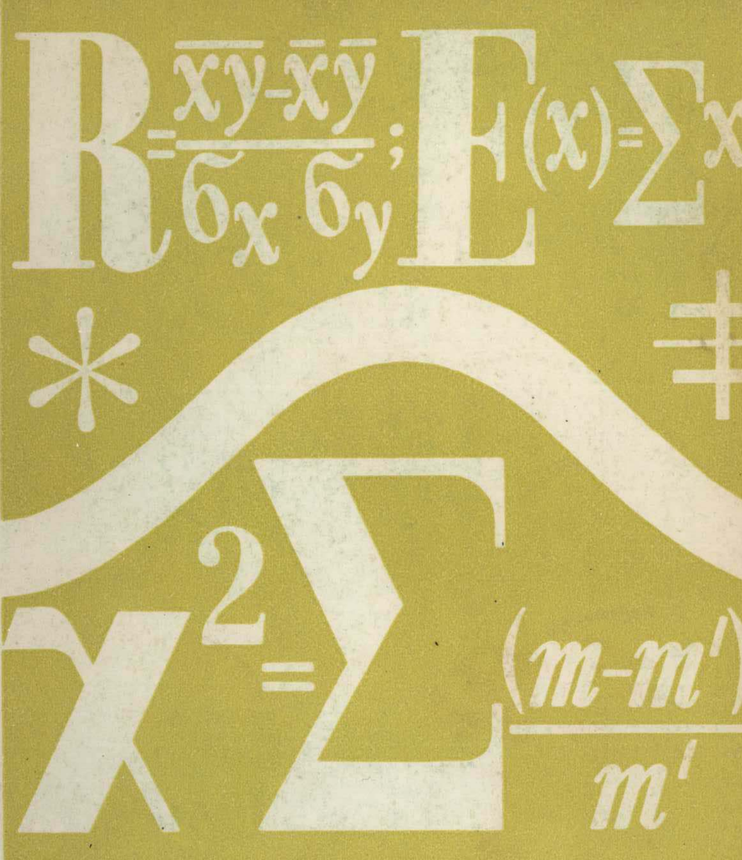
МАТЕМАТИЧЕСКАЯ СТАТИСТИКА ДЛЯ ЭКОНОМИСТОВ



В. С. Мхитарян

СТАТИСТИЧЕСКИЕ МЕТОДЫ В УПРАВЛЕНИИ КАЧЕСТВОМ ПРОДУКЦИИ

В. С. Мхитарян

СТАТИСТИЧЕСКИЕ МЕТОДЫ В УПРАВЛЕНИИ КАЧЕСТВОМ ПРОДУКЦИИ

МОСКВА «ФИНАНСЫ И СТАТИСТИКА» 1982

ББК 22.172
М93

Мхитарян В. С.
М93 Статистические методы в управлении качеством продукции. — М.: Финансы и статистика, 1982. — 119 с., ил. — (Мат. статистика для экономистов).
60 к.

Рассматриваются два направления применения статистических методов в управлении качеством продукции: разработка планов статистического (выборочного) контроля и анализ информации, накапливаемой в процессе его проведения. Приводится методика анализа результатов контроля по количественным и качественным признакам.
Для экономистов, статистиков, специалистов, занимающихся управлением качества продукции.

М $\dfrac{1702060000-153}{010(01)-82}$ 23—82

ББК 22.17
517.

ВВЕДЕНИЕ

Качество продукции является одним из важнейших показателей эффективности общественного производства, поэтому в настоящее время обращается внимание на необходимость существенного улучшения качества всех видов выпускаемой продукции, увеличения производства новых видов изделий, отвечающих современным требованиям.

Основными направлениями экономического и социального развития СССР на 1981—1985 годы и на период до 1990 года поставлена задача: «Значительно повысить качество всех видов выпускаемой продукции, расширять и обновлять ассортимент изделий в соответствии с современными требованиями развития народного хозяйства и научно-технического прогресса, а также растущими потребностями населения. Неуклонно увеличивать удельный вес продукции высшей категории качества в общем объеме ее выпуска. Активно внедрять комплексные системы управления качеством продукции»[1]. Решение такой задачи требует постоянной целенаправленной работы производственных коллективов на всех участках, начиная с момента проектирования изделия и кончая сдачей его потребителю.

