

УДК 65.012; 519,24

Г.М. Трищ

Украинская инженерно-педагогическая академия, Харьков

## УПРАВЛЕНИЕ КАЧЕСТВОМ МАШИНОСТРОИТЕЛЬНОЙ ПРОДУКЦИИ ПО МАЛЫМ ВЫБОРКАМ

*Рассмотрены возможности применения математического аппарата порядковых статистик для управления качеством в машиностроении по малому количеству статистической информации. Проведен статистический анализ полученных оценок для распределения Симпсона.*

**Ключевые слова:** эффективные оценки параметров, порядковые статистики, малая выборка.

### Введение

Для любого государства качество продукции и услуг национального производителя является приоритетным в международной конкурентной борьбе, поскольку существенно влияет на формирование внешней политики и национальной безопасности, определяет уровень жизни и стабильность национальной валюты.

Наиболее важным и определяющим в повышении конкурентоспособности продукции является машиностроение, которое служит основой материальной базы общества. Его развитию всегда придавалось и придается первостепенное значение. В настоящее время успешно внедряются сверхточная обработка материалов на основе высоких технологий, прецизионного оборудования и сверхточной измерительной аппаратуры. Прогнозируется дальнейшее повышение точности – свыше 0,001 мкм.

При управлении качеством технологических процессов в машиностроении используется информация, получаемая статистическими методами. Ее получение и обработка основываются на методологии параметрических статистик, рассчитанной на массовое и крупносерийное производство. Для современного производства, которое характеризуется, в основном, мелкосерийным выпуском изделий, необходимо создание методов работы со статистической информацией ограниченного объема и соответствующего нового нормативного обеспечения. То есть достоверную и в нужное время информацию необходимо получать при работе с малым объемом контролируемых изделий – малыми выборками. Для этого должны использоваться решения, основанные на научных принципах структурного построения, внутренней непротиворечивости и простоты, удобства в применении и легко адаптироваться к различным технологическим процессам. Это позволит обеспечить необходимую точность изготовления изделий малых партий с минимальными затратами.

**Цель статьи** – рассмотрение возможности применения математического аппарата порядковых

статистик для управления качеством в машиностроении по малому количеству статистической информации.

С развитием теории вероятностей и математической статистики статистические методы управления качеством распространились не только на методы анализа и оценки качества продукции, но и на контроль протекания технологических процессов и приемочный контроль качества продукции. Решение этих задач статистическими методами возможно двумя группами методов – методами параметрических и непараметрических статистик. При этом нельзя однозначно отдать предпочтение одному из групп методов при решении определенной задачи, так как они имеют свои достоинства и недостатки и зависят в основном от количества статистической информации. Непараметрические статистики не требуют знания закона распределения случайной величины, а используют только выборочные значения из партий деталей (генеральной совокупности). Недостаток их в том, что они требуют большого объема статистических данных (большой выборки), но основным преимуществом является их способность решать практические задачи, не зная закона распределения случайных величин показателей качества. Параметрические статистики для решения практических задач управления качеством не всегда требуют большого количества статистической информации, за счет предположения о знании закона распределения случайных величин показателей качества.

Поэтому в зависимости от двух факторов: количества статистической информации (выборка действительных значений показателей качества из генеральной совокупности) и знания закона распределения их распределения определяется группа методов для эффективного решения практических задач управления качеством.

На практике очень часто приходится работать в условиях ограниченных объемов выборок. Особенно остро это ощущают службы контроля качества предприятий, имеющих мелкосерийное производст-

во. Такое же положение существует в производстве и эксплуатации дорогостоящих и высоконадежных технических изделий.

Особенности и пути решения в условиях малых выборок задач каждого класса в отдельности для управления качеством машиностроительной продукции как дискретных случайных величин заключаются в эффективном использовании статистической информации ограниченного объема. Информационный подход оказывается весьма полезным и результативным при решении задач статистического анализа. На это указывается в работах Н. Винера, Р. Фишера, К. Шеннона, которые фактически построили фундамент современной теории информации.

В условиях ограниченной информации обычно оперируют малой выборкой, под которой понимается малое число наблюдений над случайной величиной, описывающее изучаемое явление. Но такая формулировка относительна и не всегда определяет границу между малой и большой выборкой. Различать большие и малые выборки необходимо, но точной границы между ними установить нельзя. Выборка, сделанная из совокупности с небольшим разбросом признака, может считаться большой, тогда как выборка такого же объема, произведенная из более разнородной совокупности, окажется малой. Вопрос о том, к какой категории отнести выборку, решается в каждом конкретном случае.

