис. 4.8. Общий ход корреляционной функции

Для того чтобы математическое ожидание m_x и корреляционная функция $k_x(\tau)$ были определены с удовлетворительной точностью нужно, чтобы число точек n было достаточно велико (порядка сотни, а в некоторых случаях даже нескольких сотен). Выбор длины элементарного участка Δt определяется характером изменения случайной функции. Если случайная функция изменяется сравнительно плавно, участок Δt можно выбирать большим, чем когда она совершает резкие и частые колебания. Чем более высокочастотный состав имеют колебания, образующие случайную функцию, тем чаще нужно располагать опорные точки при обработке. Ориентировочно можно рекомендовать выбирать элементарный участок Δt так, чтобы на полный период самой высокочастотной гармоники в составе случайной функции приходилось порядка 5–10 опорных точек.

Часто выбор опорных точек вообще не зависит от обрабатывающего, а диктуется темпом работы записывающей аппаратуры. В этом случае следует вести обработку непосредственно полученного из опыта материала, не пытаясь вставить между наблюденными значениями промежуточные, так как это не может повысить точности результата, а излишне осложнит обработку.

Подготовка данных

При решении многих прикладных задач необходимо знать статистические характеристики случайных процессов. Обычно эти характеристики определяют по реализациям случайного процесса, полученным экспериментально. Наиболее просто эта задача решается для стационарного случайного процесса, статистические характеристики которого не изменяются во времени.

Если стационарный случайный процесс обладает эргодическим свойством, то его статистические характеристики могут быть определены по одной реализации, наблюдаемой на достаточно большом интервале времени.

Целью работы является определение оценок математического ожидания, дисперсии, плотности распределения вероятности и корреляционной функции стационарного случайного процесса по его заданной реализации. Пусть дана реализация $x_p(t)$ случайного процесса длительностью T_p (рис. 4.9).

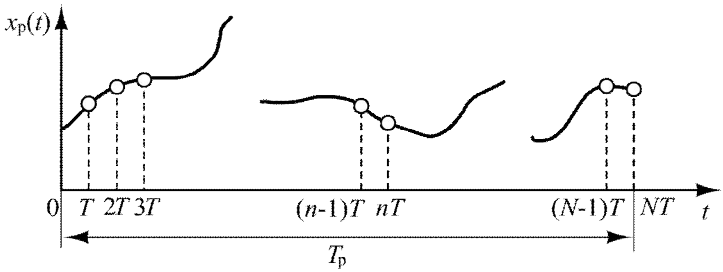

Рис. 4.9. Дискретизация реализации случайного процесса

На основании эргодического свойства математическое ожидание, дисперсия и корреляционная функция этого стационарного случайного процесса определяются соответственно выражениями:

$$m_x = \lim_{T_p \to \infty} \frac{1}{T_p} \int_0^{T_p} x_p(t) dt \qquad (4.8)$$

$$D_x = \lim_{T_p \to \infty} \frac{1}{T_p} \int_0^{T_p} \left[x_p(t) - m_x\right]^2 dt \qquad (4.9)$$

$$R_x(\tau) = \lim_{T_p \to \infty} \frac{1}{T_p} \int_0^{T_p} \left[x_p(t) - m_x\right] \cdot \left[x_p(t+\tau) - m_x\right] dt. \qquad (4.10)$$

Оценки приведенных выше статистических характеристик находятся по имеющейся реализации $x_p(t)$ заданной длины по следующим формулам (4.11)-(4.13), полученным из (4.8)-(4.10).

$$m_x^* = \frac{1}{T_p}\int_0^{T_p} x_p(t)dt \qquad (4.11)$$

$$D_x^* = \frac{1}{T_p}\int_0^{T_p} \left[x_p(t) - m_x^*\right]^2 dt \qquad (4.12)$$

$$R_x^*(\tau) = \frac{1}{T_p - \tau}\int_0^{T_p-\tau} \left[x_p(t) - m_x^*\right]\cdot\left[x_p(t+\tau) - m_x^*\right]dt\,. \qquad (4.13)$$

Расчет численных значений оценок математического ожидания m_x^*, дисперсии D_x^* и ординат функции $R_x^*(\tau)$, являющейся оценкой корреляционной функции, производится путем дискретизации реализации $x_p(t)$ случайного процесса и заменой интегралов в выражениях (4.11)-(4.13) суммами. Для этого интервал наблюдения $[0,T_p]$ случайного процесса (рис. 4.9) разбивается на N достаточно малых интервалов длительностью T. Вначале каждого из этих интервалов определяются значения

$$x(0) = x_p(0),\ x(1) = x_p(T),\ \ldots,\ x(N-1) = x_p((N-1)T).$$

При этом возникает задача выбора необходимой длины T_p реализации случайного процесса и длительности T интервала дискретизации. Необходимая длина реализации в основном определяется требуемой точностью оценивания корреляционной функции и ее свойствами. Разработаны методики, которые позволяют оценить необходимую длину реализации для получения корреляционной функции с заданной точностью непосредственно по отрезку реализации случайного процесса. Например, для определения T_p используется связь среднего числа максимумов и нулей случайного процесса в единицу времени с параметрами корреляционной функции. Шаг дискретизации T рекомендуют выбирать так, чтобы на «полупериод» (время между двумя пересечениями графиком случайной функции линии математического ожидания) реализации случайного процесса приходилось около семи дискретных значений.

