СОДЕРЖАНИЕ

Предисловие

Книга разделена на части, каждую из которых можно читать независимо от других. Тем не менее именно в такой последовательности, как расположены части и главы книги, следует ее изучать.

Настоящая книга не предназначена для беглого просмотра, а требует от читателя значительного внимания. Изложение материала требует от читателя знания основ математической статистики и классической теории надежности.

Книга может быть использована и для справок по эффективным оценкам показателей надежности различных планов испытаний.

Понимая, что книга не свободна от недостатков, авторы с признательностью примут замечания и предложения по улучшению содержания и формы изложения материала.

Введение

В математической статистике разработаны различные методы для построения точечных оценок параметров законов распределений вероятностей случайных величин: метод моментов, метод максимального правдоподобия и метод минимума расстояний [1, 2]. На практике, используя эти методы, не всегда удается построить несмещенную и эффективную оценку, если таковая существует. В общем случае правила нахождения несмещенных оценок в настоящее время не существует, и их определение требует своего рода искусства. В ряде случаев найденные несмещенные и эффективные оценки имеют весьма громоздкий вид со сложным алгоритмом вычисления [3]. Они также не всегда являются достаточно эффективными в классе всех смещенных оценок, т. е. не всегда имеют значительное преимущество перед простыми, но смещенными оценками, с точки зрения близости к оцениваемому показателю. Существующая проблема вполне решаема с помощью интегрального оценивания.

ЧАСТЬ I

ОСНОВНЫЕ ПОНЯТИЯ И МЕТОДЫ ПОСТРОЕНИЯ СТАТИСТИЧЕСКИХ ОЦЕНОК

1.1. Методы построения статистических оценок

Для определенности, не нарушая общности рассуждений, будем в основном рассматривать биномиальные испытания (план типа $NБ\tau$) и испытания с ограниченным временем испытаний и восстановлением (план типа $NB\tau$), где N – число испытуемых однотипных изделий ($N = n$ – число первоначально выставленных изделий); τ – наработка (одинаковая для каждого изделия); B – характеристика плана, означающая, что работоспособность изделия после каждого отказа в течение срока испытаний восстанавливается; $Б$ – характеристика плана, означающая, что работоспособность изделия после каждого отказа в течение срока испытаний не восстанавливается [4, 5]. При этом там, где это необходимо, будем считать, что наработка до отказа изделий подчиняется экспоненциальному закону распределений (далее з.р.) вероятностей случайных величин (далее – с.в.) с параметром T_0, где последний совпадает со средней наработкой до отказа (СНДО). Тогда расчетное значение вероятности безотказной работы (ВБР) одного изделия за заданное время τ будет определяться равенством

$$P_0(\tau) = e^{(-\tau/T_0)}. \tag{1.1}$$

Заметим, что плану испытаний типа $N\mathrm{B}\tau$ соответствует распределение Пуассона [4, 5], а плану типа $NБ\tau$ соответствует биномиальное распределение [4, 5].

Обозначим случайное число отказов через R, тогда для плана испытаний типа $NB\tau$ достаточной статистикой является число наблюденных отказов ($R = r$) [1–5]. Для плана испытаний типа $NB\tau$ случайная величина R, имеет пуассоновское распределение

$L(R \le r; \Delta)$ с параметром $\Delta = N\tau/T_0$ [1, 2]. Тогда, по определению, r – реализация с.в. R. С другой стороны, R – сумма с.в. X_i, каждая из которых есть случайное число отказов одного из N изделий $(1 \le i \le N)$, поставленных на испытания. Случайные величины X_i имеют пуассоновское распределение с параметром Δ/N, а их сумма определяет пуассоновское распределение $L(R \le r; \Delta)$ суммарного потока отказов [1, 2]:

$$L(R \le r;\Delta) = \sum\nolimits_{i=0}^{X_1+\ldots+X_N=r} e^{-\Delta}\frac{\Delta^i}{i!}. \tag{1.2}$$

Для биномиальных испытаний (план типа $NБ\tau$) достаточной статистикой является число наблюдаемых отказов ($R = r$) и суммарная наработка $S(R = r, \tau, s_i)$ [1–5], где R – случайное число отказов, s_i – моменты отказов, $i = 1, 2, ..., r$. Для биномиальных испытаний с.в. R, имеет биномиальное распределение $b_N(r)$ [2, ф. 1.4.55] с параметрами N и p, $0 \le p \le 1$, т. е. с. в. R, равная числу успехов в серии из N независимых опытов, принимает целочисленные значения 0, 1, 2, ..., N с вероятностями:

$$b_N(r) = C_N^r p^r (1-p)^{N-r}. \tag{1.3}$$

Функция распределения F_R $(R \le r,\ N,\ p)$ биномиальной с.в. R примет вид

$$F_R(R \le r, N, p) = \sum\nolimits_{k=0}^{r} b_N(k). \tag{1.4}$$

Функция распределения $F_R(R \le r,\ N,\ p)$ вычисляется через неполную бета-функцию $I_p(x, y)$ по формуле [2, ф. 1.4.57]:

$$F_R(r, N, p) = 1 - I_p(r+1, N-r) = I_{1-p}(N-r, r+1). \tag{1.5}$$

Вероятности $b_N(k)$ вычисляются через неполную бета-функцию $I_p(x, y)$ по формуле [2, ф. 1.4.58]:

$$b_N(k) = I_p(k, N-k+1) - I_p(k+1, N-k). \tag{1.6}$$

Заметим, что традиционная оценка параметра p биномиального з.р. $\hat{p}(R,\ N) = R/N$ является несмещенной и эффективной оценкой [2, пример 2.4.20]. Оценка $\hat{p}$ также является и оценкой максимального правдоподобия [2, пример 2.10.7].

Определим кратко наиболее часто встречающиеся критерии эффективности оценок [1, 2] и их отличия. В основе этих критериев лежит среднеквадратический подход сравнения оценок. Пусть T_0 не является с.в. и принадлежит множеству значений $T_0 \in G$. Для функции от параметра $\theta(T_0)$ оценка $\hat{\theta}_0(R)$ называется эффективной оценкой в классе оценок $\hat{\theta}_0 \in \Theta$, если для любой другой оценки $\hat{\theta}_0(R)$ из этого класса выполняется неравенство

$$E\left(\hat{\theta}_0\,(R) - \theta(T_0)\right)^2 \le E\left(\hat{\theta}(R) - \theta(T_0)\right)^2,$$

где E – математическое ожидание, соответствующее з.р. числа отказов для параметра $T_0 \in G$. То есть сравниваются две оценки, одна из которых после их сравнения признается эффективнее другой.

1.2. Минимаксный подход

Оценка называется минимаксной $\hat{\theta}_0(R)$, если для любой другой оценки $\hat{\theta}(R)$ неслучайного параметра $t \in G$ выполняется неравенство [1, 2]:

$$\sup_{t\in G} E_t(\hat{\theta}_0(R) - \theta(t))^2 \le \sup_{t\in G} E_t(\hat{\theta}(R) - \theta(t))^2\,. \qquad (1.7)$$

Из минимаксного подхода следует, что всегда найдется вариант, когда минимаксная оценка $\hat{\theta}_0(R)$ является лучшей только в ближайшем диапазоне $t_0 \in [t_0 - \varepsilon,\ t_0 + \varepsilon]$ минимизации самого худшего случая уклонения от оцениваемого параметра $\theta(t)$, который составляет небольшую долю рабочего диапазона $t \in [t_1,\ t_2]$. И, в то же время, уклонения минимаксной оценки $\hat{\theta}_0(R)$ могут превышать уклонения других оценок $\hat{\theta}(R)$ в более обширном рабочем диапазоне значений оцениваемого параметра $t_0 \notin [t_1,\ t_2]$ (НЕ максимального уклонения). Хотя уклонения этих оценок $\hat{\theta}(R)$ и превышают худший случай минимаксной оценки $\hat{\theta}_0(R)$ в ближайшем диапазоне $t_0 \in [t_0 - \varepsilon,\ t_0 + \varepsilon]$ ($t_0 \notin [t_1,\ t_2]$), но зато минимаксная оценка проигрывает в другом более обширном рабочем диапазоне, где ее уклонения превышают уклонения оценок $\hat{\theta}(R)$, и в этом диапазоне (НЕ максимального уклонения) минимаксная оценка $\hat{\theta}_0(R)$ теряет свою эффективность.

1.3. Байесовский подход

Суть байесовского подхода состоит в том, что неизвестный (оцениваемый) параметр T_0 (или функция от параметра $\theta(T_0)$) рассматривается как случайная величина с некоторой плотностью распределения $q(t)$, где t – реализация с. в. T_0 [1, 6]. Плотность $q(t)$ называется априорной, т. е. данной до эксперимента. Байесовский подход предполагает, что неизвестный параметр T_0 был выбран случайным образом из распределения с плотностью $q(t)$.

В соответствии с формулой Байеса плотность апостериорного (после эксперимента) распределения имеет вид [1]

$$q(t / R) = \frac{f_\Delta(r)q(t)}{f(r)}, \qquad (1.8)$$

где $f(r) = \int f_\theta(r)q(t)dt$.

