

ВОЛОДИН И. Н., НОВИКОВ А. А., СИМУШКИН С. В.

**ГАРАНТИЙНЫЙ СТАТИСТИЧЕСКИЙ КОНТРОЛЬ
КАЧЕСТВА: АПОСТЕРИОРНЫЙ ПОДХОД**

Введение

Статистический контроль качества включает в себя два аспекта: контроль за стабильностью технологического процесса производства и принятие решения о кондиционности выпускаемой продукции или, более широко, аттестацию ее качества. Мы будем рассматривать только второй аспект этой проблемы, традиционно решаемой в рамках теории проверки статистических гипотез и теории оценивания.

С. Н. Бернштейн [2], по-видимому, был первым, по крайней мере в отечественной литературе, кто обратил внимание на неудовлетворительное решение проблемы гарантийности статистического контроля качества в рамках классического (небайесовского) подхода, основанного на управлении доверительным коэффициентом при аттестации качества выпускаемой продукции. Аналогичная ситуация, естественно, имеет место и при решениях о кондиционности продукции с гарантированными ограничениями на риски изготовителя и потребителя (вероятности ошибок первого и второго рода). Суть претензий можно пояснить на следующем простом примере, который обычно приводил Л. Н. Большев в своих лекциях по вероятностным моделям (1972–1974 гг., Московский университет, не опубликовано).

Предприятие производит выборочные обследования выпускаемых за каждый технологический цикл партий своей продукции, используя некоторую процедуру контроля с риском потребителя 1%. Если предприятие всегда производит некондиционную продукцию, то ее определенная доля, но не более 1%, будет, тем не менее, поступать потребителю. Но в таком случае с точки зрения потребителя процедура контроля не является гарантийной, поскольку ему всегда поставляется продукция, не удовлетворяющая стандарту на качество, — истинный, а не мифический риск потребителя равен 100%. Естественно, случись такое на реальном производстве, оно будет немедленно остановлено. Однако, это произойдет не в соответствии с решающим правилом процедуры контроля, и тем самым будет признано, что существующая процедура неприемлема с точ-

ки зрения ее гарантийности — необходимо управлять не вероятностью пропуска продукции неудовлетворительного качества, а долей кондиционной продукции среди принятой, т.е. управлять выходным уровнем качества $Q_{\text{вых}}$. Решить такую задачу можно лишь в рамках байесовской модели производства продукции и проверки ее качества.

Одно из возможных решений проблемы реальной гарантийности процедуры контроля дает метод «последующих оценок» $Q_{\text{вых}}$, которые вычисляются периодически по мере накопления архива испытаний (см. [1]). Эти оценки предназначаются для корректировки действующей процедуры контроля таким образом, чтобы в последующем $Q_{\text{вых}}$ процедуры было приемлемо с точки зрения потребителя. Естественно, все это делается в предположении, что входной уровень $Q_{\text{вх}}$ качества производимой продукции остается неизменным, а процесс производства статистически подконтрольным.

В подходе Л. Н. Большева к этой проблеме предлагается изначально строить процедуру контроля с гарантированным уровнем выходного качества ($Q_{\text{вых}} \geq 1 - \alpha_0$, где α_0 — априори заданное малое число), оптимизируя при этом некоторые другие характеристики процедуры (ряд подобных задач решается в [12], [14]). Последующие оценки в такой постановке позволяют строить эмпирические (по Роббинсу) аналоги оптимальных процедур гарантийного контроля при известном априорном распределении контролируемого параметра.

