

УДК 005.6:658.5:519.226.3

**МЕТОДИКА СТАТИСТИЧЕСКОГО ПРИЕМОЧНОГО КОНТРОЛЯ НА ОСНОВЕ БАЙЕСОВСКОГО ПОДХОДА (ГИПЕРГЕОМЕТРИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ)****<sup>1</sup>Юдин С.В., <sup>2</sup>Протасьев В.Б., <sup>3</sup>Подкопаев Р.Ю., <sup>4</sup>Юдин А.С.**<sup>1</sup>ФГБОУ ВО «Российский экономический университет им. Г.В. Плеханова»,  
Тульский филиал, Тула, e-mail: svjudin@rambler.ru;<sup>2</sup>ФГБОУ ВО «Тульский государственный университет», Тула, e-mail: avprotasev@mail.ru;<sup>3</sup>Сертифицированный аудитор систем менеджмента, Тула, e-mail: rquant@mail.ru;<sup>4</sup>ООО «Научно-исследовательский центр оборонно-промышленного комплекса»,  
Москва, e-mail: alextula78@rambler.ru

Современное производство во многих случаях рассчитано на выпуск мелкосерийной и штучной продукции. Существующие методы статистического контроля и управления качеством предполагают наличие большого количества однородных данных, полученных, как правило, за длительный период наблюдения. Противоречие, возникающее между применяемыми методами математической статистики и потребностями производства, можно разрешить на основе использования байесовского подхода, предполагающего наличие априорной информации, и принципа «интегрального риска», позволяющего построить интервальные оценки параметров дефектности партии изделий, представленных к инспекции. Проведен анализ литературы, посвященной байесовскому подходу в обработке результатов измерений малой выборки. Предложен синтез байесовского подхода и концепции «интегрального риска». Этот подход дает возможность оценить качество партии изделий по результатам контроля малой выборки с использованием априорной информации о распределении количества или доли дефектных изделий в партии за предшествующие периоды. В статье предложена методика расчета планов статистического приемочного контроля с использованием табличного процессора MS Excel. Простота и универсальность этой методики позволяет применять ее инженерами, не являющимися специалистами в области программирования. Использование методики позволит снизить вероятность ошибки и тем самым повысить экономическую эффективность производства.

**Ключевые слова:** контроль, качество, статистические методы, менеджмент качества, управление процессами

**METHODOLOGY OF STATISTICAL ACCEPTANCE CONTROL BASED ON THE BAYES APPROACH (HYPERGEOMETRIC DISTRIBUTION)****<sup>1</sup>Ydin S.V., <sup>2</sup>Protasev V.B., <sup>3</sup>Podkopaev R.Yu., <sup>4</sup>Yudin A.S.**<sup>1</sup>Plekhanov Russian University of Economics, Tula branch, Tula, e-mail: svjudin@rambler.ru;<sup>2</sup>Tula State University, Tula, e-mail: avprotasev@mail.ru;<sup>3</sup>Certified auditor of management systems, Tula, e-mail: rquant@mail.ru;<sup>4</sup>Scientific Research Center of the Defense Industrial Complex, Moscow, e-mail: alextula78@rambler.ru

Modern production in many cases is designed for the production of small-batch and piece products. The existing methods of statistical control and quality management assume the presence of a large number of homogeneous data obtained, as a rule, for an unlimited time. The contradiction arising between the applied methods of mathematical statistics and the requirements of production can be resolved on the basis of using the Bayesian approach that assumes the existence of a priori information and the principle of «integral risk», which allows to construct interval estimates of the defects parameters of the batch of products submitted to the inspection. An analysis of the literature devoted to the Bayesian approach in the processing of small sample measurements is carried out. A synthesis of the Bayesian approach and the concept of «integral risk» is proposed. This approach makes it possible to evaluate the quality of a batch of products based on the results of a small-sample control using a priori information on the distribution of the quantity or proportion of defective products in the batch over previous periods. The article proposes a technique for calculating plans for statistical acceptance control using the MS Excel spreadsheet. The simplicity and versatility of this technique makes it possible to use it by engineers who are not specialists in programming. Using the technique will reduce the probability of error and, thereby, increase the economic efficiency of production.