Оценки математического ожидания и дисперсии

Для оценки математического ожидания, заменив в (4.11) интеграл суммой, получим $m_x^* = \frac{1}{N \cdot T}\sum_{n=0}^{N-1} x(n)\cdot T = \frac{1}{N}\sum_{n=0}^{N-1} x(n)$.

Таким образом, оценка математического ожидания равна среднему арифметическому значений $x(0), x(1), x(N-1)$ реализации случайного процесса в дискретные моменты времени.

Для оценки дисперсии из формулы (4.12) найдем

$$D_x^* = \frac{1}{N}\sum_{n=0}^{N-1}\left[x(n) - m_x^*\right]^2 .$$

Однако эта оценка дисперсии оказывается смещенной. На практике применяется формула, которая удовлетворяет условию несмещенности: $D_x^* = \frac{1}{N-1}\sum_{n=0}^{N-1}\left[x(n) - m_x^*\right]^2$.

Оценка одномерной плотности распределения

Оценка одномерной плотности распределения вероятности случайного процесса сводится к построению гистограммы распределения значений $x(0), x(1), x(N-1)$ реализации. Для этого следует разбить полученный диапазон значений случайных чисел на равные интервалы и вычислить частоты попадания случайных чисел в эти интервалы.

Большое значение при построении гистограммы имеет разбиение интервала изменения случайных чисел на интервалы группировки. При слишком большом числе интервалов группировки некоторые из них оказываются слабозаполненными и гистограмма получается изрезанной и многолепестковой. При малом числе интервалов гистограмма утрачивает детальность и становится малоинформативной. В работе рекомендуется принять $L \approx \sqrt{N}$ (L – целое).

Оценка корреляционной функции

По формуле (4.13) для значений оценки корреляционной функции в дискретные моменты времени $\tau_m = mT$ $(m = 0,1,...)$ получим следующее выражение

$$R_x^*(m) = \frac{1}{N-m} \sum_{n=0}^{N-m-1} \left[x(n) - m_x^*\right] \cdot \left[x(n+m) - m_x^*\right], \qquad (4.14)$$

где $m = 0,1,...,M-1$. Показано, что чем больше m, тем больше ошибка определения корреляционной функции. Поэтому формулу (4.14) рекомендуют использовать при $M < N/10$. Найденную по формуле (4.14) оценку корреляционной функции необходимо дополнить симметричными отсчетами для отрицательных $m = -M+1,...,-1$.

4.2. ЗАДАНИЯ ДЛЯ ЛАБОРАТОРНОЙ РАБОТЫ № 4. ОЦЕНИВАНИЕ ХАРАКТЕРИСТИК СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

В качестве объекта исследования используется последовательность случайных чисел $x(0), x(2),..., x(N-1)$, моделирующая некоторый стационарный случайный процесс. Для получения последовательности используется программа Matlab. Структура программы состоит из трёх частей: объявление входных параметров; генерирование дискретного массива с помощью функции Вейерштрасса и визуализация полученных данных в виде графика (смотрите текст программы ниже).

```
K = 4; s = 2.3; lambda=2; T = 1; N = 1000; % 1) входные параметры.
n = (0:(N-1))'; t = n*T/N; x = zeros( N,1 );  % 2) получение массива.
for n = 1:N
   for k=0:K
      x(n) = x(n) + (lambda^(s-2)*k)*sin(t(n)*2*pi*(lambda^(k)));
   end
end
figure,plot(t,x);xlabel('x');ylabel('y'); % 3) визуализация
```

Текст программы в среде Matlab R2012b для получения последовательности чисел

Входными параметрами являются: «K» – количество гармоник, которые будут моделировать процесс, «s» и «lambda» – величины меняющие форму сигнала, «N» – количество отсчётов в сигнале, «Т» – длительность времени, в течение которого существует процесс.

Получите гистограммы для определения частот попадания случайных чисел в заданные интервалы. Для этого сформируйте вектор w с составляющими $w(0), w(1), w(2),...., w(L-1)$, представляющими собой границы интервалов группировки. Эти составляющие образуются по формуле

$$w(l+1) = w(l) + \Delta w, \quad l = 0,1,2,...,L-1, \tag{4.15}$$

где

$$\Delta w = \frac{\max(x) - \min(x)}{L}, \quad w(0) = \min(x). \tag{4.16}$$

Для расчета оценки корреляционной функции используются приведенные выше формулы.