В данной статье делается обзор публикаций авторов за последнее десятилетие, посвященных развитию идей Л. Н. Большева в области контроля качества и аттестации выпускаемой продукции. Кроме теоретических аспектов этого подхода приводятся результаты по решению следующих, представляющих практический интерес, задач: 1) построение процедур контроля с фиксированным объемом испытаний, которые гарантируют заданный уровень $1 - \alpha_0$ выходного качества $Q_{\text{вых}}$ и в то же время максимизируют уровень контроля $Q_{\text{конт}}$ — долю некондиционной продукции среди отклоненной; 2) характеристика и построение процедур с минимальным объемом испытаний, который требуется для достижения заданных уровней $1 - \alpha_0$ и $1 - \alpha_1$ для $Q_{\text{вых}}$ и $Q_{\text{конт}}$; 3) построение процедур аттестации выпускаемой продукции, которые гарантируют вероятность ее кондиционности в случае приобретения потребителем и обладают минимальной вероятностью аттестационной ошибки, состоящей в том, что продукт в действительности кондиционен, но его паспортные данные не удовлетворяют потребителя.

Следующие примеры, один из которых имеет непосредственное отношение к реальной практике машиностроительного производства, иллюстрируют круг задач, обсуждаемых (и частично решаемых) в данном обзоре.

1. Контроль плавки металла по предельной прочности. Плавка металла или металлическое изделие с припуском, из которого

отбираются образцы, контролируется по предельной прочности — минимальной нагрузке на мм^2 , при которой происходит разрыв образца. Металл плавки считается кондиционным, если значение θ его предельной прочности превосходит предписанное значение θ_0 , называемое обычно «нормой». В условиях стабильного производства естественно предположить, что значение θ для каждой конкретной плавки представляет реализацию некоторой случайной величины ϑ и изменчивость значений θ от плавки к плавке обуславливается вариацией условий, при которых производится металл. Точно так же можно считать, что результат x испытания конкретного образца металла есть реализация случайной величины X и изменчивость x от образца к образцу в теле плавки обуславливается, в основном, так называемыми «ошибками испытаний».

Любая процедура контроля состоит из некоторого правила, которое определяет количество отбираемых образцов, и правила принятия решения о кондиционности металла контролируемой плавки по результатам их испытаний. Согласно объявленной выше концепции гарантийности мы считаем основной характеристикой процедуры контроля вероятность $Q_{\text{вых}}$ события $\vartheta \geq \theta_0$ при условии, что плавка металла прошла контроль (принято решение о кондиционности металла). При любом количестве испытуемых образцов правило принятия решения следует выбирать таким образом, чтобы $Q_{\text{вых}}$ было не меньше заданного уровня $1 - \alpha_0$. Вероятность $Q_{\text{вых}}$ обычно называют *выходным уровнем качества* и сопоставляют с *входным уровнем качества* $Q_{\text{вх}}$ — априорной (безусловной) вероятностью события $\vartheta \geq \theta_0$; отношение $Q_{\text{вых}}/Q_{\text{вх}}$ определяет эффективность процедуры контроля. Если обратиться к частотной интерпретации этих характеристик, то $Q_{\text{вх}}$ определяет долю кондиционных плавков, изготавливаемых в процессе производства, а $Q_{\text{вых}}$ — долю кондиционных плавков среди принятых, т.е. отосланных потребителю. Следовательно, $1 - Q_{\text{вых}}$ определяет средний размер выплат по рекламациям. Другой, не менее важной характеристикой контроля является вероятность события $\vartheta < \theta_0$ при условии, что плавка не принята. Эта условная вероятность $Q_{\text{конт}}$ называется *уровнем контроля* и трактуется как доля некондиционных плавков среди отклоненных при контроле.

Процедура контроля называется *гарантийной по выходному уровню качества (уровню контроля)*, если $Q_{\text{вых}} \geq 1 - \alpha_0$ (соответственно $Q_{\text{конт}} \geq 1 - \alpha_1$), где $1 - \alpha_0$ (соответственно $1 - \alpha_1$) — заданное ограничение на указанную характеристику. Процедура контроля называется *гарантийной*, или *полностью гарантийной*, если выполняются оба неравенства $Q_{\text{вых}} \geq 1 - \alpha_0$ и $Q_{\text{конт}} \geq 1 - \alpha_1$.