Важно иметь в виду, что к большим выборкам можно применять аппарат теории малых выборок, тогда как обратное приводит к значительным ошибкам. В сомнительных случаях для получения надежных результатов рекомендуется пользоваться аппаратом малых выборок. В больших выборках средние теснее группируются около генеральной средней, что позволяет получать более точные и надежные результаты. В малых выборках приходится довольствоваться более широкими границами для средних или меньшей достоверностью результатов. Тем не менее, теория малых выборок нашла в практике широкое распространение и применяется даже в тех случаях, когда во власти исследователя сделать выборку большой [1].

К определению малой выборки можно подходить с информационных позиций. Поскольку, как отмечалось, случайная выборка несет информацию об изучаемом явлении, то статистическая обработка есть не что иное, как извлечение информации из выборки. Если это так, то следует ответить всего на два вопроса: какое количество информации содержится в выборке заданного объема и какое количество информации необходимо для получения результата с заданной точностью и достоверностью.

Если нет возможности точно оценить количество информации, с которым имеет дело статистик, то это необходимо сделать хотя бы приближенно.

Для этой цели можно использовать следующие соображения. Во многих методах статистической обработки выборки используется идея группировки данных (гистограмма, критерий  $\chi^2$  и т.д.). При анализе выборок значительного объема с использованием этих методов обычно удается получить удовлетворительный результат. Но группировка наблюдений вызывает обычно уменьшение информации, которая извлекается из выборки. Значит, если применение такого метода в конкретной задаче удовлетворяет статистика, то, очевидно, выборка содержит избыточную информацию для заданных точности и достоверности. Отсюда следует, что, во-первых, количество информации в выборке заданного объема неразрывно связано с возможностью достичь вполне определенных точности и достоверности. Следовательно, можно говорить о существовании достаточной выборки, которая как раз и является границей, разделяющей большие и малые выборки. Во-вторых, если выборка меньше достаточной, то для устранения потерь информации при ее обработке необходимо отказаться от группировки наблюдений и перейти к методам, основанным на использовании каждой отдельной реализации (статистическая функция распределения, критерий знаков, критерий Уилкоксона и др.). Этот момент и может быть положен в основу определения малой выборки.

Таким образом, выборку можно считать малой, если при ее обработке методами, основанными на группировке наблюдений, нельзя решить задачу с требуемой точностью и достоверностью результата.

Как было сказано выше, правильность и эффективность решения практических задач определяется объемом информации о функционировании технологического процесса, которую можно разделить на консервативную и оперативную. К консервативной информации относится знание закона распределения и нахождение эффективных статистических оценок его параметров, а к оперативной – объем выборки. Под информацией, в данном случае, будем понимать отражение случайной выборкой изучаемого явления, поэтому можно сказать, что при статистическом анализе происходит извлечение информации из выборки.

Рассмотрим информационное обеспечение процесса управления качеством статистическими методами. Объем оперативной информации – объема выборки из генеральной совокупности, может быть найден из информационной модели

$$D(X) \geq \frac{1}{I(X)}, \quad (1)$$

где  $I(X)$  – количество информации относительно параметра  $X$ , которая состоит из количества консервативной информации  $I_K(X)$  и количества оперативной информации  $p$ .

Из неравенства (1) видно, что уменьшение дисперсии  $D(X)$  требует увеличения количества информации  $I(X)$ . Для этого необходимо разработать методы получения большего объема необходимой информации без увеличения объема контрольной выборки. Другими словами, для использования статистических методов управления качеством необходимо знание закона распределения и его эффективных статистических оценок. Чем точнее найдена функция плотности распределения и чем эффективнее статистические оценки, тем меньше необходимый объем выборки.

Под эффективностью статистической оценки будем понимать минимальность ее дисперсии и несмещенность. Данное положение показывает, что для уменьшения объема выборки необходимо сосредоточить внимание на двух задачах: определение закона распределения параметров качества и нахождение их эффективных статистических оценок.

Решение первой задачи не представляет труда только при условии, что количество статистической информации о показателях качества значительное. Но в условиях ограниченной информации эта возможность отсутствует.

Такое положение дел требует разработки новых методов, позволяющих эффективно использовать статистическую информацию. Основой для разработки таких методов может служить теория порядковых статистик.