Порядок выполнения задания

1. Сформировать последовательность случайных чисел $x(n), n = 0,1,...,N-1$, $N \in [500, 1000]$ согласно вариантам задания.

2. Разбить интервал изменения значений случайных чисел в последовательности $x(n), n = 0,1,...,N-1$ на L интервалов группировки, определив последовательность $w(0), w(1), w(2), w(L-1)$ по формулам (4.15)-(4.16) и построить гистограмму. Сделать вывод о характере распределения случайной последовательности.

3. Определить значение оценки m_x^* математического ожидания случайной последовательности.

4. Определить значение оценки D_x^* дисперсии случайной последовательности.

5. Рассчитать значения ординат оценки $R_x^*(m)$ корреляционной функции для $m = 0,1,...,32$. Построить график оценки корреляционной функции.

4.3. КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Что понимают под стационарной функцией?
2. Что такое реализация случайной функции?
3. Приведите примеры случайных функций.
4. Что понимают под реализацией случайного процесса?
5. Какой случайный процесс называют стационарным?

6. В чем суть эргодического свойства?

7. Поясните физический смысл понятий математического ожидания и дисперсии стационарного случайного процесса?

8. Как связаны между собой среднее квадратическое отклонение и дисперсия случайного процесса?

9. Какое свойство стационарного случайного процесса характеризует корреляционная функция?

10. Какая связь существует между дисперсией и корреляционной функцией стационарного случайного процесса?

4.4. ВАРИАНТЫ ЗАДАНИЙ

Вариант	Параметры генерации последовательности для программы в среде Matlab			
	K	s	lambda	T
1	5	2,3	2	1
2	4	4	1,7	3,5
3	5	2	1,7	2
4	5	1,5	1,3	10
5	8	2,5	1,3	10
6	12	2,5	1,25	2
7	3	1	2	2
8	4	1,1	2,2	2
9	10	2	2	0,05
10	20	3	1,2	1
11	30	3	1,1	2
12	15	2	1,1	10
13	4	1	3	0,5
14	5	10	1,9	1
15	6	0	1,7	1
16	7	-10	1,5	2
17	5	2	1,5	10
18	3	1	2,1	13
19	13	1	1,2	13
20	6	2,5	2	1,2
21	4	5	1,6	3,6
22	11	2,1	1,89	0,05
23	4	2	1,04	100
24	4	2	1,01	1000
25	3	-10	1,025	1000

ЗАКЛЮЧЕНИЕ

В процессе своей жизни человек часто сталкивается с событиями и явлениями, исход которых заранее не определен. Тем не менее в подобных ситуациях, связанных с неопределенностью, человеку необходимо принимать решения.

Обычно принятию решений предшествует анализ известных данных (на основании предшествующего опыта, здравого смысла, интуиции и т.д.). Стремясь увидеть и обосновать закономерности в неопределенных процессах, человечество выработало целый арсенал методов, которые называются математической статистикой (прикладной статистикой или анализом данных).

В данном пособии кратко рассмотрены основные направления методов математической статистики.

Выборочное наблюдение – решает задачу обобщения на всю совокупность результатов, полученных при изучении его части, например, анкетирование и т.д.

Проверка статистических гипотез – позволяет ответить на вопрос о достоверности принимаемого решения (например, обоснованность рейтинга популярности).

Дисперсионный анализ – изучает влияние факторных признаков на результативный (например, зависит ли производительность труда рабочего от стажа, возраста, стажа и возраста).

Корреляционно-регрессионный анализ – позволяет выявить связи и построить модели зависимости (например, какая зависимость существует между спросом на продукцию и курсом валюты).

Перечисленные выше методы основываются на теоретических положениях теории вероятности и быстро входят в нашу жизнь посредством пакетов прикладных математических (MathCad, MatLab) и статистических (STATISTICA) программ. Настоящее учебное пособие призвано помочь студентам в изучении основ теории вероятностей и математической статистики для решения многих практических задач.

Реальные социально-экономические, технические, технологические и другие процессы зависят от большого количества параметров, их характеризующих. Поэтому возникает необходимость использовать методы *многомерного статистического анализа*, которые обобщают методы проверки статистических гипотез, дисперсионный, корреляционно-регрессионный анализ и другие разделы на многомерный случай. Важным моментом при выборе математического инструментария является предварительный анализ данных – *разведочный анализ*, целью которого является исследование исходных данных, изучение их вероятностной и геометрической природы, рассмотрение и проверка рабочих гипотез. Круг рассмотренных выше методов обработки экспериментальных данных определяют как *прикладная статистика* (или *анализ данных*).

Более подробно ознакомиться с современными методами прикладной статистики можно по литературе [1, 5]. Эти направления мы предлагаем для дальнейшего самостоятельного изучения и применения на практике.