Качество продукции является обобщенным показателем, зависящим от уровня научных исследований и конструирования, разработки технологии производства и контроля. Требования к качеству продукции постоянно изменяются и совершенствуются в зависимости от общего научного и технического уровня развития данной отрасли.

Важная роль в решении проблемы повышения качества отводится комплексным системам управления качеством продукции, которые охватывают этапы разработки, производства и эсплуатации изделий.

Комплексная система призвана всемерно способствовать выпуску высококачественной продукции, наиболее

---

[1]Материалы XXVI съезда КПСС. М., Политиздат, 1981, с. 141.

полно удовлетворяющей запросам потребителя, получению при этом наибольшей прибыли и увеличению рентабельности производства путем повышения производительности труда, снижения себестоимости продукции. В сферу ее деятельности входит управление качеством выпускаемой продукции, комплектующих изделий и материалов, оборудования и оснастки, качеством нормативно-технической документации, а также качеством труда работников.

В комплексной системе управления статистические методы, используемые на базе современных ЭВМ, должны сыграть важную роль при сборе и анализе информации о качестве продукции и работы оборудования на всех этапах производства, а также о качестве труда подразделений и отдельных исполнителей. В зависимости от степени агрегирования эта информация может быть использована для управления качеством на уровне руководства предприятия, цеха, бригады и отдельного работника.

Соответствие показателей качества продукции установленным требованиям (стандартам, техническим условиям и другой документации) определяется путем проведения контроля. В зависимости от охвата проверяемой продукции различают сплошной и выборочный контроль. При сплошном контроле о качестве продукции судят по результатам проверки всех изделий. Однако такой контроль часто не применим из экономических соображений при крупносерийном и массовом производстве. Кроме того, в ряде случаев испытания приводят к полной или частичной потере изделиями качества (испытания на надежность, на срок службы и т. д.). В этих случаях применяют статистический (выборочный) контроль, согласно которому контролируют не все изделия, а лишь сравнительно небольшую их долю.

Статистический контроль используется как для регулирования хода технологического процесса, так и для оценки качества партий продукции. В первом случае преследуется цель предупреждения брака путем периодического наблюдения за ходом процесса и своевременного вмешательства в него. Во втором случае решается чисто контрольная задача — проверка соответствия партии сырья (входной контроль) или готовой продукции (приемочный контроль) техническим условиям. Оба эти метода базируются на выборочных процедурах математической статистики. При выборочном контроле необходимо по результатам проверки части изделий сделать вывод о качестве всей партии.

В зависимости от вида контрольной операции различают контроль по альтернативному, качественному и количественному признакам. *При контроле по альтернативному признаку* изделия по результатам измерений разбивают на годные и дефектные. При *контроле по качественному признаку* изделия классифицируются на несколько групп. Например, по результатам контроля изделие может быть отнесено к 1, 2, 3-му сорту или к браку. При *контроле по количественному признаку* измеряется числовое значение параметра.

Каждый из видов контроля имеет свои преимущества и недостатки. Например, при контроле по качественному признаку каждое наблюдение несет меньшую информацию, чем при контроле по количественному признаку, в связи с чем для получения обоснованных решений требуется большее число наблюдений. Но при качественном контроле процесс наблюдений проще и его легче автоматизировать. Кроме того, методика качественного контроля не связана с видом распределения контролируемого признака и поэтому является более универсальной.

В деле внедрения статистических методов управления качеством продукции на предприятиях большое значение имеет их стандартизация. Методология применения государственных стандартов предусматривает, что вся нормативная техническая документация на приемку продукции, контроль технологических процессов и статистический анализ качества должна выполняться в строгом соответствии с требованиями стандартов. Основная цель, преследуемая стандартизацией статистических методов контроля качества, заключается в установлении единых норм, требований и методов в области контроля и испытаний с целью получения оптимального качества продукции.

Прогрессивное значение стандартизации обусловлено тем, что стандарты отражают наиболее важные научные достижения в соответствующей области, систематическое обновление стандартов (через каждые пять лет) способствует постоянному повышению их научно-технического уровня, кроме того, стандарт исключает неправильное решение тех или иных технических вопросов.

В СССР первые стандарты по статистическим методам контроля были изданы в 1970 г. и сыграли важную роль в деле внедрения этих методов в производство.

