

ПРИМЕНЕНИЕ ОБОБЩЕННЫХ РАСПРЕДЕЛЕНИЙ ДЛЯ АНАЛИЗА ТЕХНОЛОГИЧЕСКИХ ПОГРЕШНОСТЕЙ И ПОЛУЧАЕМАЯ ПРИ ЭТОМ ЭФФЕКТИВНОСТЬ

Базой статистических методов контроля качества продукции, статистического анализа и регулирования технологических процессов (ТП) являются теория вероятностей и математическая статистика.

Поскольку все параметры производимой продукции (например, размеры, прочность и др.) имеют некоторый разброс (рассеяние) относительно средних значений, то система управления качеством должна в обязательном порядке содержать подсистему статистического анализа и регулирования ТП. Именно эта подсистема может дать реальное повышение качества продукции, правда при условии, если она будет построена на базе распределений, которые наиболее точно описывают технологические погрешности.

Если статистическое распределение отличается от предполагаемого теоретического, то это может привести к непредсказуемым последствиям.

Чтобы постоянно поддерживать ТП в управляемом состоянии, вводится статистическое регулирование ТП, например, с помощью контрольных карт. Это позволяет отслеживать показатели точности, стабильности и самый главный показатель - ожидаемый уровень брака, и удерживать их в заданных пределах.

Перед тем как внедрять статистическое регулирование ТП, необходимо провести его статистический анализ при установившемся состоянии. Целью статистического анализа является:

- установление закона распределения контролируемого параметра;
- анализ точности и стабильности ТП;
- определение значений показателей точности, стабильности ТП и возможного уровня брака при заданном конструкторском допуске.

Главным пунктом статистического анализа справедливо считается правильное установление закона распределения контролируемого параметра, поскольку он является наиболее полной характеристикой случайной величины и позволяет вычислять все необходимые показатели.

Обычно в качестве выравнивающего закона распределения технологических погрешностей принимают нормальный закон. Но всегда ли нормальный закон является наилучшим для выравнивания статистических распределений? Опыт показывает, что далеко не всегда.

Отсюда следует, что все показатели, и особенно главный из них - ожидаемый уровень брака, рассчитанные в предположении нормального закона, могут оказаться далеки от истины. Так, в случае асимметричного статистического распределения нормальный закон может дать погрешность в определении ожидаемого уровня брака до 10 раз в большую или меньшую сторону.

Следовательно, главная задача, которую необходимо всякий раз решать при статистическом анализе ТП, - это вопрос установления такого закона распределения, который наиболее точно описывает (выравнивает) статистический ряд распределения, ибо только на его основе можно с высокой точностью выполнять необходимые расчеты и давать прогнозы.

Спрашивается, а как найти это наилучшее распределение?

Вот здесь перед специалистом по качеству возникают непреодолимые барьеры, поскольку известные распределения не могут с достаточной точностью описать все многообразие статистических распределений.

В учебниках по теории вероятностей и математической статистике для решения этой задачи рекомендуется выдвигать гипотезы о выравнивающем законе распределения и проверять

каждую из них по критериям согласия в надежде на то, что подходящий закон в конце концов будет установлен. К сожалению, как показывает опыт, таким способом практически никогда нельзя подобрать наилучшее выравнивающее распределение, а в итоге мы не получаем ожидаемого экономического эффекта.

Поэтому не удивительно, что статистические методы анализа и регулирования ТП на многих предприятиях используются не по существу, а лишь формально и в результате не оказывают существенного влияния на повышение качества продукции. Отсюда - нежелание руководителей предприятий тратить необходимые средства на разработку или приобретение системы статистического анализа и регулирования ТП. А ведь она составляет одно из важнейших звеньев системы управления качеством продукции.

Этому в немалой степени способствует сложившаяся в прикладной статистике ситуация. А она сравнима с той, которая была в химии накануне создания "Периодической системы элементов" Д.И.Менделеева - было известно достаточно много элементов, в некоторой степени изучены их свойства, но еще не было главного - "Периодического Закона", позволявшего предсказывать существование еще не открытых элементов и их свойства.