Естественно, для того, чтобы вычислить $Q_{\text{вых}}$ и $Q_{\text{конт}}$ для действующей процедуры или построить новую процедуру, обладающую каким-либо преимуществом по сравнению с действующей, необходимо знать совместное распределение случайных величин ϑ и X . Информацию об этом распределении предоставляет постепенно накапливаемый архив данных по

испытаниям плавов, так что в принципе о гарантийности процедуры можно говорить только по истечении определенного промежутка времени с начала производства. К данным архива следует также подключить информацию о физических явлениях, определяющих изменчивость θ от плавки к плавке и изменчивость x от образца к образцу, то есть построить вероятностную модель изменчивости предельной прочности и ее измерений (с последующей проверкой адекватности модели данным испытаний).

Прочность металла в каждой плавке при устоявшейся технологии его производства целиком зависит от структурных особенностей его кристаллической решетки, более того, установлено, что предельная прочность прямо пропорциональна числу «правильных» кристаллов в плавке, что позволяет выдвинуть гипотезу о нормальности маргинального распределения случайной величины ϑ , которое в рамках подобных моделей принято называть априорным. Изменчивость же результатов испытаний различных образцов одной и той же плавки обуславливается не только (точнее, не столько) различием числа кристаллических дефектов в испытываемых образцах, а, в основном, их неидентичностью по отношению друг к другу, неравномерностью подачи нагрузки при их растяжении и прочими «ошибками» испытаний и измерений. Поэтому естественно предположить, что условное распределение случайной величины X при каждом фиксированном значении параметра θ также является нормальным.

Итак, мы пришли к гипотезе о справедливости так называемой модели $\mathcal{N}-\mathcal{N}$: распределение ϑ есть $\mathcal{N}(\mu, \tau^2)$ с плотностью

$$g(\theta) = \frac{1}{\sqrt{2\pi}\tau} \exp \left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\},$$

а условное распределение X при $\vartheta = \theta$ есть $\mathcal{N}(\theta, \sigma^2)$ с плотностью

$$f(x | \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(x - \theta)^2}{2\sigma^2} \right\}.$$

В этой модели параметры μ и τ характеризуют общее состояние прочностных свойств металла данной марки, а параметр σ — точностные свойства метода испытаний. Не исключено, что в σ заложена изменчивость прочности по телу плавки, и в таком случае норму θ_0 следует «подправить». Например, рабочую норму $\tilde{\theta}_0$ по заданной θ_0 следует находить из уравнения

$$\mathbf{P}\{X > \theta_0 | \vartheta = \tilde{\theta}_0\} = 1 - \Phi \left(\frac{\theta_0 - \tilde{\theta}_0}{\sigma} \right) = 1 - \gamma,$$

где γ — заданная малая вероятность. Тогда $\tilde{\theta}_0 = \theta_0 + \sigma \Phi^{-1}(1 - \gamma)$. Здесь $\Phi(x)$ — функция стандартного нормального распределения, $\Phi^{-1}(p)$ — квантиль уровня p этого распределения.

Проверку модели $\mathcal{N}-\mathcal{N}$ можно провести по данным архива

$$\begin{array}{ccccccc} x_{11}, & x_{12}, & \dots, & x_{1n_1} & & & \\ \dots\dots\dots & & & & & & \\ x_{m1}, & x_{m2}, & \dots, & x_{mn_m} & & & \end{array}$$

контроля m плавков с испытанием n_k образцов в k -й плавке, $k = 1, \dots, m$. Для этого достаточно сформировать выборку объема m , скажем, из данных испытаний первых образцов $x_{11}, \dots, x_{m1}$ в каждой плавке и проверить ее на нормальность. Согласно модели $\mathcal{N}-\mathcal{N}$ эти данные представляют выборку из нормального распределения с параметрами μ и $\sigma^2 + \tau^2$, и если такая гипотеза подтверждается, то в силу известной теоремы Крамера (см. [13, с. 66, теорема 19]) подтверждается как гипотеза о нормальности априорного распределения ϑ , так и гипотеза о нормальности условного распределения X , то есть подтверждается справедливость модели $\mathcal{N}-\mathcal{N}$.