В настоящее время система отечественных стандартов по статистическим методам управления качеством продукции охватывает такие направления, как [37]: анализ точности технологических процессов, методы регулирования технологических процессов, методы приемочного контроля, методы контроля времени безотказной работы, средства механизации и автоматизации статистического контроля, терминология по статистическим методам контроля, организация внедрения статистических методов контроля, прикладная статистика.

В качестве справочных руководств по методам статистического контроля можно рекомендовать книги Д. Коудена [23], Б. Л. Хэнсена [44], Э. Шиндовского и О. Шюрца [39]. Теоретическим основам выборочного контроля посвящена монография Ю. К. Беляева [5].

В настоящей работе рассмотрены два направления применения методов математической статистики в системе управления качеством продукции: разработка планов статистического контроля и анализ информации, накапливаемой в процессе контроля с целью повышения его эффективности. Работа иллюстрирована примерами, характеризующими возможности статистических методов в системе управления качеством.

В гл. 1 кратко, в справочной форме, изложены некоторые общие положения математической статистики, используемые при дальнейшем изложении.

В гл. 2 большое внимание уделяется методам построения планов выборочного контроля качества продукции и статистического регулирования технологических процессов, определению их характеристик, системе отечественных стандартов по статистическим методам управления качеством.

В гл. 3 и 4 рассматриваются некоторые задачи статистического анализа результатов при контроле изделий по альтернативному, качественному и количественному признакам, исследуется обобщенный показатель качества при контроле по нескольким альтернативным признакам.

В приложении приведены фрагменты статистических таблиц, необходимые для решения рассматриваемых примеров. Читателю можно рекомендовать достаточно полные статистические таблицы [6, 43].

Глава 1

ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

1.1. СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ

Методы математической статистики позволяют анализировать результаты опытов или наблюдений и строить оптимальные математико-статистические модели массовых, повторяющихся явлений, изменчивость которых обусловлена рядом неуправляемых факторов.

Будем называть *генеральной совокупностью* множество всех наблюдений, которые могут быть сделаны при данном комплексе условий. Под комплексом условий обычно понимают объективно существующие условия, определяющие те закономерности, в соответствии с которыми происходит случайное варьирование признаков при повторениях наблюдений.

Так как во многих случаях обследование всей генеральной совокупности (всех изделий, которые могли бы быть изготовлены при данном технологическом режиме) либо слишком трудоемко, либо принципиально невозможно, то в практике контроля часто ограничиваются лишь некоторой *выборкой* из генеральной совокупности, т. е. рядом наблюдений $x_1, x_2, \ldots, x_n$ исследуемой случайной величины X. При этом, как правило, подразумевают, что наблюдения $x_1, x_2, \ldots, x_n$ *случайно* отбирают из генеральной совокупности X и каждое наблюдение генеральной совокупности имеет одинаковую вероятность быть отобранным.

Оценка параметров. Задачи математической статистики практически сводятся к оценке свойств генеральной совокупности по результатам случайной выборки.

Любую функцию $Q_n^* (x_1, \ldots, x_n)$ от результатов выборочных наблюдений $x_1, x_2, \ldots, x_n$ принято называть *статистикой* (выборочной характеристикой). Статистики обычно и используются для построения статистических оценок параметров Q генеральной совокупности, когда точные значения этих параметров нам не известны.

Статистику Q_n^*, используемую как оценку параметра Q, называют *точечной оценкой*. Из точечных оценок в приложениях математической статистики наиболее часто используют среднюю арифметическую

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{1.1}$$

как оценку математического ожидания $MX = \mu$;
выборочную дисперсию

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{1.2}$$

и среднее квадратическое отклонение

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{1.3}$$

как оценки генеральной дисперсии $DX = \sigma^2$ и среднего квадратического отклонения σ,
а также начальные

$$\nu_k^* = \frac{1}{n} \sum_{i=1}^{n} x_i^k \tag{1.4}$$

и центральные

$$\mu_k^* = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^k \tag{1.5}$$

моменты до четвертого порядка включительно, т. е. $k = 1, 2, 3, 4$, как оценки соответствующих моментов генеральной совокупности.

Очевидно, что средняя арифметическая равна первому начальному моменту $\nu_1^* = \bar{x}$, а дисперсия выборки — второму центральному моменту выборки $\mu_2^* = s^2$.

