

УДК 005.6:658.5:519.226.3

МЕТОДИКА РАСЧЕТА ИНФОРМАЦИОННЫХ ПЛАНОВ СТАТИСТИЧЕСКОГО ПРИЕМОЧНОГО КОНТРОЛЯ НА ОСНОВЕ БАЙЕСОВСКОГО ПОДХОДА

¹Юдин С.В., ²Протасьев В.Б., ³Подкопаев Р.Ю., ⁴Юдин А.С.

¹ФГБОУ ВО «Российский экономический университет им. Г.В. Плеханова»,

Тульский филиал, Тула, e-mail: svjudin@rambler.ru;

²ФГБОУ ВО «Тульский государственный университет», Тула, e-mail: avprotasev@mail.ru;

³Сертифицированный аудитор систем менеджмента, Тула, e-mail: rquant@mail.ru;

⁴ООО «Научно-исследовательский центр оборонно-промышленного комплекса»,
Москва, e-mail: alextula78@rambler.ru

Повышенные требования к качеству продукции, особенно в оборонной промышленности, требуют нового подхода к решению задач контроля и управления качеством. Основные проблемы, связанные с новыми задачами, заключаются в том, что объем производства конкретных изделий и деталей, как правило, мал и исчисляется единицами. Современные методы математической статистики основываются на предположении о том, что исследуется стационарный процесс, характеристики которого не меняются со временем. Штучное и мелкосерийное производство не может обеспечить исследователя надежной экспериментальной базой, позволяющей с достаточной точностью оценить доверительные интервалы исследуемых параметров. Для решения отмеченной проблемы авторы предлагают интеграцию двух подходов: информационного и байесовского. Информационно-статистические методы, основанные на понятии «энтропия», позволяют получать достоверные результаты анализа даже в тех случаях, когда исследуемые параметры имеют негауссово распределение, а байесовский подход дает возможность учесть историю процесса со всеми его изменениями. Показано, что на основе интегрированного подхода, учитывающего предысторию исследуемого процесса, можно построить планы статистического приемочного контроля, обладающие существенно меньшими требованиями к объему контролируемой партии по сравнению с традиционными. Предложена методика проведения соответствующих расчетов.

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

METHODOLOGY OF STATISTICAL ACCEPTANCE CONTROL BASED ON THE BAYES APPROACH (HYPERGEOMETRIC DISTRIBUTION)

¹Ydin S.V., ²Protasev V.B., ³Podkopaev R.Yu., ⁴Yudin A.S.

¹Plekhanov Russian University of Economics, Tula branch, Tula, e-mail: svjudin@rambler.ru;

²Tula State University, Tula, e-mail: avprotasev@mail.ru;

³Certified auditor of management systems, Tula, e-mail: rquant@mail.ru;

⁴Scientific Research Center of the Defense Industrial Complex,
Moscow, e-mail: alextula78@rambler.ru

Increased requirements to quality of production, especially, in the defensive industry demand the new approach to the decision of problems of control and quality management. The basic problems connected with new problems, consist that the volume of output of concrete products and details, as a rule, is small and is estimated in units. Modern methods of mathematical statistics are based on the assumption that the stationary process which characteristics do not change in due course is investigated. Piece and small-scale manufacture cannot provide the researcher with the reliable experimental base allowing with sufficient accuracy to estimate confidential intervals of the studied parameters. For the decision of noted problem authors offer integration of two approaches: information and Bayesian. The information-statistical methods based on concept «entropy», allow to receive authentic results of the analysis even when the studied parameters have non-Gauss distribution, and the Bayesian approach gives the chance to consider history of process with all its changes. It is shown that on the basis of an integrated approach considering background of investigated process, it is possible to construct the plans of statistical acceptance control possessing essentially smaller requirements to volume of controllable party in comparison with the traditional. The technique of carrying out of corresponding calculations is offered.

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

Использование статистических методов для исследования и анализа промышленной продукции началось еще в XIX в. в Великобритании и других промышленно развитых странах. В конце XIX в. статистический подход к контролю и управлению начал

разрабатывать великий российский ученый М.В. Остроградский.

Можно отметить, что именно статистические методы управления технологическим процессом позволили добиться невероятного повышения доли качественных

(т.е. удовлетворяющих требованиям потребителя) изделий в общем потоке. Японское экономическое чудо, концепция «Шесть сигм», «Дао Тойота» и другие примеры дают основание утверждать о важности статистических методов исследования.