Мы располагали архивом данных по контролю $m = 560$ плавков. Из каждой плавки первоначально отбирались два образца. Если оба результата испытаний превосходили заданное число C , то наблюдения прекращались. В противном случае дополнительно испытывались еще четыре образца. Таким образом, в нашем архиве каждое n_k , $k = 1, \dots, m$, являлось реализацией случайной величины и равнялось либо 2, либо 6.

Мы проверяли гипотезу о нормальности выборок $x_{11}, \dots, x_{m1}$ и $x_{12}, \dots, x_{m2}$ по критерию хи-квадрат. Согласие с моделью $\mathcal{N}-\mathcal{N}$ (критический уровень значимости) для первой выборки было равно 20%, для второй — 60%.

В настоящее время для контроля металла и металлических изделий по механическим характеристикам на большинстве машиностроительных предприятий используется следующая двухступенчатая процедура контроля: сначала испытываются n_1 образцов (обычно $n_1 = 2$) и полученные данные сравниваются с приемочным числом C . Если $\min_{1 \leq i \leq n_1} x_i \geq C$, то плавка принимается, если же выполняется противоположное неравенство, то дополнительно испытываются $n_2 - n_1$ образцов (обычно 4, так что $n_2 = 6$) и данные $x_{n_1+1}, \dots, x_{n_2}$ испытаний этих образцов сравниваются с тем же числом C . Если $\min_{n_1+1 \leq i \leq n_2} x_i \geq C$, то плавка принимается, в противном же случае принимается решение о некондиционности плавки по значению наблюдаемой механической характеристики.

Вероятность приемки плавки с прочностью θ при нормальном условном распределении наблюдаемой случайной величины равна

$$\begin{aligned} L(\theta) = & \mathbf{P}_\theta(X_1 \geq C, \dots, X_{n_1} \geq C) + [1 - \mathbf{P}_\theta(X_1 \geq C, \dots, X_{n_1} \geq C)] \\ & \times \mathbf{P}_\theta(X_{n_1+1} \geq C, \dots, X_{n_2} \geq C) = \left[1 - \Phi\left(\frac{C - \theta}{\sigma}\right) \right]^{n_1} \end{aligned}$$

$$+ \left[1 - \left(1 - \Phi \left(\frac{C - \theta}{\sigma} \right) \right)^{n_1} \right] \left[1 - \Phi \left(\frac{C - \theta}{\sigma} \right) \right]^{n_2 - n_1}.$$

Функцию $L(\theta)$ обычно называют *оперативной характеристикой* процедуры контроля. Другие, более важные с точки зрения нашего подхода к проблеме гарантийности, характеристики двухступенчатой процедуры вычисляются с использованием формулы Байеса:

$$\begin{aligned} Q_{\text{вх}} &= \mathbf{P}(\vartheta \geq \theta_0) = 1 - \Phi \left(\frac{\theta_0 - \mu}{\tau} \right), \\ Q_{\text{вых}} &= \mathbf{P}(\vartheta \geq \theta_0 \mid \text{плавка принята}) \\ &= \frac{\frac{1}{\sqrt{2\pi}\tau} \int_{\theta_0}^{\infty} L(\theta) \exp \left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\} d\theta}{\frac{1}{\sqrt{2\pi}\tau} \int_{-\infty}^{\infty} L(\theta) \exp \left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\} d\theta}, \\ Q_{\text{конт}} &= \mathbf{P}(\vartheta < \theta_0 \mid \text{плавка отклонена}) \\ &= \frac{\frac{1}{\sqrt{2\pi}\tau} \int_{-\infty}^{\theta_0} (1 - L(\theta)) \exp \left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\} d\theta}{\frac{1}{\sqrt{2\pi}\tau} \int_{-\infty}^{\infty} (1 - L(\theta)) \exp \left\{ -\frac{(\theta - \mu)^2}{2\tau^2} \right\} d\theta}. \end{aligned}$$

Для оценки параметров μ , τ^2 и σ^2 использовался тот же, что и при проверке модели $\mathcal{N}-\mathcal{N}$, архив данных. По результатам испытаний только первых двух образцов каждой плавки в рамках сбалансированной однофакторной модели II дисперсионного анализа со случайными факторами (см., например, [3, § 10.5, с. 293–295]) получены следующие значения оценок для параметров априорного распределения: $\hat{\mu} = 122,07$, $\hat{\tau} = 2,67$, $\hat{\sigma} = 1,18$.

