
ФЕДЕРАЛЬНОЕ АГЕНТСТВО
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ



РЕКОМЕНДАЦИИ
ПО СТАНДАРТИЗАЦИИ

**Р 50.1.083—
2012**

Статистические методы
ПРИМЕРЫ ПРИМЕНЕНИЯ
Часть 5
Анализ изменчивости данных

Издание официальное



Москва
Стандартинформ
2013

Предисловие

1 РАЗРАБОТАНЫ Автономной некоммерческой организацией «Научно-исследовательский центр контроля и диагностики технических систем» (АНО «НИЦ КД»)

2 ВНЕСЕНЫ Техническим комитетом по стандартизации ТК 125 «Статистические методы в управлении качеством продукции»

3 УТВЕРЖДЕНЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 29 ноября 2012 г. № 1283-ст

4 ВВЕДЕНЫ ВПЕРВЫЕ

Информация об изменениях к настоящим рекомендациям публикуется в ежегодном указателе «Руководящие документы, рекомендации и правила», а текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящих рекомендаций соответствующее уведомление будет опубликовано в ежемесячном информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет

© Стандартинформ, 2013

Настоящие рекомендации не могут быть полностью или частично воспроизведены, тиражированы и распространены в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения	1
2 Нормативные ссылки	1
3 Термины и определения	2
4 Статистический контроль	2
5 Отбор выборки	6

Введение

Серия рекомендаций по стандартизации «Статистические методы. Примеры применения» включает пояснения к применению статистических методов в простой и доступной форме.

В рекомендациях данной серии представлены способы применения статистических методов и приемов, показано, как знание процесса производства может способствовать его улучшению и повышению качества изготавливаемой продукции.

Знание изменчивости исследуемой характеристики чрезвычайно важно для принятия решений по результатам контроля или наблюдений за параметрами процесса. При этом выводы об изменчивости наблюдаемой характеристики зависят от соотношения объема выборки и объема всей совокупности.

Совокупностью является весь набор исследуемых объектов. Партией является часть совокупности, состоящая из объектов совокупности.

Наиболее полную информацию об изменчивости наблюдаемой характеристики позволяет получить сплошной контроль. Однако во многих случаях он неприемлем из-за больших финансовых и временных затрат. Поэтому из экономических соображений мониторинг изменчивости характеристики продукции обычно проводят по ограниченному количеству наблюдений или единиц, т.е. по выборке.

В настоящих рекомендациях рассмотрено несколько типовых ситуаций исследования изменчивости наблюдаемой характеристики по выборочным данным и приведены методы их наглядного представления.

Статистические методы

ПРИМЕРЫ ПРИМЕНЕНИЯ

Часть 5

Анализ изменчивости данных

Statistical methods. Examples of application. Part 5. Variation analysis

Дата введения — 2013—12—01

1 Область применения

В серии рекомендаций «Статистические методы. Примеры применения» приведены пояснения к использованию статистических методов, применяемых в менеджменте, контроле и улучшении процессов. В настоящих рекомендациях рассмотрены простые способы анализа изменчивости данных, приведены примеры типовых случаев анализа изменчивости данных и отбора выборки.

2 Нормативные ссылки

В настоящих рекомендациях использованы нормативные ссылки на следующие документы:

ГОСТ Р 50779.10—2000 Статистические методы. Вероятность и основы статистики. Термины и определения

ГОСТ Р 50779.11—2000 Статистические методы. Статистическое управление качеством. Термины и определения

ГОСТ ISO 9000—2011 Системы менеджмента качества. Основные положения и словарь

ГОСТ Р ИСО 11648-1—2009 Статистические методы. Выборочный контроль нештучной продукции. Часть 1. Общие принципы

ГОСТ Р ИСО 11648-2—2009 Статистические методы. Выборочный контроль нештучной продукции. Часть 2. Отбор выборки сыпучих материалов

ГОСТ Р ИСО/ТО 10017—2005 Статистические методы. Руководство по применению в соответствии с ГОСТ Р ИСО 9001

Р 50.1.040—2002 Статистические методы. Планирование экспериментов. Термины и определения

П р и м е ч а н и е — При пользовании настоящими рекомендациями целесообразно проверить действие ссылочных документов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный документ, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого документа с учетом всех внесенных в данную версию изменений. Если заменен ссылочный документ, на который дана датированная ссылка, то рекомендуется использовать версию этого документа с указанным выше годом утверждения (принятия). Если после утверждения настоящих рекомендаций в ссылочный документ, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный документ отменен без замены, то положение, в котором дана ссылка на него, рекомендуется принять в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящих рекомендациях применены термины по ГОСТ Р 50779.10, ГОСТ Р 50779.11, Р 50.1.040 и ГОСТ ISO 9000.