В качестве оценки вероятности p (генеральной доли) появления какого-то события в практике статистического приемочного контроля используют выборочную долю $\omega = \frac{m}{n}$, где m — число появлений этого события в выборке объемом n.

При определении оценки $\hat{x}_{\text{med}}$ *медианы* по выборке объемом n наблюдения располагают в порядке возрастания: $x^{(1)} < x^{(2)} < \ldots < x^{(n)}$. За оценку $\hat{x}_{\text{med}}$ принимают средний член этого ряда. Если же число наблюдений четное, то в качестве оценки медианы берут среднюю арифметическую двух средних значений.

Оценка *коэффициента асимметрии*, характеризующая скошенность вариационного ряда, определяется как

$$A^* = \frac{\mu_3^*}{s^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}}, \tag{1.6}$$

а оценка эксцесса, характеризующая островершинность вариационного ряда,— по формуле

$$E^* = \frac{\mu_4^*}{s^4} - 3 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{2}} - 3. \tag{1.7}$$

Если средняя арифметическая, медиана и среднее квадратическое отклонение имеют размерность признака, то коэффициенты асимметрии и эксцесса являются безразмерными величинами.

В отличие от параметров генеральной совокупности *всякая статистика*, а следовательно, и *точечная оценка являются случайной величиной*. Например, если доля изделий 1-го сорта в партии, состоящей из $N = 1000$ изделий, оценивается по выборке объемом $n = 100$, то, повторив несколько раз соответствующее выборочное обследование, мы увидим, что эта оценка, т. е. доля изделий 1-го сорта в выборке, варьирует случайным образом от одной выборки к другой. В то же время существует некоторое истинное значение доли (генеральная доля) изделий 1-го сорта в данной партии, около которого и происходит это случайное варьирование ее статистических оценок.

В математической статистике в зависимости от задачи статистику рассматривают либо как случайную величину, либо как число (конкретную реализацию случайной величины). Возникает вопрос, каким требованиям должны отвечать точечные оценки, чтобы их можно было считать в каком-то определенном смысле «хорошими»? Эти требования характеризуются понятиями несмещенности, состоятельности и эффективности.

Оценку Q_n^* называют *несмещенной*, если при любом объеме выборки n ее математическое ожидание равно оцениваемому параметру Q, т. е. $MQ_n^* = Q$. Смещением оценки B_n называют разность $B_n = MQ_n^* - Q$. Смещение характеризует систематическую погрешность, которая, вообще говоря, зависит от объема выборки.

Можно показать, что средняя арифметическая $\overline{x}$ есть несмещенная оценка математического ожидания $M\overline{x} = \mu$, а выборочная дисперсия s^2 является смещенной оценкой генеральной дисперсии $DX = \sigma^2$. В самом деле,

$$Ms^2 = M\left(\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n}\right) = \frac{n-1}{n}\sigma^2,$$

т. е. $Ms^2 \neq \sigma^2$.

Смещение оценки $B_n = Ms^2 - \sigma^2 = -\frac{1}{n}\sigma^2$ монотонно убывает при неограниченном возрастании n, т. е. $B_n \to 0$ при $n \to \infty$, что важно при исследовании состоятельности оценки. Отсюда ясно, что *требование несмещенности особенно существенно при малом количестве наблюдений.*

Однако оценку s^2 можно легко исправить. Несмещенной оценкой генеральной дисперсии σ^2 является скорректированная выборочная дисперсия

$$\hat{s^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2.$$

Оценка Q_n^* параметра Q называется *состоятельной*, если по мере роста числа наблюдений n (т. е. при $n \to N$ — в случае конечной генеральной совокупности объемом N или при $n \to \infty$ в случае бесконечной генеральной совокупности) она стремится (сходится по вероятности[1]) к оцениваемому параметру Q, т. е. $\lim\limits_{n \to \infty} Q_n^* = Q$.

Условием состоятельности оценок является стремление к нулю при неограниченном возрастании n смещения B_n и

[1]«Сходимость по вероятности» означает, что при любом как угодно малом $\varepsilon > 0$ вероятность события $\{Q_n^* - \varepsilon < Q < Q_n^* + \varepsilon\}$ стремится по мере возрастания объема выборки n к единице.