Следовательно, по аналогии с химией главная задача прикладной статистики на данном этапе, без решения которой невозможно разработать эффективную систему статанализа ТП, - это создание своего рода "Периодической системы распределений", исследование их свойств, разработка критериев для установления закона распределения случайной величины по статистическому распределению, наконец, разработка новых методов оценивания параметров, что представляет собой весьма сложную задачу.

Проблема построения универсальных вероятностных моделей для выравнивания широкого класса статистических распределений обсуждается не впервые.

Уже в 1895г. английский статистик К.Пирсон предложил свое семейство непрерывных распределений, заданное в виде дифференциального уравнения. Это семейство распределений он получил путем выравнивания дискретного гипергеометрического распределения. Им же был предложен метод моментов для нахождения оценок параметров выравнивающих (теоретических) распределений (назовем его классическим методом моментов).

Существенным недостатком семейства распределений К.Пирсона по нашему мнению является отсутствие обобщенной плотности, представленной в явном виде, что сильно ограничивает возможности его использования на практике. Кроме того, метод моментов не позволяет находить оценки параметров тех распределений, в том числе принадлежащих семейству К.Пирсона, которые не имеют моментов высших порядков (3-го или 4-го). Поэтому разработка системы непрерывных распределений, более широкой, чем семейство кривых К.Пирсона, а также новых методов оценивания параметров имела бы большое значение как в теоретических, так и прикладных исследованиях, в том числе для оценки состояния технологических процессов.

Работы в этом направлении проводились в течение длительного времени. В 1912г. английский статистик Р.Фишер предложил новый метод оценивания параметров практически любых распределений - метод наибольшего правдоподобия. Однако и этот метод имеет серьезные недостатки и главный из них - отсутствие критериев для установления типа выравнивающего распределения.

За последние 100 лет с года опубликования К.Пирсоном своего семейства распределений предпринимались неоднократные попытки разработать другие семейства распределений, но несмотря на огромные затраты труда (не менее тысячи человеко-лет!), более удачных решений предложено не было. Тогда некоторые ученые пришли к выводу, что построение обобщенных распределений невозможно и продолжали интенсивно развивать непараметрические методы.

Анализ показывает, что основными причинами, которые задержали создание теории универсальных (обобщенных) распределений на многие десятки лет, являются следующие.

Первая причина - это уверенность в невозможности построения универсальных распределений.

Вторая причина - устоявшееся мнение, что совокупностью известных распределений, включая семейство кривых К.Пирсона, можно описать все или почти все статистические распределения.

Но главной причиной, не позволившей разработать теорию универсальных распределений, является отсутствие общего подхода, т.е. предпринимались попытки решить одну из частных задач, например, выравнивание дискретных распределений, вместо решения общей задачи, которую можно сформулировать так: построить систему непрерывных распределений, включающую как частные случаи широко известные классические непрерывные распределения, в том числе семейство кривых К.Пирсона или по крайней мере значительную его часть.

Переформулировка задачи по выравниванию отдельных дискретных распределений в более общую задачу - построить систему непрерывных распределений - дает возможность использовать любые известные распределения с целью их обобщения. При этом обобщенному распределению будут присущи все свойства отдельных распределений.

Мой многолетний опыт по обработке статистических распределений различных случайных величин в областях строительства, техники, биологии, информатики, математической лингвистики и др. позволяет сделать следующие выводы:

- никакое одно семейство непрерывных распределений не может с достаточной точностью описать все многообразие статистических распределений. Для этого необходимо иметь несколько семейств распределений, причем, более широких, чем семейство кривых К.Пирсона;

- существующие критерии для установления типа выравнивающей кривой не гарантируют точности решения этой задачи;

- известные методы оценивания параметров далеки от совершенства и оказываются практически непригодными для некоторых вновь построенных семейств распределений.