Следует отметить, что на современном этапе управление производством, фирмой, регионом практически невозможно без системного подхода, разрабатывающего методики анализа целей, методы и модели совершенствования организационной структуры, управления функционированием объектов.

В зависимости от информации об изучаемом объекте применяют следующие методы: мозговой атаки, построения сценариев, экспертной оценки, математической логики, теории множеств, теории игр, прикладной статистики, математического программирования и т.д. Разумеется большинство методов пересекается. При этом статистические методы в рамках системного анализа являются одним из возможных подходов перевода словесного описания модели изучаемого объекта в формальное, для решения задач управления и принятия решений.

Моделирование в большой степени искусство, овладеть которым можно только, решая практические задачи. Целью настоящего изложения является возможный путь поиска истины, двигаясь по которому можно приобрести неоценимый опыт.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. *Айвазян С.А.* Прикладная статистика. Основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1983. С. 472.

2. *Вентцель Е.С.* Теория вероятностей / Е.С. Вентцель. – М.: Высш. шк., 2001. – С. 575.

3. *Вентцель Е.С.* Теория вероятностей и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. – М.: Наука, 1988.

4. *Гмурман В.Е.* Теория вероятностей и математическая статистика / В.Е. Гмурман. – М.: Высш. шк., 1999. – С. 479.

5. *Дубров А.М.* Многомерные статистические методы: учебник / А.М. Дубров, В.С. Мхитарян, Л.И. Трошин. - М.: Финансы и статистика, 1998. – С. 352.

6. *Калинина В.Н.* Математическая статистика: учебник / В.Н. Калинина, В.Ф. Панкин. – М.: Высш. шк., 2001. – С. 336.

7. *Кассандрова О.Н.* Обработка результатов наблюдений / О.Н. Кассандрова, В.В. Лебедев. – М.: Наука, 1970. – С. 104.

8. *Колесников А.Ф.* Основы математической обработки результатов измерений / А.Ф. Колесников. – Томск: ТГУ, 1963. – С. 49.

9. *Кремер Н.Ш.* Теория вероятностей и математическая статистика / Н.Ш. Кремер. – М.: ЮНИТИ–ДАНА, 2007. – С. 551.

10. *Плескунин В.И.* Теоретические основы организации и анализа выборочных данных в эксперименте: учеб. пособие / В.И. Плескунин, Е.Д. Воронина. – Л.: ЛЭУ, 1979. – С. 232.

11. *Румшинский Л.З.* Математическая обработка результатов эксперимента: справ. руководство / Л.З. Румшинский. – М.: Наука, 1971. – С. 192.

12. *Рыжов Э.В.* Математические методы в технологических исследованиях / Э.В. Рыжков, О.А. Горленко. – Киев: Наук. думка, 1990. – С. 184.

13. *Смирнов Н.В.* Курс теории вероятностей и математической статистики для технических приложений / Н.В. Смирнов, И.В. Дунин-Барковский. – М.: Наука, 1969.

14. *Степанов М.М.* Аппроксимация функций: метод. указания к лабораторным работам по дисциплине «Информатика» /

М.М. Степанов, Н.Н. Потапова, Т.В. Ерещенко. – Волгоград: ВолгГАСУ, 2012.

15. *Сухов А.Н.* Математическая обработка результатов измерений: учеб. пособие/А.Н. Сухов. – М.: МИСИ, 1982. – С. 89.

16. *Тихонов В.И.* Статистический анализ и синтез радиотехнических устройств: учеб. пособие для вузов / В.И. Тихонов, В.Н. Харисов. – М.: Радио и связь, 2004.

17. *Третьяк Л.Н.* Обработка результатов наблюдений: учеб. пособие / Л.Н. Третьяк. – Оренбург. ГОУ ОГУ, 2004. – С. 171.

18. *Тюрин Ю.Н.* Статистический анализ данных на компьютерах. под ред. В.Э. Фигурнова / Ю.Н. Тюрин, А.А. Макаров. – М.: ИНФРА – М, 1998.

19. *Ферстер Э.* Методы корреляционного и регрессионного анализа: пер. с нем. / Э. Ферстер, Б. Ренц. – М.: Финансы и статистика, 1983.

20. *Чкалова О.Н.* Основы научных исследований / О.Н. Чкалова. – Киев: Вища шк., 1978. – С. 120.

21. *Веников В.А*. Электрические системы. Математические задачи электроэнергетики / В.А. Веников [и др.]. – М.: Высш. шк., 1981. – С. 288.