При норме $\theta_0 = 119,00$ подстановка этих значений в уравнения для $Q_{\text{вх}}$, $Q_{\text{вых}}$ и $Q_{\text{конт}}$ (см. выше) дала следующие значения состоятельных (объем архива $m \rightarrow \infty$) оценок характеристик двухступенчатой ($n_1 = 2$, $n_2 = 6$, $C = 119,00$) процедуры контроля: $\hat{Q}_{\text{вх}} = 0,875$, $\hat{Q}_{\text{вых}} = 0,989$, $\hat{Q}_{\text{конт}} = 0,595$.

Если отказаться от испытаний дополнительных 4-х образцов на второй ступени и использовать для контроля одноступенчатую процедуру по минимальному значению в выборке из двух наблюдений, то $\hat{Q}_{\text{вых}} = 0,989$, $\hat{Q}_{\text{конт}} = 0,587$. Следовательно, введение второй ступени практически не оказывает влияния на характеристики процедуры.

Как показано в [17] (см. также пример после теоремы 2.3 в § 2), в случае нормального распределения величины X при фиксированном θ и и любого априорного распределения ϑ процедура с фиксированным объемом испытаний n , гарантирующая заданный уровень $1 - \alpha_0$ выходного качества и максимизирующая уровень контроля $Q_{\text{конт}}$, должна быть

основана на выборочном среднем $\bar{x}$: плавка принимается, если $\bar{x} > C$, где приемочное число C определяется из уравнения $Q_{\text{вых}}(C) = 1 - \alpha_0$. В случае модели $\mathcal{N}-\mathcal{N}$ с использованием оценок параметров априорного распределения вместо их точных (неизвестных) значений такая процедура асимптотически гарантирует заданный уровень $1 - \alpha_0$ для $Q_{\text{вых}}$ и асимптотически максимизирует величину $Q_{\text{конт}}$ при неограниченном возрастании объема архива m . Для данных нашего архива с $n = 2$ и $1 - \alpha_0 = 0,99$ (то есть для фактического уровня выходного качества для рассмотренной выше двухступенчатой процедуры) приемочное число оптимальной процедуры $C = 119,5$, а $Q_{\text{конт}} = 0,639$.

Процедура контроля, основанная на $\bar{x}$, является также оптимальной (асимптотически оптимальной) среди всех полностью гарантийных процедур с фиксированным объемом испытаний. Если потребовать $1 - \alpha_0 = 0,99$, $1 - \alpha_1 = 0,90$, то по данным нашего архива требуемое число испытаний $n = 12$, а приемочное число $C = 119,02$; при ограничениях $1 - \alpha_0 = 0,95$ и $1 - \alpha_1 = 0,90$ необходимый объем выборки $n = 2$, а $C = 118,32$ (!). Мы обсудим этот замечательный факт (приемочное число ниже нормы) в § 5; там же см. продолжение данного примера в части гарантийной аттестации плавки металла по предельной прочности.

О последовательных гарантийных процедурах контроля, обладающих высокой вероятностью остановки после испытаний первого образца, см. § 4.

Отметим, что приведенные выше расчеты носят в основном иллюстративный характер. В действительности норма $\theta_0 = 110$, так что оценка $\hat{Q}_{\text{вх}} = 0,999$, из чего можно сделать вывод о нецелесообразности контроля плавки по предельной прочности по крайней мере до тех пор, пока технологический процесс статистически подконтролен. Однако, на тех же образцах в процессе испытаний замеряются еще три механические характеристики металла, две из которых (относительные сужение и удлинение) имеют резко асимметричные априорные распределения, так что гарантийный контроль качества в рамках модели $\mathcal{N}-\mathcal{N}$ становится невозможным.