4 Статистический контроль

4.1 Общие положения

Контроль в режиме реального времени должен обеспечивать эффективное управление процессом. Из экономических соображений мониторинг изменчивости характеристики продукции обычно проводят по ограниченному количеству наблюдений или единиц, т.е. по выборке.

Важно исследовать изменения между единицами продукции или наблюдениями в выборке, учитывая соотношение объема выборки к общему количеству исследуемых объектов. Это относится к параметрам процесса, например, таким как температура розлива металла, содержание химического элемента в материале или геометрические размеры деталей. Во всех случаях после отбора выборки оценки, относящиеся ко всей совокупности исследуемых объектов, можно получить только после статистической обработки выборочных данных.

Совокупностью является весь набор исследуемых объектов. Партией является часть совокупности.

На рисунке 1 представлены результаты проверок, проведенных на исследуемых объектах, отбираемых в процессе производства равномерно во времени, по пять одновременно. Данные могут представлять собой, например, объем электролита в емкости или минимальную температуру выполнения операции на определенном оборудовании и т. п. Результаты проверки каждой выборки из пяти элементов показаны на рисунке 1 в виде точек, соответствующих результатам измерений. Представлены шесть случаев, пронумерованных на рисунке 1 от 1 до 6, в каждом из которых показаны 12 выборок из последовательных партий. Ниже точек, представляющих данные по каждому случаю, приведены графики, изображающие плотность теоретического распределения контролируемой характеристики, соответствующие данным всех выборок из 12 партий.

Случай 1. Между образцами 12 выборок существуют значительные различия. Однако имеется определенная стабильность изменений от выборки к выборке, в отличие от случая 5. В случае 1 нет признаков неустойчивости процесса во времени и можно предположить, что выборки принадлежат одной партии.

Если изменчивость является устойчивой, то:

- а) по результатам измерений характеристики элементов случайной выборки можно оценить функцию распределения характеристики процесса;
- б) если функция распределения известна (на основе данных экспериментов или эксплуатации), можно спрогнозировать изменчивость от одной случайной выборки к другой.

Необходимо отметить, что ситуации а) и б) являются противоположными.

При формировании оценки, описанной в а), больший объем выборки позволяет получить более достоверную оценку. Например, в случае 1 оценки среднего и стандартного отклонения по 12 выборкам объема 5 являются более точными, чем по единственной выборке объема 5. От достоверности полученной оценки зависят принимаемые решения на основе полученного распределения. Примером возможных задач является построение предикционных и толерантных интервалов.

Пример 1 — (Ситуация, соответствующая перечислению б)). Для безопасности и удобства транспортировки автомобильные аккумуляторы поставляют сухозаряженными вместе с электролитом в пластиковых емкостях. При недостатке электролита рабочие пластины аккумулятора будут не полностью покрыты, при избытке электролита он может вылиться из аккумулятора. Из практического опыта известно, что объем электролита, заливаемого в емкость, изменяется в соответствии с нормальным распределением со средним 729,0 мл и стандартным отклонением 2,0 мл соответственно. Статистическая теория позволяет сделать приведенные ниже утверждения.

1) Вероятность того, что средний объем электролита в емкости в случайной выборке из шести емкостей будет менее 726,5 мл, составляет 0,0011, т. е. такой результат в среднем может быть получен в одной из 900 выборок.

2) Вероятность того, что хотя бы одна емкость в выборке из шести емкостей будет содержать менее 726,5 мл электролита, составляет 0,4883. Это событие приблизительно в 400 раз более вероятно, чем предыдущий результат.

Если характеристика распределения соответствует случаю 1 (рисунок 1), процесс является стабильным или статистически управляемым. В этом случае методы статистической теории позволяют решать задачи прогнозирования, упомянутые в перечислениях а) и б).

Управляемое состояние процесса показано на рисунке 1 (случай 1).