дисперсии DQ_n^* оценки (т. е. $B_n \to 0$ и $DQ_n^* \to 0$ при $n \to \infty$).

Можно показать, что средняя арифметическая $\bar{x}$, выборочная дисперсия s^2 и относительная частота ω есть состоятельные оценки соответственно математического ожидания μ, генеральной дисперсии σ^2 и вероятности p.

При выполнении условия состоятельности имеет смысл увеличивать объем наблюдений, так как это приводит к повышению точности оценивания.

Дисперсия выборочной оценки связана с еще одним ее важным свойством — эффективностью. *Несмещенная оценка Q_n^* параметра Q называется эффективной*, если среди прочих несмещенных оценок того же параметра она обладает наименьшей дисперсией. Требование эффективности оценки основано на логическом правиле, заключающемся в том, что если имеется несколько несмещенных оценок параметра, то следует отдать предпочтение той из них, которая подвержена меньшим случайным колебаниям около неизвестного нам значения оцениваемого параметра (меньшему рассеянию) при переходе от одной выборки объема n к другой, т. е. оценке с наименьшей дисперсией $D(Q_n^*)$.

Например, если выборка взята из генеральной совокупности X, имеющей нормальное распределение, то при $n \to \infty$ дисперсия средней арифметической $D\bar{x} = \frac{1}{n}\sigma^2$ в $\pi/2$ раз меньше, чем дисперсия медианы $Dx_{\text{med}} = \frac{\pi\sigma^2}{2n}$. Отсюда следует, что в случае нормального распределения для получения с помощью медианы оценки генеральной средней μ той же точности, что и оценка, полученная с помощью $\bar{x}$, необходимо увеличить объем выборки в $\pi/2 = 1{,}57$ раза.

Эффективность оценки зависит от закона распределения исследуемой совокупности и объема выборки n.

Наиболее оптимальным методом получения хороших оценок является *метод наибольшего правдоподобия*. Он основан на использовании функции правдоподобия

$$L = L(x_1, x_2, \ldots, x_n;\ Q) = f(x_1, Q) \cdot f(x_2, Q) \cdot \ldots \cdot f(x_n, Q),$$

где $f(x, Q)$ — функция вероятностей (для дискретной случайной величины) или плотность вероятностей (для непрерывной случайной величины); $Q = (Q_1, Q_2, \ldots, Q_k)^T$ —

вектор параметров закона распределения, подлежащих оцениванию по выборке.

Функцию правдоподобия можно рассматривать как вероятность (или плотность вероятностей) совместного появления результатов выборки ($x_1, x_2, \ldots, x_n$). За оценки наибольшего правдоподобия принимают значения $\hat{Q}_1, \hat{Q}_2, \ldots$ $\ldots, \hat{Q}_k$, максимизирующие функцию правдоподобия. Обычно эти оценки находят, решая систему k уравнений $\frac{\partial \ln L}{\partial Q_j} = 0$, где $j = 1, 2, \ldots, k$.

Оценка $\hat{Q}$ наибольшего правдоподобия эффективна, если параметр Q имеет эффективную оценку, состоятельна, имеет асимптотически нормальное распределение, с математическим ожиданием Q и конечной дисперсией.

Интервальные оценки. Точечная оценка без указания степени точности и надежности малоинформативна, так как наблюдаемые значения статистики есть лишь значения случайной величины. Она может существенно отличаться от оцениваемого параметра при малом объеме выборки, что приводит к грубым ошибкам.

Интервальной оценкой параметра Q называют такой интервал ($Q_n^{*(1)}, Q_n^{*(2)}$), относительно которого можно утверждать с определенной, близкой к единице вероятностью γ, что он содержит неизвестное значение параметра Q. Величину γ называют *доверительной вероятностью* или *надежностью оценки* параметра Q; $Q_n^{*(1)}$ и $Q_n^{*(2)}$ — некоторые функции от результатов выборочных наблюдений $x_1, x_2, \ldots, x_n$. Разность $h = Q_n^{*(2)} - Q_n^{*(1)}$ между верхней и нижней границами доверительного интервала называют *шириной доверительного интервала, а величину* $\delta = h/2$ — *точностью оценки.*

Ширина доверительного интервала зависит от объема выборки n и от величины доверительной вероятности: она уменьшается с ростом n и увеличивается с приближением доверительной вероятности к единице.