ПРИЛОЖЕНИЯ

Приложение 1

Таблица значений функции Лапласа $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\, dt$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,00	0,000000	0,31	0,121700	0,62	0,232400
0,01	0,004000	0,32	0,125500	0,63	0,235700
0,02	0,008000	0,33	0,129300	0,64	0,238900
0,03	0,012000	0,34	0,133100	0,65	0,242200
0,04	0,016000	0,35	0,136800	0,66	0,245400
0,05	0,019900	0,36	0,140600	0,67	0,248600
0,06	0,023900	0,37	0,144300	0,68	0,251700
0,07	0,027900	0,38	0,148000	0,69	0,254900
0,08	0,031900	0,39	0,151700	0,70	0,258000
0,09	0,035900	0,40	0,155400	0,71	0,261100
0,10	0,039800	0,41	0,159100	0,72	0,264200
0,11	0,043800	0,42	0,162800	0,73	0,267300
0,12	0,047800	0,43	0,166400	0,74	0,270300
0,13	0,051700	0,44	0,170000	0,75	0,273400
0,14	0,055700	0,45	0,173600	0,76	0,276400
0,15	0,059600	0,46	0,177200	0,77	0,279400
0,16	0,063600	0,47	0,180800	0,78	0,282300
0,17	0,067500	0,48	0,184400	0,79	0,285200
0,18	0,071400	0,49	0,187900	0,80	0,288100
0,19	0,075300	0,50	0,191500	0,81	0,291000
0,20	0,079300	0,51	0,195000	0,82	0,293900
0,21	0,083200	0,52	0,198500	0,83	0,296700
0,22	0,087100	0,53	0,201900	0,84	0,299500
0,23	0,091000	0,54	0,205400	0,85	0,302300
0,24	0,094800	0,55	0,208800	0,86	0,305100
0,25	0,098700	0,56	0,212300	0,87	0,307800
0,26	0,102600	0,57	0,215700	0,88	0,310600
0,27	0,106400	0,58	0,219000	0,89	0,313300
0,28	0,110300	0,59	0,222400	0,90	0,315900
0,29	0,114100	0,60	0,225700	0,91	0,318600
0,30	0,117900	0,61	0,229100	0,92	0,321200

Продолжение прил. 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0,93	0,323800	1,29	0,401500	1,65	0,450500
0,94	0,326400	1,30	0,403200	1,66	0,451500
0,95	0,328900	1,31	0,404900	1,67	0,452500
0,96	0,331500	1,32	0,406600	1,68	0,453500
0,97	0,334000	1,33	0,408200	1,69	0,454500
0,98	0,336500	1,34	0,409900	1,70	0,455400
0,99	0,338900	1,35	0,411500	1,71	0,456400
1,00	0,341300	1,36	0,413100	1,72	0,457300
1,01	0,343800	1,37	0,414700	1,73	0,458200
1,02	0,346100	1,38	0,416200	1,74	0,459100
1,03	0,348500	1,39	0,417700	1,75	0,459900
1,04	0,350800	1,40	0,419200	1,76	0,460800
1,05	0,353100	1,41	0,420700	1,77	0,461600
1,06	0,355400	1,42	0,422200	1,78	0,462500
1,07	0,357700	1,43	0,423600	1,79	0,463300
1,08	0,359900	1,44	0,425100	1,80	0,464100
1,09	0,362100	1,45	0,426500	1,81	0,464900
1,10	0,364300	1,46	0,427900	1,82	0,465600
1,11	0,366500	1,47	0,429200	1,83	0,466400
1,12	0,368600	1,48	0,430600	1,84	0,467100
1,13	0,370800	1,49	0,431900	1,85	0,467800
1,14	0,372900	1,50	0,433200	1,86	0,468600
1,15	0,374900	1,51	0,434500	1,87	0,469300
1,16	0,377000	1,52	0,435700	1,88	0,469900
1,17	0,379000	1,53	0,437000	1,89	0,470600
1,18	0,381000	1,54	0,438200	1,90	0,471300
1,19	0,383000	1,55	0,439400	1,91	0,471900
1,20	0,384900	1,56	0,440600	1,92	0,472600
1,21	0,386900	1,57	0,441800	1,93	0,473200
1,22	0,388300	1,58	0,442900	1,94	0,473800
1,23	0,390700	1,59	0,444100	1,95	0,474400
1,24	0,392500	1,60	0,445200	1,96	0,475000
1,25	0,394400	1,61	0,446300	1,97	0,475600
1,26	0,396200	1,62	0,447400	1,98	0,476100
1,27	0,398000	1,63	0,448400	1,82	0,465600
1,28	0,399700	1,64	0,449500	1,83	0,466400