4.2 Неустойчивая изменчивость

Изменчивость в случае 1 является устойчивой, а изменчивость в случае 5 является наиболее неопределенной. Процесс при отсутствии неслучайной причины, из которого отобраны выборки 4 и 11 в случае 5, находится в статистически управляемом состоянии. Точечный график показывает, что могут быть два источника изменчивости, которые воздействуют поодиночке или одновременно. Плотность распределения, построенная по данным 12 партий, демонстрирует недостаточную однородность. Изменчивость в случае 5 скорее всего представляет собой выход из статистически управляемого состояния. Без дальнейшего изучения факторов, влияющих на изменчивость от партии к партии, прогнозирование изменчивости для продукции данного процесса невозможно.

4.3 Систематическая изменчивость

Другой ситуации соответствует случай 3, где выборки 3, 7 и 11 отличаются от остальных. Здесь присутствует систематическое повторение данных, которого не было в случае 5. Возможно наличие двух причин изменчивости. Если это так, желательно устранить одну из них. Однако если это невозможно, то всю серию партий можно разделить на две гомогенные группы, в пределах которых процесс является устойчивым. Следовательно, к каждой группе могут быть применены статистические методы.

Пример 2 — Пусть в случае 3 выборки 1, 5, 9 представляют собой части партии продукции, изготовленной процессом № 1, выборки 2, 6, 10 — процессом № 2, выборки 3, 7, 11 — процессом № 3, выборки 4, 8, 12 — процессом № 4. Если партии сформированы путем объединения продукции, изготовленной каждым процессом, то выборка объема 20, представляющая собой случай 4, является репрезентативной для производства в целом. Этот пример помогает понять различие между простой случайной выборкой и репрезентативной выборкой. В случае репрезентативной выборки вся продукция разделена на гомогенные (однородные) части, из которых отобрана случайная выборка с объемом, пропорциональным объему данной части. Эти подвыборки затем объединяют в репрезентативную выборку. Таким образом, в описанном примере, четверть выборки объема 20 (пять единиц продукции) отобраны из партии, изготовленной каждым из четырех процессов. При отборе простой случайной выборки все возможные выборки заданного объема имеют одинаковую вероятность отбора. Таким образом, при отборе простой случайной выборки объема 20 в выборку может вообще не попасть продукция какого-то процесса (например, процесса № 3) или наоборот, половину или большую часть выборки будет составлять продукция, изготовленная процессом № 3. Очевидно, что такие выборки нежелательны для рассматриваемого примера.

Пример 3 — Рассмотрим изготовление латунных листов, толщина которых на краях меньше чем в центре вследствие особенностей процесса производства. Если листы разрезают на более узкие полосы, то толщина металла в полосе зависит от места, из которого вырезана полоса. Если такие изменения выявлены, продукцию следует разделить на однородные части в соответствии с целями отбора выборки. Однако, если такие изменения не выявлены, то возможно, что для испытаний необходимо отобрать выборки, состоящие из образцов, взятых только из середины листа или только с краев, или из разных мест листа. В этой ситуации имеет место случай 5, не позволяющий принимать достоверные решения по результатам измерений выборочных единиц.

Различия между случаем 1 и случаем 3 очень важны. В первом случае любая выборка может быть использована для получения информации обо всей продукции производственного процесса. В случае 3 необходимо обеспечить получение репрезентативной выборки. В случае 1 изменчивость внутри выборки и между выборками из отдельных партий не отличается от того, что, возможно, ожидалось, если бы серия случайных выборок была отобрана из партии, сформированной объединением и перемешиванием продукции из отдельных партий. В другом случае ситуация аналогична случаю, когда выборочные единицы отобраны из нескольких различных источников и доля единиц, отобранных из каждого источника, пропорциональна количеству единиц продукции, поступивших из этого источника. При этом отбор единиц продукции, поступивших из конкретного источника, является случайным.