Рассмотрим сущность порядковых статистик. Если  $x_1, \dots, x_n$  – выборка объема  $n$  ( $n$  – совокупность независимых и одинаково распределенных случайных величин) из генеральной совокупности, имеющей распределение  $F(x)$ , то  $x_{(1)} \leq \dots \leq x_{(n)}$  – соответствующие упорядоченные величины, называется порядковыми статистиками из  $F(x)$ . Если для обычной выборки энтропия отлична от нуля, то для порядковых статистик она равна нулю за счет упорядочения  $x_{(1)}, x_{(2)} \dots x_{(n)}$ . Таким образом, упорядочение величин  $x_i$  дает дополнительную информацию о процессе без изменения объема выборки. Это делает применение порядковых статистик более эффективными при решении практических задач. Заметим, что дополнительную информацию получают, в основном, за счет усложнения зависимостей, описывающих статистические оценки, и имеющейся взаимосвязи порядковых статистик между собой.

Пусть имеется выборка объема  $n$ . Перенумеруем выборочные значения в порядке возрастания их величины:  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ . В частности, наименьшим значением будет тогда  $x_{(1)}$ . Полученное после нумерации  $i$ -е значение  $x_{(i)}$  называется  $i$ -й порядковой статистикой. В этом смысле медиана и квантили являются порядковыми статистиками (медиана, например, равна  $\frac{1}{2}(n+1)$ -му наблюдению, или сред-

нему значению  $\frac{1}{2}n$ -го и  $\frac{1}{2}(n+1)$ -го наблюдений).

Порядковыми статистиками являются также наименьшее ( $i=1$ ) и наибольшее ( $i=n$ ) значения в выборке. Разность  $R = x_{(n)} - x_{(1)}$  называется размахом выборки (выборочным размахом).

Действие по упорядочению выборки называется ранжированием. Выборка, содержащая одни и те же элементы, может быть реализована в опыте  $n!$  разными способами в зависимости от порядка следования элементов. Если опыт поставлен правильно, все возможные реализации равновероятны и среди них с вероятностью  $1/n!$  может оказаться выборка, ранжированная уже в процессе формирования. Эта вероятность очень мала даже для небольших объемов выборки. Отсюда следует, что операция упорядочения превращает ранжированную выборку в уникальный объект.

Отличие ранжированной выборки от исходной количественно можно оценить общепринятой мерой беспорядка – энтропией. Так как число реализаций выборки  $n!$ , а все реализации равновероятны, то в этом случае энтропия

$$H = -\log_2(1/n!) = \log_2 n! \quad (2)$$

Ранжированная выборка обладает энтропией, равной нулю [2]. То есть уменьшение энтропии происходит в процессе упорядочения. Энтропия системы уменьшается в результате поступления информации, причем уменьшение и равно количеству поступившей информации. Отсюда следует, что в процессе упорядочения выборка как бы наполняется большим количеством информации. Эта информация может полезно быть использована, в частности, для получения оценок параметров распределений по небольшому количеству испытаний.

Известно, что если на действительный размер изготовления детали на настроенном станке влияет закономерно изменяющаяся погрешность, возрастающая сначала замедленно, а затем ускоренно, то распределение размеров происходит по закону Симпсона [3, 4]. Это распределение может иметь место при совместном действии размерного износа режущего инструмента с сильно выраженной фазой начального износа и увеличении силы резания в конце стойкости инструмента в результате его прогрессирующего затупления. В таких условиях количество факторов уменьшается, так как процесс механической обработки является полустатистическим или окончательным, при котором множество факторов нейтрализовались на предыдущей операции (черновой).

Для модели Симпсона, соответствующей рассеиванию размеров изготовления деталей по 6–7-му квантилю точности, предлагаемая в [5] оценка фактического поля рассеяния состоятельная и смещенная:

$$\omega = 2\sqrt{6} \cdot S, \quad (3)$$

где  $S$  – эмпирический стандарт, явно смещенный в сторону уменьшения истинного значения, и это смещение уменьшается с ростом  $n$ . Это смещение в среднем составляет для модели распределения Симпсона при  $n = 5$  – 5%, при  $n = 10$  – 2,5%, и при  $n = 50$  – 0,4%.