Следует заметить, что основные принципы контроля и управления качеством основаны на использовании больших выборок, исчисляющимися десятками, а то и тысячами.

В современном производстве, особенно в оборонной промышленности, когда работа ведется по заказам, согласно которым нужно произвести единицы изделий, такой подход становится неадекватным.

Решению этой проблемы посвящены, в частности, работы [1, 2].

Исследования авторов привели к выводу о необходимости разработки специальных методов контроля и управления качеством в случае мелкосерийного и штучного производства. В качестве методической базы предлагается использовать байесовский подход, применение которого описано в работах [1, 2], и методы математической теории информации, представленных в работах [3, 4].

Целью исследования предлагаемой к рассмотрению работы является попытка интеграции байесовского и информационно-статистического подхода к построению планов статистического приемочного контроля (СПК), в случае контроля партий малого объема.

Материалы и методы исследования

Теоретические основы применения информационного подхода к расчету планов статистического приемочного контроля изложены в работе [4]. Они базируются на фундаментальном труде С. Кульбака [3].

В предисловии к книге С. Кульбака академик А. Колмогоров писал: «аналитический аппарат теории информации был

создан тогда, когда здание математической статистики было в своих основных, находящихся наиболее широкое применение частях уже построено и кодифицировано. Но навыки мысли и аналитический аппарат теории информации должны, по-видимому, привести к заметной перестройке этого здания» [3, с. 5–6].

В последние годы этот подход стал применяться в самых разнообразных отраслях. Примером этому могут служить работы А.Д. Урсула [5], О.В. Цветкова [6], Е.Г. Адрианова и др. [7], Роберта Грэя [8] и других.

Как можно заметить из работ Д.В. Гаскарова, В.И. Шаповалова [2], Н.П. Кривенко [9] и других, достоверные результаты при анализе малых выборок не могут быть получены стандартными подходами математической статистики. Во всех случаях необходимо наличие некоторой априорной информации, что приводит к необходимости использования байесовского подхода.

Можно показать [4], что информационно-статистический подход при расчете планов СПК обеспечивает большую надежность и точность оценки уровня дефектности партии.

Вышеизложенное дает основания полагать, что интеграция информационного и байесовского подходов позволит адекватно оценивать качества малых партий.

В работе [4] рассмотрен информационный подход к построению планов статистического приемочного контроля. Он использует следующие понятия:

q – вероятность получения бракованного изделия;

$p = 1 - q$ – вероятность изготовления годной детали.

Энтропия такого процесса равна

$$H(q) = -q \cdot \ln q - (1 - q) \cdot \ln(1 - q). \quad (1)$$

Основная характеристика плана СПК может быть вычислена по следующей формуле:

$$L(H) = L(q; n; d) = \frac{1}{\sqrt{2\pi\sigma(H^*)}} \int_{-\infty}^{H_0} \exp\left[-\frac{(H' - H(q))^2}{2\sigma^2(H^*)}\right] dH'. \quad (2)$$

При помощи оперативной характеристики можно определить риск поставщика α и риск потребителя β .

Риск поставщика определяется как вероятность забраковать партию, в то время как входной уровень дефектности равен приемочному уровню q_n . Отсюда $\alpha = 1 -$

$L(q_n)$. Риск потребителя β определяется как вероятность принятия партии с уровнем дефектности равным браковочному уровню q_6 : $\beta = L(q_6)$.

Как показано в работе [4], приемочное и браковочное числа можно связать с объемом выборки n следующими уравнениями:

$$\left. \begin{aligned} -\frac{d_n}{n} \ln \frac{d_n}{n} - \left(1 - \frac{d_n}{n}\right) \ln \left(1 - \frac{d_n}{n}\right) &= H(q_n) + t_{1-\alpha} \sqrt{\frac{a^2(q_n) - H^2(q_n)}{n}} \\ -\frac{d_6}{n} \ln \frac{d_6}{n} - \left(1 - \frac{d_6}{n}\right) \ln \left(1 - \frac{d_6}{n}\right) &= H(q_6) - t_{1-\beta} \sqrt{\frac{a^2(q_6) - H^2(q_6)}{n}} \end{aligned} \right\} \quad (3)$$

Здесь t_γ – γ -квантиль нормального распределения.

Концепция «интегрального риска»

Суть концепции «интегрального риска» заключается, в отличие от традиционного подхода точечных оценок, в том, что рассматривается вероятность принятия или забракования партии для интервалов уровней дефектности: вероятность забраковать партию с дефектностью меньше, чем q_n (интегральный риск поставщика α), и вероятность принять партию с дефектностью большей, чем q_6 : (интегральный риск потребителя).