Термин «расслоенная выборка» часто неправильно используют в качестве синонима термину «репрезентативная выборка». Фактически репрезентативная выборка является простым частным случаем расслоенной выборки. Как и репрезентативную выборку, расслоенную выборку используют в ситу-

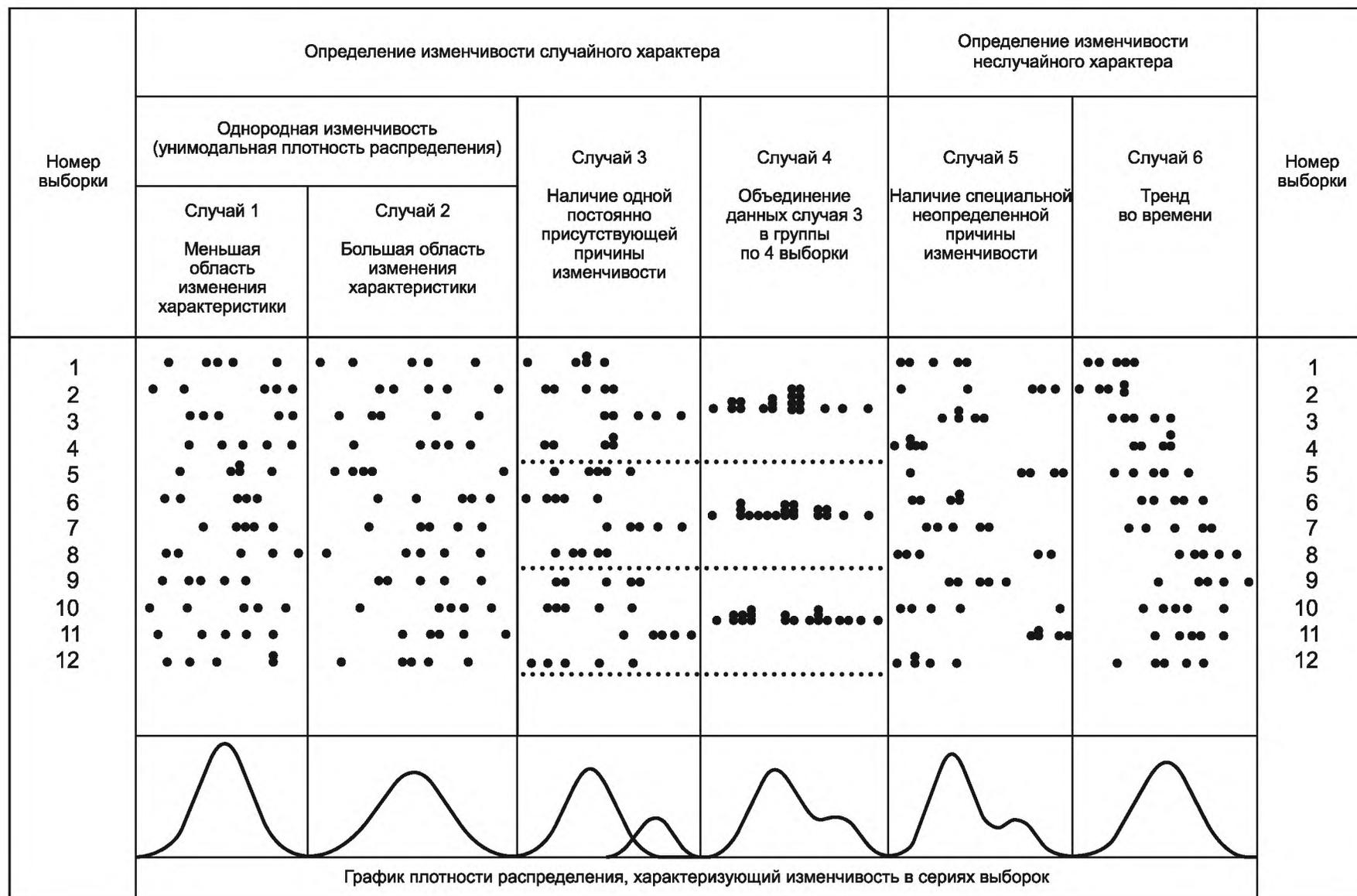


Рисунок 1 — Типы изменчивости в выборках

ации, когда продукция может быть разделена на относительно гомогенные слои (страты). На этом сходство заканчивается. При отборе выборки из каждой страты (каждого слоя) для расслоенной выборки учитывают стоимость отбора выборочной единицы из каждой страты и полученные ранее оценки изменчивости в пределах каждой страты. Отбор делают с целью уменьшения затрат на достижение заданной точности оценки среднего исследуемой характеристики или достижения максимальной точности в пределах заданных затрат. Поэтому отбор репрезентативной выборки представляет собой частный случай расслоенной выборки, при котором изменчивость и стоимость отбора выборочной единицы одинаковы для всех страт.