**Keywords:** control, quality, statistical methods, quality management, process management

Статистические методы контроля и управления качеством широко применяются во всех отраслях промышленности. Благодаря широкому внедрению этих методов добились успеха японские, корейские, китайские корпорации. Еще раньше статистические методы применялись и применяются в странах Западной Европы и Северной Америки. Современное производство немислимо без их использования.

Существующие методы статистического контроля и управления качеством, управления технологическим процессом основаны на предположении о наличии стабильного процесса, происходящего в постоянных условиях длительное время, возможности получить неограниченную выборку изделий и заготовок. Это положение с переходом на мелкосерийное, и тем более штучное, производство существенно меняется, и ис-

пользуемые стандартные методы статистических выводов перестают работать.

Основная проблема, возникающая при контроле малых выборок, заключается в существенном увеличении доверительных интервалов, поскольку их ширина пропорциональна  $1/\sqrt{n}$ , где  $n$  – объем выборки. Если точность оценок при  $n = 100$  приблизительно равна 10% (приблизительно потому, что в оценке ширины доверительного интервала присутствуют и другие множители), то при  $n = 10$  – 32%. Особенно существенны ошибки в оценке параметров при проведении контроля партий изделий по качественному признаку.

До настоящего времени эта проблема в математической статистике не разрешена, несмотря на большое количество работ в этой области.

Цель исследования: исследования С.В. Юдина и др. [1, 2] привели к выводу о необходимости разработки специальных методов контроля и управления качеством в случае мелкосерийного и штучного производства. Это связано с тем, что многие предприятия выпускают штучную и мелкосерийную продукцию, номенклатура которой постоянно меняется.

В монографии Д.В. Гаскарова и В.И. Шаповалова 1978 г. [3] впервые на высоком научном уровне рассмотрен вопрос об оценке параметров случайных чисел и проверки статистических гипотез по итогам данных, полученных при анализе малых выборок. Авторы использовали метод минимума энтропии и математический аппарат обобщенных функций. Они показали возможность при анализе малых выборок проводить оценки параметров распределений и даже проверку статистических гипотез, как о виде закона распределения, так и о наличии статистических связей.

Тем не менее точность оценок и доверительные вероятности, получаемые в результате предложенных методик, не могут быть применены в условиях высоких требований к изделиям ВВТ.

В 1989 г. была опубликована работа В.П. Савчука, посвященная байесовским методам статистического оценивания [4]. Эти методы основаны на важном принципе: информация меняет вероятности.

Целью исследования является разработка методов статистического приемочного контроля, основанного на гипергеометрическом распределении, учитывающие историю процесса с целью минимизировать риски при контроле малых партий.

#### Материалы и методы исследования

Формула Байеса уже давно, с конца XVIII века, применяется исследователями в самых разных обла-

стях, начиная с простых задач оценки вероятностей тех или иных событий на основе знаний о прошлом, заканчивая оценкой надежности сложных систем и построения систем искусственного интеллекта.

Уникальная формула Байеса имеет вид

$$P(H_k / A) = \frac{P(H_k)P(A / H_k)}{\sum_{i=1}^n P(H_i)P(A / H_i)}. \quad (1)$$

Здесь  $A$  – наблюдаемое событие,  $\{H_i, i = \overline{1, n}\}$  – система гипотез.

Вероятности  $p(H_i)$  называются априорными (известными до опыта) вероятностями, а вероятности  $p(H_i/A)$  – апостериорными (полученными эмпирически после опыта) вероятностями.

Важнейшей проблемой в байесовском подходе является определение вероятностей, как события  $p(A)$ , так и условных вероятностей  $p(H_i/A)$ .