Окончание прил. 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1,74	0,459100	1,93	0,473200	2,58	0,495100
1,75	0,459900	1,94	0,473800	2,60	0,4953
1,76	0,460800	1,95	0,474400	2,62	0,495600
1,77	0,461600	1,96	0,475000	2,64	0,495900
1,78	0,462500	1,97	0,475600	2,66	0,496100
1,79	0,463300	1,98	0,476100	2,56	0,494800
1,80	0,464100	1,99	0,476700	2,58	0,495100
1,81	0,464900	2,00	0,477200	2,60	0,4953
1,82	0,465600	2,02	0,478300	2,62	0,495600
1,83	0,466400	2,04	0,479300	2,64	0,495900
1,84	0,467100	2,06	0,480300	2,66	0,496100
1,85	0,467800	2,08	0,481200	2,66	0,496100
1,86	0,468600	2,10	0,482100	2,68	0,496300
1,87	0,469300	2,12	0,483000	2,70	0,496500
1,88	0,469900	2,14	0,483800	2,72	0,496700
1,89	0,470600	2,16	0,484600	2,74	0,496900
1,90	0,471300	2,18	0,485400	2,76	0,497100
1,91	0,471900	2,20	0,486100	2,78	0,497300
1,92	0,472600	2,22	0,486800	2,80	0,497400
1,93	0,473200	2,24	0,487500	2,82	0,497600
1,94	0,473800	2,26	0,488100	2,84	0,497700
1,95	0,474400	2,28	0,488700	2,86	0,497900
1,96	0,475000	2,30	0,489300	2,88	0,498000
1,97	0,475600	2,32	0,489800	2,90	0,498100
1,98	0,476100	2,34	0,490400	2,92	0,498200
1,82	0,465600	2,36	0,490900	2,94	0,498400
1,83	0,466400	2,38	0,491300	2,96	0,498500
1,84	0,467100	2,40	0,491800	2,98	0,498600
1,85	0,467800	2,42	0,492200	3,00	0,498650
1,86	0,468600	2,44	0,492700	3,20	0,499310
1,87	0,469300	2,46	0,493100	3,40	0,499660
1,88	0,469900	2,48	0,493400	3,60	0,499841
1,89	0,470600	2,50	0,493800	3,80	0,499928
1,90	0,471300	2,52	0,494100	4,00	0,499968
1,91	0,471900	2,54	0,494500	4,50	0,499997
1,92	0,472600	2,56	0,494800	5,00	0,499999

Приложение 2

Критические точки распределения χ^2 – Пирсона

Число степеней свободы k	Уровень значимости a		
	0,010	0,025	0,050
1	6,60000	5,00000	3,80000
2	9,20000	7,40000	6,00000
3	11,30000	9,40000	7,80000
4	13,30000	11,10000	9,50000
5	15,10000	12,80000	11,10000
6	16,80000	14,40000	12,60000
7	18,50000	16,00000	14,10000
8	20,10000	17,50000	15,50000
9	21,70000	19,00000	16,90000
10	23,20000	20,50000	18,30000
11	24,70000	21,90000	19,70000
12	26,20000	23,30000	21,00000
13	27,70000	24,70000	22,40000
14	29,10000	26,10000	23,70000
15	30,60000	27,50000	25,00000
16	32,00000	28,80000	26,30000
17	33,40000	30,20000	27,60000
18	34,80000	31,50000	28,90000
19	36,20000	32,90000	30,10000
20	37,60000	34,20000	31,40000
21	38,90000	35,50000	32,70000
22	40,30000	36,80000	33,90000
23	41,60000	38,10000	35,20000
24	43,00000	39,40000	36,40000
25	44,30000	40,60000	37,70000
26	45,60000	41,90000	38,90000
27	47,00000	43,20000	40,10000
28	48,30000	44,50000	41,30000
29	49,60000	45,70000	42,60000
30	50,90000	47,00000	43,80000

Окончание прил. 2

Число степеней свободы k	Уровень значимости a		
	0,950	0,975	0,990
1	0,00390	0,00098	0,00016
2	0,10300	0,05100	0,02000
3	0,35200	0,21600	0,11500
4	0,71100	0,48400	0,29700
5	1,15000	0,83100	0,55400
6	1,64000	1,24000	0,87200
7	2,17000	1,69000	1,24000
8	2,73000	2,18000	1,65000
9	3,33000	2,70000	2,09000
10	3,94000	3,25000	2,56000
11	4,57000	3,82000	3,05000
12	5,23000	4,40000	3,57000
13	5,89000	5,01000	4,11000
14	6,57000	5,63000	4,66000
15	7,26000	6,26000	5,23000
16	7,96000	6,91000	5,81000
17	8,67000	7,56000	6,41000
18	9,39000	8,23000	7,01000
19	10,10000	8,91000	7,63000
20	10,90000	9,59000	8,26000
21	11,60000	10,30000	8,90000
22	12,30000	11,00000	9,54000
23	13,10000	11,70000	10,20000
24	13,80000	12,40000	10,90000
25	14,60000	13,10000	11,50000
26	15,40000	13,80000	12,20000
27	16,20000	14,60000	12,90000
28	16,90000	15,30000	13,60000
29	17,70000	16,00000	14,30000
30	18,50000	16,80000	15,00000