4.4 Систематические изменения во времени

Случай 6 представляет собой ситуацию, для которой характерны систематические изменения качества продукции во времени. Если причины этих отклонений найдены, то, если причины этих отклонений не могут быть устранены, можно разделить изготовленную процессом продукцию на потоки со стабильным качеством. Примерами могут быть изменения, вызванные колебаниями температуры или влажности (например, в текстильном производстве) или индивидуальными особенностями операторов. Однако если причины изменчивости во времени не известны, то прогноз изменений процесса на основе результатов измерений в нескольких партиях невозможен. Таким образом, выборки 1 и 2 или 8 и 9 являются репрезентативными для определения распределения процесса.

4.5 Статистическая неопределенность

Если общий объем продукции получен из нескольких источников, каждый источник находится в состоянии статистической управляемости и соотношение количества единиц продукции, поступивших из каждого источника, известно, то для оценки качества продукции в целом на основе отобранных выборок могут быть использованы статистические методы. Если недостаточно информации для правильного отбора репрезентативной выборки, определить изменчивость статистическими методами невозможно. Это может быть вызвано особенностями применяемого в производстве оборудования при наличии нескольких процессов или поставщиков, или изменений во времени, вызванных сезонными колебаниями качества сырья.

4.6 Другие (ненормальные) распределения

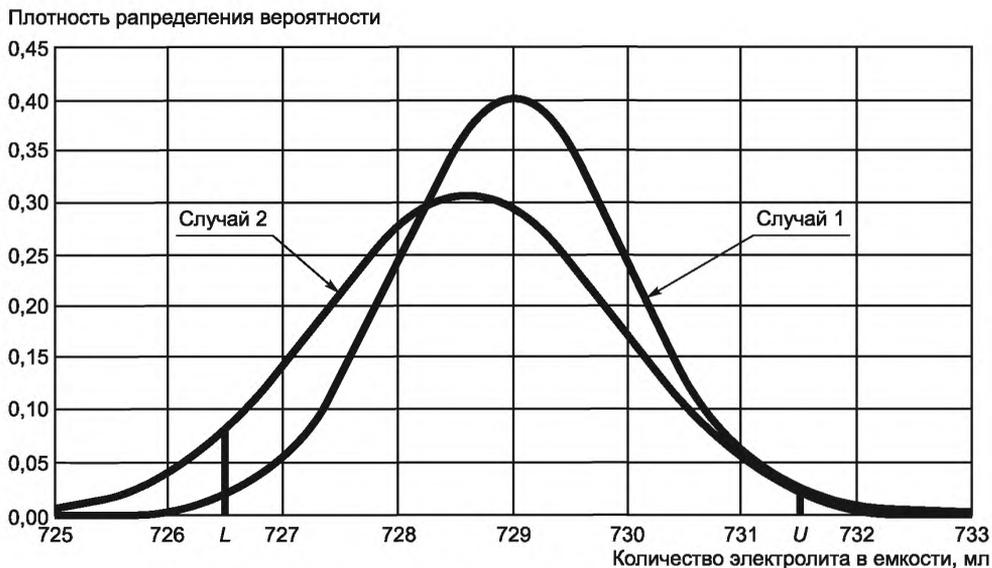
Необходимо помнить, что изменчивость процесса, находящегося в состоянии статистической управляемости не всегда может быть описана нормальным распределением. В большом количестве случаев, связанных с производством продукции, часто используют нормальное распределение. Однако в большинстве практических ситуаций кривая плотности распределения измеряемой характеристики не является симметричной, например, распределения срока службы и усилия разрыва, как правило, имеют длинный хвост справа. Но если распределение сохраняется от партии к партии, то процесс находится в статистически устойчивом состоянии.

4.7 Уровень качества и возможности процесса

Изменчивость в случае 1 и в случае 2 выглядит постоянной от партии к партии. Однако для подтверждения стабильности этого недостаточно. Постоянство дисперсии не указывает, на каком уровне качества продукции (высоком или низком) функционирует процесс. Чтобы оценить уровень качества продукции, необходима информация о среднем и изменчивости процесса. Источником такой информации являются выборочные оценки среднего и стандартного отклонения процесса. Эта информация позволяет обнаружить отклонения в работе процесса и часто позволяет за счет регулировки процесса поддерживать его на необходимом уровне. Для мониторинга процесса могут быть использованы контрольные карты.