Для построения интервальных оценок необходимо знать закон распределения статистики Q_n^*.

Рассмотрим законы распределения наиболее часто употребляемых в практике статистического контроля и анализа выборочных характеристик. Предварительно кратко остановимся на нормальном законе распределения.

Непрерывная случайная величина X, принимающая значения от $-\infty$ до $+\infty$, имеет нормальное распределение, если ее плотность определяется как

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{1.8}$$

Функция распределения нормального закона имеет вид

$$F(x)=\int\limits_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(u-\mu)^2}{2\sigma^2}}\,\mathrm{d}u, \tag{1.9}$$

где числовые характеристики μ и σ характеризуют соответственно математическое ожидание и среднее квадратическое отклонение случайной величины X.

При нормальном распределении медиана совпадает с генеральной средней $x_{\text{med}}=\mu$, а коэффициенты асимметрии и эксцесса равны нулю ($A=0$, $E=0$).

Для удобства пользования нормальным распределением имеются специальные таблицы плотности нормированного нормального распределения $f(t)$ и соответствующей интегральной функции $\Phi(t)$:

$$f(t)=\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{t^2}{2}} \tag{1.10}$$

и

$$\Phi(t)=\frac{2}{\sqrt{2\pi}}\int\limits_{0}^{t}\mathrm{e}^{-\frac{u^2}{2}}\,\mathrm{d}u, \tag{1.11}$$

где $t=\frac{x-\mu}{\sigma}$ — нормированная нормально распределенная случайная величина с $\mathrm{M}t=0$ и $\mathrm{D}t=1$.

Простые приемы позволяют использовать эти таблицы и применительно к нормальным законам с произвольными значениями параметров μ и σ. Вероятность попадания случайной величины X с параметрами μ и σ в интервал от x_1 до x_2 равна:

$$P(x_1<X<x_2)=\frac{1}{2}\,[\Phi(t_2)-\Phi(t_1)], \tag{1.12}$$

где $t_1=\frac{x_1-\mu}{\sigma}$, $t_2=\frac{x_2-\mu}{\sigma}$.

Величины $\Phi(t_1)$ и $\Phi(t_2)$ находятся по таблице интегральной функции $\Phi(t)$ по значениям t_1 и t_2 (табл. П.1).

Пользуясь соотношением (1.12) и табл. П.1, можно определить, что вероятности попадания нормально распределенной величины X в интервалы $(\mu - \sigma, \mu + \sigma)$, $(\mu - 2\sigma, \mu + 2\sigma)$ и $(\mu - 3\sigma, \mu + 3\sigma)$ соответственно равны 0,68269; 0,95450 и 0,99730. Вероятность нахождения случайной величины в интервале $(\mu - 3\sigma, \mu + 3\sigma)$ весьма близка к единице (0,9973). Поэтому «трехсигмовые» границы $\mu \pm 3\sigma$ принимают за границы предельных возможных значений нормально распределенной случайной величины. В практике статистического контроля «трехсигмовые» границы (правило «трех сигм») широко используют при статистическом регулировании технологических процессов.

Законы распределения выборочных характеристик. Для любой случайной выборки объема n из генеральной совокупности X нормально распределенной с математическим ожиданием μ и дисперсией σ^2, выборочные характеристики имеют следующие законы распределения:

$$1)\ T = \frac{\bar{x} - \mu}{\sigma}\sqrt{n} \tag{1.13}$$

— нормированный нормальный закон распределения $N(0,1)$ с математическим ожиданием, равным нулю ($MT = 0$), и дисперсией, равной единице ($DT = 1$);

$$2)\ \frac{\bar{x} - \mu}{s}\sqrt{n-1} \tag{1.14}$$

— распределение Стьюдента (t-распределение) с $n - 1$ степенями свободы;

$$3)\ \frac{x_i - \mu}{\sigma} \tag{1.15}$$

— нормированное нормальное распределение $N(0,1)$;

$$4)\ \frac{x_i - \mu}{s}\sqrt{\frac{n-1}{n}} \tag{1.16}$$

— распределение Стьюдента (t-распределение) с $n - 1$ степенями свободы;

$$5)\ \frac{ns^2}{\sigma^2} \tag{1.17}$$

— распределение χ^2 с $n - 1$ степенями свободы.