ТРИ СИСТЕМЫ НЕПРЕРЫВНЫХ РАСПРЕДЕЛЕНИЙ

В 1975г. мною было получено обобщенное распределение, заданное четырехпараметрической плотностью вида

$$p(t) = Nt^{\gamma-1} (1 - \alpha t^{\beta})^{\frac{1}{u}-1}, |$$

которое включает как частные случаи множество известных распределений, в том числе почти все семейство кривых К.Пирсона, и может претендовать на роль универсального закона распределения в математической статистике.

На базе этого обобщенного распределения были получены другие распределения и в итоге построены три системы непрерывных распределений.

Наряду с системами непрерывных распределений построена система дискретных распределений, включающая как частные случаи известные законы: биномиальный, Пуассона, отрицательный биномиальный, логарифмический и др., а также взаимосвязанная с ней система кривых роста новых событий.

Каждая система непрерывных распределений задана тремя обобщенными четырехпараметрическими плотностями, обладающими определенными свойствами и поэтому пригодными для описания своего круга статистических распределений.

Первая система непрерывных распределений содержит так называемый параметр сдвига,

благодаря которому кривая распределения может перемещаться по горизонтальной оси без изменения своей формы, меняя лишь значения параметра сдвига. Она включает нормальный закон как частный случай.

Первая система непрерывных распределений используется для выравнивания и прогнозирования статистических распределений таких случайных величин, последующие значения которых образуются из предыдущих путем их изменения (сдвига) на некоторую постоянную величину. При этом форма кривой распределения не изменяется. Но изменяются среднее и коэффициент вариации.

Средние значения таких случайных величин растут во времени по линейному закону, но могут иметь и другой закон роста.

Первой системой непрерывных распределений описывается, например, статистическое распределение работников по возрасту, распределение образцов бетона по прочности, а также статистические распределения производственных погрешностей.

Вторая система непрерывных распределений может быть получена на базе первой системы и также задается тремя обобщенными плотностями. Она содержит как частный случай логарифмически нормальный закон.

Эта система непрерывных распределений используется для выравнивания и прогнозирования распределений таких неотрицательных случайных величин, последующие значения которых образуются из предыдущих путем их умножения на некоторую постоянную величину. При этом форма кривой распределения изменяется, но неизменным остается коэффициент вариации.

Средние значения логарифмов таких случайных величин изменяются во времени по линейному закону, а сами случайные величины - по показательному закону.

Второй системой непрерывных распределений описывается, например, статистическое распределение работающих по размерам заработной платы, результаты испытаний на надежность и другие ряды распределения.

Аналогично получается третья система непрерывных распределений (из первой или второй системы).

При использовании обобщенных (универсальных) распределений задача установления типа выравнивающей кривой по статистическому ряду распределения значительно упрощается, поскольку не требуется выдвижения множества гипотез о выравнивающем распределении и проверки каждой из них по критериям согласия. Но чтобы это преимущество обобщенных распределений использовать практически, необходимо разработать критерии, которые позволяли бы однозначно устанавливать тип выравнивающей кривой, а также создать совершенно новые методы оценивания параметров, поскольку известные методы в данном случае оказываются мало пригодными.

Следовательно, построение систем непрерывных распределений - это лишь начало большой работы, включающей их исследование, классификацию, разработку новых общих критериев для установления типа выравнивающей кривой, разработку совершенно новых методов оценивания параметров и, наконец, доведение полученных результатов до программной реализации.

Без решения всех перечисленных задач невозможно успешное использование обобщенных распределений на практике.

В результате многолетних исследований мною были найдены критерии, которые зависят лишь от двух параметров формы, а также разработан общий для трех систем распределений метод оценивания параметров, причем этот метод оказался устойчивым, т.е. мало чувствительным к выбросам на концах статистического распределения.

С помощью двух критериев по заранее построенной номограмме устанавливается тип выравнивающей кривой и находятся оценки двух параметров формы. Оценки двух других параметров рассчитываются по простым формулам.