Пусть известна априорная функция плотности вероятностей доли дефектных изделий в партии $w(q)$, предъявляемой к испытаниям. Тогда можно записать следующие уравнения [1, 4]:

$$\left. \begin{aligned} \alpha &= \int_0^{q_6} [1 - L(q)] \cdot w(q) dq \\ \beta &= \int_{q_6}^1 L(q) \cdot w(q) dq \end{aligned} \right\} \quad (4)$$

Опытным путем функция плотности вероятностей получается путем дискретизации области изменения величины q на интервалы, как правило, постоянной ширины Δq и определения частот попадания в каждый интервал.

Разобьем множество возможных значений величины q на k равных интервалов точками $\{0; 1/k; 2/k; \dots; (k-1)/k; 1\}$. Пусть в каждом интервале наблюдается n_i попаданий значений $q \in \left[\frac{i-1}{k}; \frac{i}{k}\right]$.

Тогда можно записать

$$p(q_i) = \frac{n_i}{n}, \quad i = \overline{1, k}. \quad (5)$$

С другой стороны,

$$p(q_i) = w(q_i) dq_i. \quad (6)$$

Тогда можно записать

$$\left. \begin{aligned} \alpha &= \sum_{i=1}^k (1 - L(q_i)) \cdot p(q_i) = \sum_{i=1}^k \left(1 - L\left(\frac{i}{n}\right)\right) \cdot \frac{n_i}{n} \\ \beta &= \sum_{i=1}^k L(q_i) \cdot p(q_i) = \sum_{i=1}^k L\left(\frac{i}{n}\right) \cdot \frac{n_i}{n} \end{aligned} \right\} \quad (7)$$

Методика расчета информационных планов СПК на основе концепции «интегрального риска»

Пусть в результате 100 наблюдений за дефектностью партий изделий получены следующие результаты (табл. 1).

Таблица 1

Эмпирическое распределение доли дефектных изделий

i	q_i	Количество n_i	Доля $p_i = n_i/n$
1	0,005	1	0,01
2	0,01	1	0,01
3	0,015	3	0,03
4	0,02	10	0,1
5	0,025	40	0,4
6	0,03	10	0,1
7	0,035	9	0,09
8	0,04	8	0,08
9	0,045	7	0,07
10	0,05	6	0,06
11	0,055	5	0,05
12	0,06	0	0
ИТОГО:		100	1

Разместим исходные данные в таблицу Excel (рис. 1).

В табл. 2 представлена расшифровка значений ячеек.

На рис. 2 представлена копия экрана с вычисленными значениями оперативной кривой информационного плана СПК

На рис. 3 приведены результаты расчетов интегральных рисков.

Результаты исследования и их обсуждение

Анализ результатов расчета интегральных рисков (рис. 3) приводит к следующим выводам.

Если задать $\alpha = 0,05$, то можно отметить, что ближайшее значение $\alpha = 0,06$ (ячейка О6). Шестая строка соответствует значению $q = 0,025$. Это говорит о том, что вероятность забраковать партию с дефектностью $q \leq 0,025$ равна 0,06.

При $\beta = 0,05$ ближайшее расчетное значение равно $\beta = 0,035$ (ячейка Q12). Двенадцатая строка соответствует значению $q = 0,055$. Это говорит о том, что вероятность принять партию с дефектностью $q \geq 0,055$ равна 0,035.

	A	B	C	D	E	F	G	H	I
1	q	К-во	Доля	Приемочное число		Объем партии N		Объем выборки n	
2	0,005	1	0,01	1		100		10	
3	0,01	1	0,01						
4	0,015	3	0,03	H0	q0=d/n	ШАГ Δq			
5	0,02	10	0,1	0,325083	0,1	0,005			
6	0,025	40	0,4						
7	0,03	10	0,1	DH0					
8	0,035	9	0,09	0,04345					
9	0,04	8	0,08						
10	0,045	7	0,07						
11	0,05	6	0,06						
12	0,055	5	0,05						
13	0,06	0	0						
14	0,065	0	0						
37	Всего:	100	1						