Для модели распределения Симпсона А. Сарханом и Б. Гринбергом только для выборок объема до 5 были найдены методом Ллойда коэффициенты оптимальных линейных оценок математического

ожидания  $\theta_1^*$  и стандартного отклонения  $\sigma^* = \frac{\theta_2^*}{2\sqrt{6}}$

по усеченным выборкам с функцией плотности:

$$f(x) = \frac{4}{\theta_2^2} \left( \frac{1}{2} \theta_2 - |x - \theta_1| \right); \quad |x - \theta_1| \leq \frac{1}{2} \theta_2, \quad (4)$$

где  $\theta_1$  – математическое ожидание и  $\theta_2$  – размах.

Пользуясь методом получения оценок параметров распределения, применяя порядковые статистики, была получена оценка параметра  $\theta_2$  модели Симпсона [6]. Она имеет вид:

$$\hat{\theta}_2 = \frac{60 \sum_{i=1}^n (2i - n - 1) X(i)}{7n(n-1)}, \quad (5)$$

и является несмещенной и состоятельной.

Для данной модели существует еще одна несмещенная и состоятельная оценка параметра  $\theta_2$  через эмпирический размах [6]:

$$\hat{\theta}_2 \approx T(X(n) - X(1)), \quad (6)$$

где

$$T = - \left[ \frac{1}{2^n (2n+1)} + \sum_{k=1}^n \frac{(-1)^k C_n^k}{2^k (2k+1)} \right]^{-1}. \quad (7)$$

Статистический анализ существующих оценок для распределения Симпсона, проведенный методом статистического моделирования для объемов выбор-

ки  $n = 5$ ,  $n = 10$  и  $n = 50$ , показал, что предлагаемые оценки, (5) и (6), несмещенные и наименьшую дисперсию из этих оценок имеет для любых  $n$  оценка (6). Оценки (5) и (3) имеют при каждом  $n$  почти одинаковые дисперсии. Поэтому оценка (5) лучше оценки (3) только за счет ее несмещенности. Эффективность оценки (3) по отношению к оценке (5) при  $n = 5$  равна 0,97, при  $n = 10$  – 0,96, а при  $n = 50$  она равна 0,87.

## Вывод

Исходя из выше сказанного, можно сделать вывод, что теория порядковых статистик может стать научной основой создания методологии управления качеством по малому количеству статистической информации. Проведенный статистический анализ существующих оценок параметров для модели Симпсона позволяет эффективно принимать решения при управлении качеством.

## Список литературы

1. Дружинин Н.К. Выборочный метод и его применение в социально-экономических исследованиях / Н.К. Дружинин. – М., 1970. – 280 с.
2. Ефимов А.Н. Порядковые статистики – их свойства и приложения / А.Н. Ефимов. – М.: Знание, 1980. – С. 64.
3. Ковишов А.Н. Технология машиностроения / А.Н. Ковишов. – М.: Машиностроение, 1987. – 320 с.
4. Воробьев Л.Н. Технология машиностроения и ремонт машин / Л.Н. Воробьев. – М.: Высшая школа, 1981. – 334 с.
5. Триц Р.М. Применение оптимальных линейных оценок для оценки точности механической обработки / Р.М. Триц, Ю.А. Яновский, Н.Ю. Ламнауэр // Вісник НТУ „ХПИ”: Збірник наукових праць. – Х., 2004. – № 28. – С. 157-164.
6. Ламнауэр Н.Ю. Анализ точности механической обработки высокоточных деталей / Н.Ю. Ламнауэр, Р.М. Триц, Ю.А. Яновский // Восточно-европейский журнал передовых технологий. – 2005. – № 2(13). – С. 56-59.

Поступила в редколлегию 9.12.2010

**Рецензент:** д-р техн. наук, проф. И.П. Захаров, Харьковский национальный университет радиоэлектроники, Харьков.

## КЕРУВАННЯ ЯКІСТЮ МАШИНОБУДУВАЛЬНОЇ ПРОДУКЦІЇ ПО МАЛИХ ВИБІРКАХ

Г.М. Триш

*Розглянуто можливості застосування математичного апарату порядкових статистик для керування якістю в машинобудуванні по малій кількості статистичної інформації. Проведено статистичний аналіз отриманих оцінок для розподілу Симпсона.*

**Ключові слова:** ефективні оцінки параметрів, порядкові статистики, мала вибірка.

## QUALITY OF MACHINE-BUILDING PRODUCTS MANAGEMENT ON SMALL SELECTIONS

G.M. Trisch

*Possibilities of application of mathematical vehicle are considered index statistician for a quality management in an engineer on a few of statistical information. The statistical analysis of the got estimations is conducted for distributing of Simpson.*

**Keywords:** effective estimations of parameters, index statisticians, small selection.