Для облегчения расчетов по каждой системе распределений автором были созданы программы под общим названием SNR (Системы непрерывных распределений). В течение ряда лет (с 1992г.) они обкатывались на разнообразном статистическом материале и совершенствовались.

Наличие обобщенных распределений и методов оценивания параметров, доведенных до программной реализации, значительно упрощает и ускоряет процедуру нахождения наилучшей выравнивающей кривой распределения, поскольку она вычисляется на основе статистических данных за один прием без перебора отдельных частных случаев. Этим достигается новый, значительно более высокий уровень точности решения прикладных задач, связанных с выравниванием статистических распределений производственных погрешностей.

Рассмотрим пример статистической обработки результатов замера контролируемого параметра по программе SNR1MM97. Были измерены отклонения от номинального размера некоторой детали диаметром 50мм. Нижняя и верхняя границы конструкторского допуска соответственно равны -0.012, 0.012 мм. Ширина конструкторского допуска равна 0.024мм. Объем выборки - 100 замеров.

Расчеты по программе показали, что уровень брака равен 2.77%, причем, на левой границе 0.06%, а на правой 2.71%. При этом коэффициент рассеяния (отношение ширины поля рассеяния при доверительной вероятности $P=0.9973$ к ширине конструкторского допуска) равен 1.06 (см. распечатку). В этом случае для полного исключения брака требуется уменьшить рассеяние. Однако только за счет смещения центров рассеяния и допуска можно добиться значительного уменьшения брака. Если сместить кривую распределения влево на 2мкм, то ожидаемый уровень брака снизится до 0.56%, причем он будет примерно одинаковым на обеих границах конструкторского допуска (по 0.28%). Для дальнейшего уменьшения брака необходимо уменьшить ширину поля рассеяния контролируемого параметра, что приведет к уменьшению коэффициента рассеяния.

В рассмотренном примере выравнивающим оказалось бэта-распределение (тип 1.1).

Если бы для описания технологических погрешностей был использован нормальный закон, то расчетный процент брака оказался бы равным 4.456%, т.е. в 1.6 раза выше фактического. И это при небольшой асимметрии выравнивающей кривой!

Обработка с помощью соответствующих программ реальных статистических данных, полученных на различных предприятиях Республики Беларусь, показала, что они хорошо выравниваются обобщенными распределениями, при этом ожидаемый уровень брака колеблется от нуля до нескольких десятков процентов, хотя предельно допустимый уровень брака в промышленности не должен превышать 0.27%.

Использование обобщенных распределений, алгоритмов, методов и программных средств может дать весьма значительный экономический эффект, а без их использования трудно наладить выпуск конкурентоспособной продукции.

За счет чего достигается экономический эффект при использовании обобщенных распределений?

1. Уменьшаются вероятности ошибок первого и второго рода, т.е. когда при налаженном процессе принимается решение о его разладке и наоборот, при разлаженном - что он налажен. Эти ошибки велики при неправильно установленном законе распределения технологических погрешностей, например, при постоянном использовании нормального закона.

2. Точно установленный закон распределения позволяет оценить:

- возможности производственного оборудования;
- эффективность управляющих воздействий;
- эффективность ремонта, наладки и т.д.;

- он обучает производственный персонал, дает ему в руки достоверную информацию для принятия правильных решений;

- дает объективную количественную характеристику качества труда (например, ожидаемый уровень брака) для каждого рабочего у станка, бригады, цеха и ее изменение во времени.

ПОКАЗАТЕЛИ СТАТИСТИЧЕСКОГО РАСПРЕДЕЛЕНИЯ (SNR1MM97)

К-во наблюдений	100
К-во интервалов	20
Ширина интервала	1
Среднее	3.92
Дисперсия	22.4736
Центр. момент 3-го порядка	-37.34583
Центр. момент 4-го порядка	1408.746
Ср. квадратич. отклонение	4.740633
Коэф-т вариации %	120.9345
Показатель асимметрии	.1228758
Показатель островершинности	2.78925
L1=	2.616711