Само апостериорное распределение параметра $\theta(T_0)$ будем обозначать через Q_R. Тогда байесовская оценка, соответствующая априорному распределению Q с плотностью $q(t)$, имеет вид

$$\hat{\theta}_Q(R) = E(\theta(T_0) \mid R) = \int \theta(t)q(t \mid R)dt = \int \theta(t)Q_R(dt). \qquad (1.9)$$

В силу свойств условного математического ожидания байесовская оценка минимизирует среднеквадратическое уклонение $E(\hat{\theta}_Q(R) - \theta(T_0))^2$. Для сравнения байесовской оценки на множестве других оценок $\hat{\theta}(R)$ выполняется неравенство

$$E(\hat{\theta}_Q(R) - (\theta(T_0))^2 \le E(\hat{\theta}(R) - (\theta(T_0))^2 =$$
$$= \int E_t(\hat{\theta}(R) - (\theta(t))^2 q(t)dt. \qquad (1.10)$$

Отметим еще раз, что для байесовской оценки безусловное среднеквадратическое уклонение (см. формулу (1.10))

$$E(\hat{\theta}_Q(R) - (\theta(T_0))^2 = \int E_t(\hat{\theta}_Q(R) - \theta(t))^2 q(t)dt \qquad (1.11)$$

принимает наименьшее возможное значение. Соотношение (1.9) показывает, что байесовская оценка минимизирует среднее значение. Недостатком байесовского подхода является обязательное знание плотности априорного з.р. случайного параметра T_0 (см. формулы (1.8)–(1.11)). С одной стороны, эти заложенные в правило предварительные знания несут в себе однократные финансовые издержки, а с другой – позволяют минимизировать объем испытаний [6], что в рамках стабильного производства дает им конкурентные преимущества [6].

Отметим полезные связи между минимаксными и байесовскими оценками. Если существует оценка θ_1 и распределение Q такие, что при всех t выполняется неравенство

$$E(\hat{\theta}_1(R)-\theta(t))^2=\int E_t(\hat{\theta}_Q(R)-\theta(t))^2 q(t)dt,$$

то оценка θ_1 – минимаксная [1]. В действительности всегда выполняется равенство, и в этом случае байесовская оценка является минимаксной [1].

1.4. Интегральный подход в процессе поиска эффективных оценок

Интегральный подход отработан в основном для планов испытаний типа $NB\tau$ [7–13], $NБ\tau$ [14–18] и плана испытаний с добавлением [19, 20].

В основе интегрального подхода лежит построение правила выбора эффективной оценки $\hat{\theta}_0\,(R)$, заданного на сумме значений относительных смещений оценок $\hat{\theta}(R)$ от функции над параметром з.р. $\theta(T_0=t)$, а именно $b/\theta(t) = (\mathrm{E}(\hat{\theta}(R)) - \theta(t))/(\theta(t))$. В этом случае самым разумным является построение критерия выбора эффективной оценки $\hat{\theta}_0\,(R)$ на множестве оценок $\hat{\theta}(R, N, \tau)\in\Theta$, основанном на суммарном квадрате относительных смещений математического ожидания исследуемых оценок $E\hat{\theta}(R, N, \tau)$ от функции над параметром $\theta(T_0)$ для всех возможных значений T_0, N и τ.

Например, на основе изложенного для плана испытаний типа $NB\tau$ [7–13] в качестве критерия получения эффективной оценки можно построить функционал (далее – $A(\hat{\theta})$):

$$A(\hat{\theta})=\int_0^\infty\left(\frac{b(\theta)}{\theta(T_0)}\right)^2\partial\Delta, \qquad (1.12)$$

где $\Delta = N\tau/T_0$ – параметр пуассоновского з.р., характеризующий поток отказов [1]; $T_0 = N\tau/\Delta$, $b(\theta) = \{E\hat{\theta}(R, N, \tau) - \theta(T_0)\}^2$ – смещение, $t\in[t_1,\ \ t_2\,]$.

Воспользовавшись свойствами пуассоновского потока с параметром Δ, определим формулу математического ожидания оценок $\hat{\theta}(R, N, \tau)$ [1]:

$$E\hat{\theta}(R,N,\tau) = \sum\nolimits_{k=0}^{\infty} \hat{\theta}(R,N,\tau)e^{-\Delta}\frac{\Delta^k}{k!}. \qquad (1.13)$$

Эффективная оценка $\hat{\theta}_0(R)$ должна обладать минимальной величиной функционала $A(\hat{\theta}_0)$. Из определения интегральной оценки следует, что ее выбор основан на минимизации суммы квадратов относительных усредненных смещений от оцениваемого параметра (или функции от параметра) на всем диапазоне значений, принимаемых этим параметром, и на всем диапазоне значений, которые могут принимать количество испытуемых изделий N и время испытаний τ.