Приложение 3

Таблица значений $t_\gamma = t(\gamma, n)$

n \ γ	0,95	0,99	0,999
5	2,78	4,60	8,61
6	2,57	4,03	6,86
7	2,45	3,71	5,96
8	2,37	3,50	5,41
9	2,31	3,36	5,04
10	2,26	3,25	4,78
11	2,23	3,17	4,59
12	2,20	3,11	4,44
13	2,18	3,06	4,32
14	2,16	3,01	4,22
15	2,15	2,98	4,14
16	2,13	2,95	4,07
17	2,12	2,92	4,02
18	2,11	2,90	3,97
19	2,10	2,88	3,92
20	2,093	2,861	3,883
25	2,064	2,797	3,745
30	2,045	2,756	3,659
35	2,032	2,720	3,600
40	2,023	2,708	3,558
45	2,016	2,692	3,527
50	2,009	2,679	3,502
60	2,001	2,662	3,464
70	1,996	2,649	3,439
80	1,991	2,640	3,418
90	1,987	2,633	3,403
100	1,984	2,627	3,392
120	1,980	2,617	3,374
∞	1,960	2,576	3,291

Приложение 4

Таблица значений $q = q(\gamma, n)$

n \ γ	0,95	0,99	0,999
5	1,37	2,67	5,64
6	1,09	2,01	3,88
7	0,92	1,62	2,98
8	0,8	1,38	2,42
9	0,71	1,2	2,06
10	0,65	1,08	1,8
11	0,59	0,98	1,6
12	0,55	0,9	1,45
13	0,52	0,83	1,33
14	0,48	0,78	1,23
15	0,46	0,73	1,15
16	0,44	0,7	1,07
17	0,42	0,66	1,01
18	0,4	0,63	0,96
19	0,39	0,6	0,92
20	0,37	0,58	0,88
25	0,32	0,49	0,73
30	0,28	0,43	0,63
35	0,26	0,38	0,56
40	0,24	0,35	0,50
45	0,22	0,32	0,46
50	0,21	0,30	0,43
60	0,188	0,269	0,38
70	0,174	0,245	0,34
80	0,161	0,226	0,31
90	0,151	0,211	0,29
100	0,143	0,198	0,27
150	0,115	0,160	0,211
200	0,099	0,136	0,185
250	0,089	0,120	0,162

Приложение 5

Критические точки распределения Стьюдента

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,1	0,05	0,02	0,01	0,002	0,001
1	6,31	12,7	31,82	63,7	318,3	637,0
2	2,92	4,30	6,97	9,92	22,33	31,6
3	2,35	3,18	4,54	5,84	10,22	12,9
4	2,13	2,78	3,75	4,60	7,17	8,61
5	2,01	2,57	3,37	4,03	5,89	6,86
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,40
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,03	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,95
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,73	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,51	3,79
23	1,71	2,07	2,50	2,81	3,49	3,77
24	1,71	2,06	2,49	2,80	3,47	3,74
25	1,71	2,06	2,49	2,79	3,45	3,72
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,71	2,05	2,47	2,77	3,42	3,69
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

Окончание прил. 5

Число степеней свободы k	Уровень значимости α (двусторонняя критическая область)					
	0,1	0,05	0,02	0,01	0,002	0,001
28	1,70	2,05	2,46	2,76	3,40	3,66
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,17	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29
	0,05	0,025	0,01	0,005	0,001	0,0005
	Уровень значимости α (односторонняя критическая область)					

Приложение 6

Критические точки распределения F Фишера – Снедекора. Уровень значимости $\alpha = 0{,}01$ (k_1 – число степеней свободы большей дисперсии, k_2 – число степеней свободы меньшей дисперсии)

k_2 \ k_1	1	2	3	4	5	6
1	4052	4999	5403	5625	5764	5889
2	98,49	99,01	90,17	99,25	99,33	99,30
3	34,12	30,81	29,46	28,71	28,24	27,91
4	21,20	18,00	16,69	15,98	15,52	15,21
5	16,26	13,27	12,06	11,39	10,97	10,67
6	13,74	10,92	9,78	9,15	8,75	8,47
7	12,25	9,55	8,45	7,85	7,46	7,19
8	11,26	8,65	7,59	7,01	6,63	6,37
9	10,56	8,02	6,99	6,42	6,06	5,80
10	10,04	7,56	6,55	5,99	5,64	5,39
11	9,86	7,20	6,22	5,67	5,32	5,07
12	9,33	6,93	5,95	5,41	5,06	4,82
13	9,07	6,70	5,74	5,20	4,86	4,62
14	8,86	6,51	5,56	5,03	4,69	4,46
15	8,68	6,36	5,42	4,89	4,56	4,32
16	8,53	6,23	5,29	4,77	4,44	4,20
17	8,40	6,11	5,18	4,67	4,34	4,10