Изменчивость процесса в случае 1 и в случае 2 представлена кривыми плотности нормального распределения, приведенными под точечными графиками на рисунке 1. Однако эти случаи различны. Случаю 2 соответствует большая изменчивость отдельных единиц в выборке, чем в случае 1. Нормальное распределение полностью определяется средним μ и стандартным отклонением σ . Если μ_1 и σ_1 — среднее и стандартное отклонение процесса для случая 1, а μ_2 и σ_2 — для случая 2, то различие между случаем 1 и случаем 2 состоит в том, что $\sigma_2 > \sigma_1$.

Если случай 1 и случай 2 представляют изменчивость объема электролита в емкостях, наполняемых двумя установками с $\sigma_1 = 1,0$ мл, $\sigma_2 = 1,3$ мл, $\mu_1 = 729,0$ мл и $\mu_2 = 728,6$ мл (это невозможно увидеть на рисунке, поскольку масштаб не указан), то в обоих случаях, статистическая теория позволяет спрогнозировать долю емкостей, количество электролита в которых лежит в пределах заданных границ. Предположим, что требования имеют вид $L = 726,5$ мл и $U = 731,5$ мл. Эта ситуация показана на рисунке 2.



Примечание — Плотности нормальных распределений с параметрами $\mu_1 = 729,0$; $\sigma_1 = 1,0$ и $\mu_2 = 728,6$; $\sigma_2 = 1,3$.

Рисунок 2 — Сопоставление данных двух установок по розливу электролита для аккумуляторной батареи

Очевидно, что возможности установок по заполнению емкостей различны. При этом заполнение емкостей первой установкой является более однородным и более приемлемым. Можно заметить, что доля емкостей, содержащих количество электролита менее нижней границы, для случая 2 является во много раз больше, чем для случая 1. Поэтому в случае 1 возможности процесса выше, т. е. качество продукции этого процесса выше, чем в случае 2.

Для описания изменчивости процесса, когда кривая его плотности распределения не меняется от выборки к выборке, используют термины «стабильный», «статистически управляемый» и «устойчивый». Заключение о стабильности процесса в значительной степени зависит от метода отбора выборки. Например, в случае непрерывного производства отбор выборки через короткие интервалы времени позволяет выявить нестабильность, которую при отборе выборок через более продолжительные интервалы времени обнаружить невозможно.

5 Отбор выборки

В некоторых случаях, когда поставщик уверен, что его процесс производства находится в статистически управляемом состоянии, он может принять простую процедуру отбора выборки, отбирая в выборку каждую 500-ю или 1000-ю единицу продукции. Слабым местом такой процедуры является временной интервал между отбором выборочных единиц, продолжительность которого может совпасть с периодом колебаний качества продукции, если он существует. В таком случае необходимо изменить периодичность отбора единиц в выборку, иначе сделанные выводы могут быть ошибочными. Примерами возможных причин изменчивости качества продукции являются: изменения температуры производственной среды, усталость или невнимательность операторов в конце смены, или периодические пополнения сырья для изготовления продукции, поступающего от разных поставщиков.

Часто проблема намного сложнее. Это обычно имеет место, когда отбор выборки выполняют не только для определения приемлемости партии, но и для определения сорта и цены продукции перед ее поставкой потребителю. Поскольку часто нет доказательств того, что уровень качества поставок от конкретного поставщика остается постоянным, важно предусмотреть метод отбора выборки, который обеспечивает получение надежной оценки качества продукции для каждой партии, даже если она неоднородна, т. е. необходимо обеспечить отбор репрезентативной выборки из каждой партии.

Пример 4 — Отбор случайной выборки керамических плиток выполняют, выбирая плитки из каждой партии так, чтобы у каждой плитки в партии был равный шанс быть отобранной в выборку. На отобранных плитках необходимо указывать партию, из которой они отобраны.

Если существует возможность, что плитки содержат вредные включения, например, в виде кальция или пирита, подготовка образцов для испытаний должна быть изменена так, чтобы обеспечить наличие в выборке плиток, содержащих включения, для формирования репрезентативной выборки.