Данные вероятности могут быть определены только из длительных наблюдений за историей процесса. Чем дольше длится наблюдение, тем точнее наши оценки.

Следует отметить, что любые методы, основанные на использовании априорных вероятностей для исчисления апостериорных вероятностей, можно классифицировать как байесовские.

В работе [4] показано, что байесовский подход можно применять и в случае непрерывных распределений, причем многие оценки совпадают с теми же, найденными по принципу максимума энтропии.

Несмотря на важность этой работы для изучения возможностей байесовского подхода при оценке вероятностей интересующих нас событий и проверке статистических гипотез, в ней уделено очень мало внимания проблеме «малой выборки».

В 2008 г. в работе И.Е. Ишемгужина и др. [5] были рассмотрены критерии согласия для малой выборки при оценке надежности оборудования. Показано, что можно, применяя специальные критерии согласия, с достаточной надежностью идентифицировать закон распределения в задачах оценки надежности.

Недостатком этой статьи является недостаточная теоретическая проработка и отсутствие единого подхода в случае разных распределений.

Большая и теоретически важная работа Б.И. Сухорученкова 2010 г. по анализу малых выборок [6] содержит интересный авторский материал. К сожалению, в работе выборка объемом  $n = 30$  считается малой. Для таких выборок нет нужды применять особые методы.

В работе М.П. Кривенко [7] используются информационная мера Кульбака – Лейблера, информация Фишера, параметрические и непараметрические статистические оценки в приложении к многомерным распределениям. Несмотря на важные теоретические результаты, практическое использование предложенных методов затруднено в связи со сложностью математического аппарата и тем, что в инженерной практике редко приходится использовать многомерные распределения.

В работах С.В. Юдина и др. [1, 2] на базе теории информации и байесовского подхода предложены новые методы, обладающие высокой надежностью, действительно применимые к анализу малых выборок, особенно в задачах оценки качества партии по данным выборки.

В статье С.В. Юдина [1] представлены основные теоретические положения нового подхода, основанные на знании априорной информации о качестве продукции и параметрах технологического процесса, а также концепции «интегрального риска».

*Принцип интегрального риска и гипергеометрическое распределение*

В работах [1, 2] был представлен принцип интегрального риска, который заключается в том, что при анализе процедуры статистического выборочного контроля учитываются интегральные вероятности принять партию с дефектностью, превышающей заданную (риск Потребителя  $\beta$ ), или отклонить партию с дефектностью ниже граничной (риск Поставщика  $\alpha$ ).

Поскольку рассматриваемый ниже подход использует априорное распределение вероятности доли брака  $q$  на множестве контролируемых партий, то апостериорная вероятность интересующих контролера событий определяется с использованием обобщенного байесовского подхода.

При контроле партий малого объема, как правило, используется точное гипергеометрическое распределение числа дефектных изделий в выборке:

$$p(n, d; N, D) = \frac{C_D^d \cdot C_{N-D}^{n-d}}{C_N^n}, \quad (2)$$

где  $n$  – объем выборки;  $d$  – количество дефектных изделий в выборке;  $N$  – объем партии;  $D$  – количество дефектных изделий в партии.

Пусть  $c$  – приемочное число. Тогда оперативная кривая плана контроля, т.е. вероятность принять партию с количеством дефектов  $D$ , равна

$$L\left(q = \frac{D}{N} \middle| N, n, c\right) = \sum_{i=0}^c \frac{C_D^i C_{N-D}^{n-i}}{C_N^n}. \quad (3)$$

При традиционном подходе риск Изготовителя (Поставщика)  $\alpha$  определяется как вероятность забраковать партию с количеством дефектных изделий равных  $D_n$ :

$$\alpha = 1 - L\left(q = \frac{D_n}{N} \middle| N, n, c\right) = 1 - \sum_{i=0}^c \frac{C_{D_n}^i C_{N-D_n}^{n-i}}{C_N^n}. \quad (4)$$

Риск Потребителя  $\beta$  определяется как вероятность принять партию с количеством дефектных изделий  $D_6$ :

$$\beta = L\left(q = \frac{D_6}{N} \middle| N, n, c\right) = \sum_{i=0}^c \frac{C_{D_6}^i C_{N-D_6}^{n-i}}{C_N^n}. \quad (5)$$

Здесь  $D_n$  и  $D_6$  – приемочное и браковочное количества дефектных изделий в партии ( $D_n < D_6$ ).