2. Контроль по альтернативному признаку. Классическая вероятностная модель контроля по альтернативному (или, как иногда говорят, по качественному) признаку связана со случайным отбором изделий из партии конечного объема N с последующим разделением отобранных изделий на два класса — годные и дефектные. Общее число D дефектных изделий в партии неизвестно; имеется ограничение θ_0 на допустимую долю $\theta = D/N$ дефектных изделий, и статистическая задача контроля качества состоит в различении двух гипотез: $\theta > \theta_0$ (партия отклоняется от приемки) и $\theta \leq \theta_0$ (партия принимается — отсылается потребителю).

Если формирование партий изделий не соотносится с технологическими циклами в их изготовлении, а происходит путем простого накоп-

ления определенного количества «штампующих» изделий, то вполне разумно выдвинуть гипотезу о биномиальном априорном распределении величины D . При справедливости такой гипотезы каждое изделие с одинаковой вероятностью p может быть дефектным, причем, пока действует выбранная процедура контроля, величина p должна оставаться постоянной. Различные планы контроля в рамках такой гипергеометрическо-биномиальной модели подробно разбираются в монографии [1] (советуем особо обратить внимание на утверждения § 2.3). Замечательно, что в рамках этой модели $Q_{\text{вх}} = Q_{\text{вых}} = p$ при любой процедуре контроля, то есть любой контроль (кроме, естественно, контроля за стабильностью технологического процесса производства изделий) обесмысливается — он не способен изменить долю дефектных изделий в принятой партии. К сожалению, до сих пор на практике бытует большое число стандартизированных методик контроля по альтернативному признаку, в которых умалчивается об априорном распределении D , а гарантийность контроля соотносится с ограничениями на риски потребителя и изготовителя.

Существенно другая вероятностная картина возникает в ситуации, когда формирование партии изделий соотносится с определенным технологическим циклом, например, когда партия содержит металлические изделия одной и той же плавки и контролируемый признак связан с прочностными свойствами металла. Рассмотрим, для простоты, случай, когда объем N партии значительно превосходит объем инспектируемых изделий. В таком случае процедура контроля осуществляется в схеме испытаний Бернулли с постоянной вероятностью θ «успешного» испытания (вероятностью дефектности наугад взятого образца из партии изделий). Решение о принятии партии признается безошибочным, если $\theta \leq \theta_0$, где θ_0 — предельная допустимая доля дефектных изделий в партии.

В отличие от предыдущих примеров в данной ситуации зачастую не представляется возможным использовать какие-либо физические законы при спецификации априорного распределения θ . Однако на отрезке $[0,1]$ возможных значений θ существует удивительно богатое семейство распределений, члены которого могут обладать правой или левой асимметрией, отрицательным или положительным эксцессом, могут иметь U -образную форму и пр. Это семейство бета-распределений с плотностью

$$g(\theta) = \frac{1}{B(\lambda_1, \lambda_2)} \theta^{\lambda_1-1} (1-\theta)^{\lambda_2-1}, \quad 0 \leq \theta \leq 1; \quad \lambda_1 > 0, \quad \lambda_2 > 0,$$

(об этой модели и других, используемых в практике контроля, см. монографию [1]). Параметры λ_1 и λ_2 бета-распределения довольно просто оцениваются по архиву данных с помощью метода моментов.

Как следует из результатов статьи [17] (см. также § 2 данного обзора), при любом (последовательном, многоступенчатом и пр.) плане

отбора изделий из партии решение о ее приемке должно быть основано на числе X дефектных изделий в выборке: партия принимается, если $X \leq C$. При фиксированном плане отбора приемочное число C определяется по заданному ограничению $1 - \alpha_0$ на $Q_{\text{вых}}$; полностью гарантийная процедура требует спецификации плана отбора и определения C по совокупности ограничений $1 - \alpha_0$ и $1 - \alpha_1$ на $Q_{\text{вых}}$ и $Q_{\text{конт}}$ соответственно.