РАСПРЕДЕЛЕНИЕ ТИПА 1.1 С ПАРАМЕТРАМИ

AU=2.696148E-02 K=8.403986 U=.2352945 L=-20.71284 N=1.367198E-10

Случайная величина X задана на интервале $-20.71284 < X < 16.37711$

$$P(X) = N * (X-L)^{(K-1)} * (1-AU*(X-L))^{(1/U-1)}$$

ПОКАЗАТЕЛИ УРОВНЯ КАЧЕСТВА (SNR1MM97)

Кол-во наблюдений	100
Среднее из опыта W1	3.92
Среднее квадратич. отклонение (с.к.о.) Sx	4.740633
Центр поля допуска T0	0
Нижняя граница поля допуска TN	-12
Верхняя граница поля допуска ТВ	12
Ширина поля допуска TX=ТВ-TN	24
Отношение допуска к с.к.о. (TX/SX)	5.062615
Смещение центров рассеян. и допуска E=W1-T0	3.92
Ширина поля рассеяния RX прм P=.9973	25.42721
Нижняя граница поля рассеяния	-11.01229
Верхняя граница поля рассеяния	14,41492
Отношение поля рассеяния к с.к.о. (RX/SX)	5.363674
Показатель уровня настройки KN=E/TX	.163333
Показатель точности KT=RX/TX	1.059467
Прогнозируемый процент брака Q%	2.76925
в т.ч. на нижней границе поля допуска	.060416
на верхней границе поля допуска	2.706635

3. Применение обобщенных распределений позволяет регулировать ТП с любым законом распределения технологических погрешностей.

4. Любые сложные расчеты легко выполняются по соответствующим программам с помощью ПЭВМ.

5. Имеется возможность получения качественной продукции при коэффициенте рассеяния, близком к единице.

6. Обобщенные распределения позволяют решать другие задачи, которые нельзя решить традиционными методами, например, прогнозирование статистической структуры выборки; нахождение законов распределения суммы N независимых случайных величин или среднего N случайных величин; вычисление числа N , при котором закон распределения суммы можно считать нормальным и т.д.

7. Обобщенные распределения позволяют широко использовать наиболее общую характеристику точности ТП - ожидаемый процент брака. С помощью программы нетрудно подобрать оптимальное смещение центров рассеяния и допуска, при котором ожидаемый процент брака будет минимальным.

8. Программы с не меньшим успехом могут быть использованы при статистическом регулировании ТП, а также при обработке результатов испытаний на надежность.

9. При использовании обобщенных распределений значительно повышается эффективность статистических методов и их роль в повышении качества продукции, возрастает доверие к ним со стороны руководителей предприятий.

Таким образом, обобщенные распределения, реализованные в программных средствах, являются мощным инструментом для статанализа и регулирования ТП, а также для решения многих других задач.

В заключение следует отметить, что многие проблемы, связанные с практическим использованием обобщенных распределений, еще ждут своего решения. На это потребуются много труда и времени.

Чтобы ускорить применение этих методов на практике, необходимо решить многие задачи, в том числе:

- организовать подготовку специалистов по современным методам статанализа на базе обобщенных распределений путем включения в вузовскую программу соответствующего спецкурса;

- включить этот спецкурс в программу постоянно действующих курсов повышения квалификации;

- довести имеющиеся программные средства до широкого практического использования на предприятиях РБ;

- организовать издание учебных пособий, методических рекомендаций, монографий по использованию обобщенных распределений в науке, технике, производстве;

- разработать новые или откорректировать существующие ГОСТы по статанализу и регулированию ТП с учетом использования обобщенных распределений и соответствующих программных средств.

Эти предложения были рассмотрены в Администрации Президента и одобрены ею.

Решение перечисленных задач важно не только с точки зрения повышения качества продукции, но и для утверждения приоритета Республики Беларусь в вопросах разработки современных методов статанализа и регулирования ТП на базе обобщенных распределений, ибо нигде в мире эти задачи до сих пор не решены, а у нас они уже доведены до программной реализации.