Недостатком этого подхода является точность оценки, в то время как необходима интервальная оценка вероятности принятия бракованной партии. Потребителя, как правило, интересует общая вероятность принятия партии с неудовлетворительным уровнем дефектности, т.е. интегральный риск».

**Результаты исследования и их обсуждение**

Пусть известна гистограмма распределения дефектных изделий в партии, исходя из результатов исследований или результатов контроля партий в прошлом:

$f_i$  ( $i = 0 \dots N$ ). Здесь  $f_i$  – количество партий, в которых количество дефектных изделий равно  $i$ ;  $N$  – объем партии. Пусть  $k = \sum_{i=0}^N f_i$  – общее количество опытов. Тогда распределение вероятностей имеет вид

$$p_i = P(D = i | N) = \frac{f_i}{k} \quad (i = 0 \dots N).$$

Пусть заданы интегральные риски Потребителя  $\beta$  и Поставщика  $\alpha$ . Тогда для интегральных рисков имеем следующую систему уравнений:

$$\left. \begin{aligned} \sum_{j=0}^{D-1} p_j \left[ \sum_{i=0}^c \left( 1 - \frac{C_j^i C_{N-j}^{n-i}}{C_N^n} \right) \right] &= \alpha \\ \sum_{j=D}^N p_j \left[ \sum_{i=0}^c \left( \frac{C_j^i C_{N-j}^{n-i}}{C_N^n} \right) \cdot p_i \right] &= \beta \end{aligned} \right\}. \quad (6)$$

Здесь  $D$  – предельно допустимое количество дефектных изделий в партии.

Система уравнений (6) дает возможность определить объем выборки  $n$  и приемочное число  $c$ , что определяет план контроля с заданными характеристиками.

*Методика расчета гипергеометрических планов СПК на основе байесовского подхода*

Входными данными для расчета оперативной кривой плана СПК, основанного на гипергеометрическом распределении, являются следующие числа: объем партии  $N$ ; объем выборки  $n$ ; предельно допустимое количество дефектных изделий в партии  $D_{max}$ ; приемочное число  $c$ .

Справочными числами являются риски изготовителя и потребителя.

На рис. 1 представлен скриншот листа MS Excel с расчетами.

Пусть из истории наблюдений за готовыми изделиями известно распределение вероятности наблюдения заданного количества дефектных изделий  $D$  в выборке объемом  $N$ . Эти данные представлены в столбце ‘G’.

В столбцах ‘H’ и ‘I’ приведены результаты расчетов интегральных рисков Поставщика и Потребителя. Эти расчеты осуществлялись в соответствии с формулой.

На рис. 2 в большем масштабе представлены интересующие нас результаты.

Если исходить из предельно допустимого количества дефектных изделий в партии  $D_{max} = 5$ , то в ячейке ‘H17’ интегральный риск  $\alpha = 0,068$  – это вероятность того, что будет забракована партия с количеством дефектных изделий 5 и менее; интегральный риск  $\beta = 0,005$  (ячейка ‘I18’) – это вероятность того, что будет принята партия с числом дефектных изделий 6 и более.