Отметим, что в этом примере при оптимизации характеристик контроля, как и в теории Неймана–Пирсона (см. [10, гл. 9]), существенную роль играет рандомизация принятия решения о приемке партии в случае $X = C$. Более подробное описание рандомизационной процедуры контроля в рамках модели Би–Ве (биномиальное распределение X и бета-распределение ϑ) с выводом приближенных формул для расчета n и C в случае, когда априорное бета-распределение имеет ярко выраженный максимум в граничной точке θ_0 , дано в § 4 данного обзора.

Приведенные примеры дают представление лишь о той части нашего обзора, которая касается приложений апостериорного подхода к проблемам гарантийности контроля качества, в то время как его основное содержание посвящено математическим обоснованиям понятия асимптотической (по мере накопления архива) гарантийности и построению оптимальных процедур при «известном априорном распределении». Мы выделили в кавычки последнюю фразу, понимая, что это исключительно редкий случай в практике контроля качества, но это тот фундамент, как критерий Неймана–Пирсона различения простых гипотез, опираясь на который, можно строить эмпирические процедуры контроля, основанные на оценках априорного распределения по архиву данных и обладающие определенной асимптотической оптимальностью.

Наши публикации в этой области носят, в основном, весьма абстрактный характер и направлены на построение общей теории статистического вывода на последовательностях статистических экспериментов (см. для сравнения близкую нам в идейном плане статью Хилла [18]). Статистический контроль качества — это тот объект приложения наших построений, в котором статистик действительно имеет дело с реальной (а не воображаемой, как в большинстве «научных» приложений) последовательностью статистических экспериментов. Поэтому основная цель данного обзора — изложить общую теорию апостериорного подхода применительно к простому и наглядному объекту статистических исследований.

В § 1 вводится понятие адаптивной стратегии — последовательности статистических процедур, применяемых для обследования последовательно поступающих объектов контроля. Каждая процедура определяется правилом останова статистического эксперимента при контроле очередного объекта и правилом принятия решения о его кондиционности. Эти правила зависят как от результатов текущего эксперимента, так и от данных обследования предшествующих объектов контроля. Такие

данные образуют «архив», постепенное накопление которого дает возможность с возрастающей точностью и надежностью оценить априорное распределение контролируемого параметра и тем самым обеспечить гарантийность адаптивной стратегии контроля качества.

Понятие гарантийности основано на определении d -рисков стратегии: имея дело с реальной последовательностью статистических экспериментов и принимаемых решений, мы вводим d -риски как доли ошибочных решений в подпоследовательностях экспериментов, завершаемых соответственно принятием (решение d^0) или отклонением (решение d^1) контролируемой продукции. Стратегия является гарантийной, если эти доли не превосходят заданных уровней α_0 и α_1 , или, что то же, $Q_{\text{вых}} \geq 1 - \alpha_0$, $Q_{\text{конт}} \geq 1 - \alpha_1$.

В этом же параграфе дается постановка задачи на оптимальный выбор стратегии и указывается схема решения такой задачи. По своей сути эта схема восходит к эмпирическому байесовскому подходу Роббинса. Сначала мы решаем задачу на оптимум при известном априорном распределении, а затем строим эмпирический, основанный на оценках d -рисков по архиву данных, аналог «простой» оптимальной стратегии, то есть стратегии, состоящей из последовательности одинаковых процедур с правилами, зависящими от данных только текущего эксперимента. Таким образом, простая стратегия является по существу обычной процедурой статистического вывода, оптимизация которой открывает путь к построению оптимальной стратегии контроля.

В § 2 рассматривается задача оптимизации простых стратегий — процедур контроля — при известном априорном распределении. Устанавливается своеобразный аналог «леммы Неймана–Пирсона», в котором указывается вид правила принятия решения, гарантирующего при любом заданном правиле остановки требуемый уровень $1 - \alpha_0$ выходного качества и максимизирующий $Q_{\text{конт}}$. Понятно, что это решает проблему построения гарантийной процедуры контроля с минимальным объемом выборки в классе всех процедур, правила остановки которых не зависят от результатов наблюдений (процедур с фиксированным объемом испытаний). Для некоторых частных случаев в рамках модели $\mathcal{N}-\mathcal{N}$ (см. п. 1 введения) выводятся явные формулы для необходимого объема выборочных обследований.