Рис. 1. Таблица исходных данных

	J	K	L	M	N
1	Оперативная кривая		MH	DH	t
2	0,993555		0,031479	0,013939	2,486786
3	0,968635		0,056002	0,020904	1,861098
4	0,937835		0,077883	0,025872	1,536854
5	0,906204		0,098039	0,029687	1,317738
6	0,875123		0,116907	0,032715	1,150946
7	0,844962		0,134742	0,035162	1,015064
8	0,815782		0,151714	0,037156	0,899406
9	0,78754		0,167944	0,038784	0,797915
10	0,760167		0,183521	0,04011	0,706838
11	0,733587		0,198515	0,041181	0,623697
12	0,707732		0,212982	0,042035	0,54677
13	0,68254		0,226968	0,0427	0,474813
14	0,657959		0,240509	0,043201	0,4069

Рис. 2. Расчет оперативной кривой информационного плана СПК

	O	P	Q	R	S
1	Интегральный риск α		Интегральный риск β	q	
2	6,44515E-05		0,84196		0,005
3	0,000378103		0,832024		0,01
4	0,002243041		0,822338		0,015
5	0,011622615		0,794203		0,02
6	0,061573472		0,703582		0,025
7	0,077077245		0,353533		0,03
8	0,093656882		0,269037		0,035
9	0,110653674		0,195617		0,04
10	0,127442018		0,132613		0,045
11	0,143426816		0,079402		0,05
12	0,158040241		0,035387		0,055
13	0,158040241		0		0,06

Рис. 3. Результаты расчета интегральных рисков

Таблица 2

Описание таблицы исходных данных

№ п/п	Обозначение	Адреса
1	q – изменяемая величина границ интервалов дискретизации	A1 ... A36
2	К-во (n_i) – количество опытов с данными значениями q	B1 ... B36
3	Доля (p_i) – оценка вероятности наблюдения значения q_i ; $p_i = n_i/n$	C1 ... C36
4	Приемочное число (d_n)	D2
5	Объем партии N	F2
6	Объем выборки n	H2
7	$q_0 = d_n/n$ – эталонное значение	E5
8	ШАГ Δq – ширина интервала дискретизации	F5
9	$H_0 = -q_0 \ln q_0 - (1 - q_0) \ln(1 - q_0)$ – используемый в (1,28) предел интегрирования при расчете информационного плана контроля	D5
10	$DH_0 = (q_0 \ln^2 q_0 + (1 - q_0) \ln^2(1 - q_0) - H_0^2) / n$	D8

Если учесть объем партии $N = 100$, то можно сделать вывод, что с вероятностью 0,06 в партии находится не более $Nq = 2,5$ не годных изделий в среднем; с вероятностью 0,035 в партии находится не менее 3,5 не годных изделий в среднем.

Можно исходить не из заданных интегральных рисков, а из заданных уровней q_n и q_0 , определяя интегральные риски.

При необходимости можно менять все три основных параметра: объем партии N , объем выборки n , приемочное число d . Варьируя эти три параметра, можно получить оптимальный для текущих условий план контроля.

Выводы

Проведенные исследования дают основание утверждать, что описанный выше подход, основанный на интеграции байесовского подхода и информационно-статистического, позволяет получить адекватный результат при малых объемах выборки.

Список литературы

1. Юдин С.В. Некоторые проблемы статистического контроля качества и методы их решения // Фундаментальные исследования. 2015. № 10–2. С. 324–329.
2. Гаскаров Д.В., Шаповалов В.И. Малая выборка. М.: Статистика, 1978. 248 с.
3. Кульбак С. Теория информации и статистика. М.: Наука, 1967. 408 с.
4. Григорович В.Г., Юдин С.В., Козлова Н.О., Шильдин В.В. Информационные методы в управлении качеством. М.: Стандарты и качество, 2001. 208 с.
5. Урсул А.Д. Природа информации: философский очерк. Челябинск: гос. акад. культуры и искусств; Науч.-образоват. центр «Информационное общество»; Рос. гос. торгово-эконом. ун-т; Центр исслед. глоб. процессов и устойчивого развития. 2-е изд. Челябинск, 2010. 231 с.
6. Цветков О.В. Энтропийный анализ данных в физике, биологии и технике. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2015. 202 с.
7. Андрианова Е.Г., Мельников С.В., Раев В.К. Диссипация и энтропия в физических и информационных системах // Фундаментальные исследования. 2015. № 8. С. 233–238.
8. Gray R.M. Entropy and Information Theory. Second Edition. Springer: N.Y., Dordrecht, Heidelberg, London. 2011. 409 p.
9. Кривенко Н.П. Прикладные методы оценивания распределения многомерных данных малой выборки. М.: ИПИ РАН, 2011. 146 с.