	A	B	C	D	E	F	G	H	I	
1	План СПК на основе Гипергеометр распр									
2	Определение приемочного "с" и браковочного "d" чисел									
3	Задано:									
4	Риск Изготовителя $\alpha=$						0,05			
5	Граничное значение оперативной кривой $=1-\alpha$						0,95			
6	Объем выборки $n=$						10			
7	Объем партии $N=$						100			
8	Предельно допустимое количество									
9	дефектных изделий в партии $D_{\max}=$						5			
10	Приемочное число $c=$						0			
11	$L(D_{\max})=$	0,583752		D		L(D)	Вероятность наблюдения заданного количества дефектных изделий D в выборке объемом N	Интегральный риск $\alpha$	Интегральный риск $\beta$	
12					0	1	0,6	0	0,927166327	
13					1	0,9	0,2	0,02	0,327166327	
14					2	0,809091	0,1	0,039090909	0,147166327	
15					3	0,726531	0,05	0,052764378	0,066257236	
16					4	0,651631	0,02	0,059731767	0,029930706	
17					5	0,583752	0,02	0,06805672	0,016898095	
18					6	0,522305	0,01	0,072833673	0,005223047	
19					7	0,46674	0	0,072833673	0	
20					8	0,416553	0	0,072833673	0	
21					9	0,371276	0	0,072833673	0	
22					10	0,330476	0	0,072833673	0	

Рис. 1. Результаты расчетов в MS Excel

	E	F	G	H	I
11	D	L(D)	Вероятность наблюдения заданного количества дефектных изделий D в выборке объемом N	Интегральный риск $\alpha$	Интегральный риск $\beta$
12	0	1		0,6	0
13	1	0,9		0,2	0,02
14	2	0,809091		0,1	0,039090909
15	3	0,726531		0,05	0,052764378
16	4	0,651631		0,02	0,059731767
17	5	0,583752		0,02	0,06805672
18	6	0,522305		0,01	0,072833673
19	7	0,46674		0	0,072833673
20	8	0,416553		0	0,072833673
21	9	0,371276		0	0,072833673
22	10	0,330476		0	0,072833673

Рис. 2. Результаты расчета интегральных рисков

Если установить оба риска равные 0,05, то с такой вероятностью мы забракуем партию с числом дефектных изделий 3 и меньше и примем партию с числом дефектных изделий 4 и более.

Представленный подход имеет широкую область применения, как в гражданской промышленности, так и на предприятиях ОПК.

## Выводы

В результате проведенных исследований можно отметить, что использование байесовского подхода, базирующегося на принципе «информация меняет вероятности», а также концепции «интегрального риска», дает возможность с высокой точностью и надежностью оценить качество



контролируемой партии, оценить риски, как Потребителя, так и Поставщика. Представлена методика расчета планов СПК, основанных на гипергеометрическом распределении, учитывающая историю работы предприятия.

#### Список литературы

1. Юдин С.В. Некоторые проблемы статистического контроля качества и методы их решения // *Фундаментальные исследования*. 2015. № 10–2. С. 324–329.
2. Григорович В.Г., Юдин С.В., Козлова Н.О., Шильдин В.В. Информационные методы в управлении качеством. М.: РИА «Стандарты и качество», 2001. Вып. 1 (10). 208 с.
3. Гаскаров Д.В., Шаповалов В.И. Малая выборка. М.: Статистика, 1978. 248 с.
4. Савчук В.П. Байесовские методы статистического оценивания: надежность технических объектов. М.: Наука, 1989. 322 с.
5. Ишемгузин И.Е., Атнагулов А.Р., Зотов А.Н., Ишемгузин Е.И. Специальные критерии согласия для малой выборки при оценке надежности нефтепромышленного оборудования // *Нефтегазовое дело*. 2008. № 1. URL: [http://ogbus.ru/authors/IshemguzhinIE/IshemguzhinIE\\_1.pdf](http://ogbus.ru/authors/IshemguzhinIE/IshemguzhinIE_1.pdf) (дата обращения: 17.08.2018).
6. Сухорученков Б.И. Анализ малой выборки. Прикладные статистические методы. М.: Вузовская книга, 2010. 384 с.
7. Кривенко М.П. Прикладные методы оценивания распределения многомерных данных малой выборки. М.: ИПИ РАН, 2011. 146 с.