Что же касается минимизации среднего объема наблюдений в классе всевозможных последовательных гарантийных процедур контроля, то этот аналог классической задачи Вальда до сих пор не имеет решения. В § 2 устанавливаются параллели с соответствующими последовательными процедурами отношения правдоподобия в классической теории проверки гипотез — такой параллелизм проливает свет на основную трудность в решении задачи Вальда в рамках апостериорного подхода к проблеме гарантийности процедур контроля качества.

В § 3 устанавливается правомочность выбранной схемы построения

оптимальных стратегий: показывается, что введение эмпирических аналогов простых стратегий позволяет строить оптимальные адаптивные стратегии в соответствии с определениями § 1. Основное содержание параграфа посвящено анализу асимптотической (при неограниченном возрастании объема архива) гарантийности эмпирических стратегий и наследованию оптимальных свойств адаптивными стратегиями, «порожденными» эмпирическими аналогами оптимальных простых стратегий.

В § 4 мы возвращаемся к простым стратегиям и исследуем асимптотику необходимого объема наблюдений в случае, когда априорное распределение имеет резко выраженный максимум в точке θ_0 , разграничивающей проверяемые гипотезы $\theta \leq \theta_0$ и $\theta > \theta_0$, то есть когда априорное распределение сосредотачивается на границе «нормы» и неопределенность в выборе решения о приемке продукции достигает своего апогея. Этот случай аналогичен схеме сближающихся гипотез (континуальных альтернатив) в классической теории статистического вывода, и может быть именно поэтому здесь удастся продвинуться в решении оптимальных задач значительно дальше, чем в § 2. Находится асимптотика необходимого объема наблюдений в классе гарантийных процедур с фиксированным числом испытаний. В классе всех последовательных гарантийных процедур приводится суррогат асимптотического решения задачи Вальда — минимизируется не средний объем наблюдений, а некоторые функционалы, характеризующие количество информации, поступающей из данных эксперимента в пользу гипотезы кондиционности (соответственно, некондиционности) контролируемой продукции.

Наконец, § 5 посвящен вопросам аттестации (паспортизации) выпускаемой продукции. В этом разделе мы пунктуально следуем идеям С. Н. Бернштейна [2] в постановке задачи доверительного оценивания; решение проблемы оптимальной доверительной аттестации есть компромисс между требованиями покупателя и желанием производителя сбыть максимум продукции. Для простоты рассматривается только задача построения верхней доверительной границы. Как и в классическом случае, в апостериорном подходе задача доверительного оценивания двойственна задаче проверки гипотез: периодически приобретая выпускаемую продукцию, покупатель должен быть уверен, что доля «хорошей» продукции среди приобретенной не опускается (в среднем) ниже номинального уровня $1 - \alpha_0$, в то время как производитель стремится минимизировать долю «хорошей» продукции среди отклоненной покупателем по причине ошибочной аттестации. В такой постановке задача оптимизации доверительного оценивания по существу решена в § 2. Асимптотические результаты § 4 дают решения соответствующих задач по построению асимптотически оптимальных доверительных границ. Исследуются также соотношения найденных границ с байесовскими доверительными границами, которые определяются как квантили апостери-

орного распределения. Байесовские границы далеки от оптимальных, что особенно заметно при больших объемах испытаний.

Данный обзор посвящен только результатам, имеющим отношение к проблемам контроля качества и аттестации выпускаемой продукции. О других исследованиях в области апостериорного подхода к проблеме статистического вывода см. [6], [7], [19], [20]. В своем обзоре мы не касаемся других многочисленных исследований по проблемам контроля качества. Традиционная тематика этой области приложений математической статистики достаточно полно отражена в монографии [1]: новые исследования, за исключением разве лишь применений теории полезности и субъективного байесовского вывода, лежат в русле идей, затронутых в этой книге.