Если материал не состоит из дискретных единиц и не является гомогенным, при отборе выборки возникают проблемы. В этом случае необходимо отбирать маленькие равные партии материала из большого количества различных частей общей массы. Если материал движется по конвейеру, равные части его могут быть отобраны через равные интервалы времени непосредственно с конвейера. Обычно на практике объединяют эти части и формируют начальные пробы, которые затем смешивают, уменьшают и формируют пробы материала, подходящие для анализа в лаборатории. Целью этих действий является получение выборки, которая является настолько репрезентативной, насколько возможно.

Основное требование правильного отбора выборки состоит в том, чтобы обеспечить всем частям нештучного материала в партии равные возможности быть отобранными, стать частью выборки и затем попасть в объединенную выборку для анализа. Любое отклонение от этого требования приводит к потере достоверности полученных оценок. Неверная схема отбора выборки не обеспечивает получение репрезентативной выборки.

Лучшим местом отбора проб, удовлетворяющим вышеупомянутому требованию, является место, где нештучный материал ссыпается с ленточного конвейера. Здесь может быть отобрано полное поперечное сечение транспортируемого материала, что позволяет получить репрезентативную выборку.

Пробы для оценки влажности должны быть отобраны и взвешены незамедлительно. Если это невозможно, то пробы должны быть сохранены в воздухонепроницаемых контейнерах с минимумом свободного воздушного пространства, чтобы минимизировать любое изменение влажности материала, и должны быть незамедлительно подготовлены для анализа.

Минимизация смещения, связанного с отбором выборки и подготовкой образцов очень важна. В отличие от точности полученных оценок, которая может быть повышена за счет увеличения количества измерений или повторения измерений, это смещение не может быть уменьшено таким образом. Следовательно, минимизация или устранение возможных смещений являются крайне важной задачей. Источниками смещения, которое может быть полностью устранено за счет правильного отбора выборки и системы подготовки образцов, могут быть частичная утрата выборки, загрязнение и неправильное извлечение образцов. К источникам смещений, которые могут быть минимизированы, но не полностью устранены, относятся изменение влажности материала, потеря пыли и разрушение частиц (при определении размера частиц).

Если из каждой партии отобрана только одна конечная выборка для лабораторного анализа, нет возможности оценить достоверность оценок исследуемых характеристик. Эта проблема может быть решена при помощи двух выборок. Выборки составляют путем последовательного размещения первичных проб поочередно в два контейнера, где первичная проба — количество материала, отобранного с помощью одного действия устройства, отбирающего пробу. Возможны четыре ситуации.

а) Если партии поставляют часто, изменения качества продукции могут быть определены по большому количеству партий приблизительно равной массы, на основе исследования каждой партии отдельно и составлением двух выборок для каждой партии.

б) Если большие партии поставляют не часто, изменения качества продукции могут быть определены по единственной партии делением ее, по крайней мере, на 10 частей приблизительно равной массы и составлением двух выборок для каждой партии.

в) Если часто поставляют небольшие партии, изменения качества продукции могут быть определены по нескольким партиям приблизительно равной массы делением всех партий, по крайней мере, на 10 партий приблизительно равной массы и составлением двух выборок для каждой партии.

г) При отборе выборки из партии, поставляемой в фургонах, когда первичные пробы отбирают из всех фургонов, изменения качества продукции могут быть определены по каждой партии отдельно составлением двух выборок для каждой партии.

Ключевые слова: статистический контроль, неустойчивая изменчивость, систематическая изменчивость, статистическая неопределенность, возможности процесса, отбор выборки

Редактор *С.Д. Золотова*
Технический редактор *В.Н. Прусакова*
Корректор *М.М. Малахова*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 03.10.2013. Подписано в печать 17.10.2013. Формат 60 × 84 $\frac{1}{8}$. Гарнитура Ариал.
Усл. печ. л. 1,40. Уч.-изд. л. 1,10. Тираж 124 экз. Зак. 1168.

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.
www.gostinfo.ru info@gostinfo.ru
Набрано во ФГУП «СТАНДАРТИНФОРМ» на ПЭВМ.
Отпечатано в филиале ФГУП «СТАНДАРТИНФОРМ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.