Критические точки распределения F Фишера – Снедекора. Уровень значимости $\alpha = 0{,}01$ (k_1 – число степеней свободы большей дисперсии, k_2 – число степеней свободы меньшей дисперсии)

k_2 \ k_1	7	8	9	10	11	12
1	5928	5981	6022	6056	6082	6106
2	99,34	99,36	99,36	99,40	99,41	99,42
3	27,67	27,49	27,34	27,23	27,13	27,05
4	14,98	14,80	14,66	14,54	14,45	14,37
5	10,45	10,27	10,15	10,05	9,96	9,89
6	8,26	8,10	7,98	7,87	7,79	7,72
7	7,00	6,84	6,71	6,62	6,54	6,47
8	6,19	6,03	5,91	5,82	5,74	5,67
9	5,62	5,47	5,35	5,26	5,18	5,11
10	5,21	5,06	4,95	4,85	4,78	4,71
11	4,88	4,74	4,63	4,54	4,46	4,40
12	4,65	4,50	4,39	4,30	4,22	4,16
13	4,44	4,30	4,19	4,10	4,02	3,96
14	4,28	4,14	4,03	3,94	3,86	3,80
15	4,14	4,00	3,89	3,80	3,73	3,67
16	4,03	3,89	3,78	3,69	3,61	3,55
17	3,93	3,79	3,68	3,59	3,52	3,45

Критические точки распределения F Фишера – Снедекора. Уровень значимости $\alpha = 0{,}05$ (k_1 – число степеней свободы большей дисперсии, k_2 – число степеней свободы меньшей дисперсии)

k_2 \ k_1	1	2	3	4	5	6
1	161	200	216	225	230	234
2	18,5	19,00	19,16	19,25	19:30	19,33
3	10,13	9,55	9,28	9,12	9,01	8,94
4	7,71	6,94	6,59	6,39	6,26	6,16
5	6,61	5,79	5,41	5,19	5,05	4,95
6	5,99	5,14	4,76	4,53	4,39	4,28
7	5,59	4,74	4,35	4,12	3,97	3,87
8	5,32	4,46	4,07	3,84	3,69	3,58
9	5,12	4,26	3,86	3,63	3,48	3,37
10	4,96	4,10	3,71	3,48	3,33	3,22
11	4,84	3,98	3,59	3,36	3,20	3,09
12	4,75	3,88	3,49	3,26	3,11	3,00
13	4,67	3,80	3,41	3,18	3,02	2,92
14	4,60	3,74	3,34	3,11	2,96	2,85
15	4,54	3,68	3,29	3,06	2,90	2,79
16	4,49	3,63	3,24	3,01	2,85	2,74
17	4,45	3,59	3,20	2,96	2,81	2,70

Критические точки распределения F Фишера – Снедекора. Уровень значимости $\alpha = 0{,}05$ (k_1 – число степеней свободы большей дисперсии, k_2 – число степеней свободы меньшей дисперсии)

k_2 \ k_1	7	8	9	10	11	12
1	237	239	241	242	243	244
2	19,36	19,37	19,38	19,39	19,40	19,41
3	8,88	8,84	8,81	8,78	8,76	8,74
4	6,09	6,04	6,00	5,96	5,93	5,91
5	4,88	4,82	4,78	4,74	4,70	4,68
6	4,21	4,15	4,10	4,06	4,03	4,00
7	3,79	3,73	3,68	3,63	3,60	3,57
8	3,50	3,44	3,39	3,34	3,31	3,28
9	3,29	3,23	3,18	3,13	3,10	3,07
10	3,14	3,07	3,02	2,97	2,94	2,91
11	3,01	2,95	2,90	2,86	2,82	2,79
12	2,92	2,85	2,80	2,76	2,72	2,69
13	2,84	2,77	2,72	2,67	2,63	2,60
14	2,77	2,70	2,65	2,60	2,56	2,53
15	2,70	2,64	2,59	2,55	2,51	2,48
16	2,66	2,59	2,54	2,49	2,45	2,42
17	2,62	2,55	2,50	2,45	2,41	2,38

ОГЛАВЛЕНИЕ

Учебное издание

Тобоев Вячеслав Андреевич
Картузова Татьяна Вячеславовна
Толстов Михаил Сергеевич

ПРИКЛАДНОЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ ДЛЯ ИНЖЕНЕРОВ

Учебное пособие

Редактор А.Н. Антонова
Компьютерная верстка и правка М.С. Толстова

Согласно Закону № 436-ФЗ от 29 декабря 2010 года
данная продукция не подлежит маркировке

Подписано в печать 30.06.14. Формат $60 \times 84/16$.
Бумага газетная. Печать офсетная. Гарнитура Times.
Усл. печ. л. 6,27. Уч.-изд. л. 6,12. Тираж 200 экз. Заказ № 491.

Издательство Чувашского университета
Типография университета
428015 Чебоксары, Московский просп., 15