Для двух независимых выборок из нормальных генеральных совокупностей X и Y статистики имеют следующие законы распределения:

$$6)\ \frac{\bar{x}-\bar{y}}{\sqrt{\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}}}\sqrt{\frac{n_1\cdot n_2}{n_1+n_2}} \qquad (1.18)$$

— распределение Стьюдента (t-распределение) с n_1+n_2- -2 степенями свободы, где $\bar{x}$ и $\bar{y}$ — выборочные средние двух независимых выборок объемом n_1 и n_2 из генеральных совокупностей X и Y, имеющих нормальное распределение с одинаковыми, но неизвестными параметрами μ и σ, s_1^2 и s_2^2 — выборочные дисперсии соответственно первой и второй выборок;

$$7)\ F=\frac{\hat{s}_1^2}{\hat{s}_2^2} \qquad (1.19)$$

— распределение Фишера — Снедекора (F-распределение) с n_1-1 и n_2-1 степенями свободы, где $\hat{s}_1^2>\hat{s}_2^2$ и $\hat{s}^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$; $\hat{s}_1^2$ и $\hat{s}_2^2$ — исправленные выборочные дисперсии двух независимых выборок объемом соответственно n_1 и n_2, взятых из нормальных совокупностей с одинаковыми средними квадратическими отклонениями.

1.2. СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

При разработке планов статистического контроля и регулирования качества, а также при статистическом анализе полученных результатов приходится решать задачи сравнения. Например, сравнение точности работы станка или измерительной системы со стандартом, сравнение качества продукции нескольких станков и т. д. Для решения этих задач по данным выборочных наблюдений используют статистические критерии, т. е. методы статистической проверки гипотез. Рассмотрим некоторые общие положения.

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

Нулевой гипотезой (H_0) называют выдвинутую гипотезу, которую нужно проверить, а гипотезу (H_1), противоположную нулевой, называют *конкурирующей*.

Под *статистическим критерием* понимают однозначно определенное правило, устанавливающее условия, при которых проверяемую гипотезу (H_0) по результатам наблюдений следует либо отвергнуть, либо не отвергнуть. Основу критерия составляет специально составленная выборочная характеристика (статистика) $Q_n^* = Q_n^*(x_1, x_2, \ldots, x_n)$, точное или приближенное распределение которой известно, где $x_1, x_2, \ldots, x_n$ — выборочные наблюдения. Каждый критерий разбивает все множество возможных значений статистики на две непересекающиеся области: критическую область и область принятия гипотезы. Если наблюдаемое значение статистики критерия попадает в критическую область, то гипотезу (H_0) отвергают. В противном случае гипотезу не отвергают. При использовании статистического критерия возможны четыре случая:

— гипотеза H_0 верна и ее принимают согласно критерию;

— гипотеза H_0 неверна и ее отвергают согласно критерию;

— гипотеза H_0 верна, но ее отвергают согласно критерию (*ошибка первого рода*);

— гипотеза H_0 неверна, но ее принимают согласно критерию (*ошибка второго рода*).

Статистический критерий не дает логического доказательства или опровержения гипотезы. Так как гипотеза (H_0) проверяется по результатам выборочных наблюдений, то неизбежно имеет место некоторая, хотя бы и очень малая, вероятность ошибочного решения как первого, так и второго рода. При неограниченном увеличении объема выборки и использовании обоснованного критерия возможно добиться как угодно малых вероятностей обеих ошибок. Однако наиболее часто в практических задачах контроля имеют дело с фиксированным объемом выборки, когда задаются вероятностью ошибки первого рода — так называемым *уровнем значимости* α.

Выбор уровня значимости α зависит от сопоставления потерь, связанных с ошибками первого и второго рода. Следует учитывать, что уменьшение α приводит к росту вероятности ошибки второго рода. Поскольку в большинстве практических задач определить величину потерь оказывается весьма затруднительно, то, как правило, пользуются некоторыми стандартными уровнями значимости. К таким значениям можно отнести: $\alpha = 0{,}1$; $0{,}05$; $0{,}01$; $0{,}005$; $0{,}001$. Наиболее часто используют значение $\alpha = 0{,}05$, которое означает, что в среднем в 5 случаях из ста мы можем допустить ошибку и отвергнуть гипотезу H_0, в то время как она справедлива.