Таким образом, интегральный подход учитывает все факторы, влияющие на выбор эффективной оценки. Интегральный подход наиболее интересен в случае, когда оценки $\hat{\theta}(R, N, \tau)$ принадлежат классу смещенных оценок $b \geq 0$. Эффективные оценки, полученные минимизацией функционала (формула (1.12)), будем называть интегральными эффективными оценками по смещению (или просто – эффективными оценками). Для несмещенных оценок существует классический вариант поиска эффективных оценок – по уклонению оценки параметра от его истинного значения [1, 2].

Дополнительным критерием поиска эффективных оценок, использующим интегральный подход, является минимизация функционала (далее – $B(\hat{\theta})$), основанного на суммировании математических ожиданий квадратов относительных уклонений оценок $\hat{\theta}(R, N, \tau)$ от функции над параметром $\theta(T_0)$ для всех возможных значений T_0, N и τ, а именно

$$B(\hat{\theta}) = \int\limits_0^{\infty} \left(\frac{1}{\theta(T_0)}\right)^2 E\{\hat{\theta}(R,N,\tau) - \theta(T_0)\}^2 \partial\Delta. \qquad (1.14)$$

Отличие оценок, эффективных по уклонению (например байесовские оценки ($\hat{\theta}_Q(R)$), от интегральной оценки, эффективной по смещению, выражено равенством [1]:

$$E(\hat{\theta}_Q(R) - \theta(T_0))^2 = D\left(\hat{\theta}_Q(R)\right) + b^2,$$

где $D(\hat{\theta}_Q(R))$ – дисперсия. То есть байесовская оценка $\hat{\theta}_Q(R)$ минимизирует среднеквадратическое уклонение за счет минимальной суммы дисперсии и квадрата смещения. Однако во многих случаях в классе смещенных оценок можно найти оценку, у которой уклонение от параметра (функции от параметра) меньше, чем у несмещенной и эффективной оценки по уклонению. Во всех случаях, когда существует эффективная (несмещенная) оценка, существует смещенная оценка более точная, чем эффективная, т. е. с меньшим квадратом ошибки. Однако смещенными оценками обычно не пользуются, чтобы избежать систематических ошибок при небольшом числе опытов N. При большом же N заметного выигрыша в точности по сравнению с эффективной (несмещенной) оценкой не получается. Поэтому эффективными несмещенными оценками пользуются всегда, когда они существуют [70, с. 284].

Для задач интегрального оценивания по смещению важно (первично) не минимальное рассеивание оценок от параметра, а минимальное смещение. Таким образом, классические оценки (несмещенные и эффективные по уклонению) и интегральные оценки (эффективные по смещению) решают задачи, в основе решения которых лежит одна и та же числовая характеристика точности оценки – среднеквадратическое отклонение. Совместное решение этих задач осуществляет поиск эффективных оценок.

Рассмотрим общий случай. Из построения формул (1.12) и (1.14) следует, что для различных планов испытаний в общем случае для множества значений переменных $N \in [1; 10]$ и $\tau \in [\tau_1; \tau_2]$ минимизация функционалов $A(\hat{\theta}(R, N, \tau_i))$ и $B(\hat{\theta}(R, N, \tau_i))$ даст множество частных эффективных оценок $\hat{\theta}(R, N, \tau_i)$. Чтобы найти усредненную эффективную оценку, необходимо ее поиск осуществлять минимизацией расширенных функционалов следующего вида:

$$A_{10}\left(\hat{\theta}(R, N, \tau)\right) = \frac{1}{10}\sum\nolimits_{N=1}^{10} \lim_{\delta_\tau \to 0}\frac{1}{I}\sum\nolimits_{i=1}^{I} A_{N,\tau_i}(\hat{\theta}), \qquad (1.15)$$

$$B_{10}\left(\hat{\theta}(R, N, \tau)\right) = \frac{1}{10}\sum\nolimits_{N=1}^{10} \lim_{\delta_\tau \to 0}\frac{1}{I}\sum\nolimits_{i=1}^{I} B_{N,\tau_i}(\hat{\theta}), \qquad (1.16)$$

где

$$A_{N,\tau_i}(\hat{\theta}) = \int_0^\infty \frac{1}{\theta^2}\{E\hat{\theta}(R, N, \tau_i) - \theta\}^2 \partial\theta,$$

$$B_{N,\tau_i}(\hat{\theta}) = \int_0^\infty \frac{1}{\theta^2}E\{\hat{\theta}(R, N, \tau_i) - \theta\}^2 \partial\theta,$$

δ_τ – шаг суммирования; $I = (\tau_2 - \tau_1)/\delta_\tau$ – число шагов суммирования.

Так как величины функционалов $A_{10}(\hat{\theta}(R, N, \tau))$ и $B_{10}(\hat{\theta}(R, N, \tau))$ с изменением объема испытаний и границ интервала τ_1 и τ_2 могут стремиться как к нулю, так и к бесконечности, то следует ограничиваться их рабочим диапазоном. Реальное время испытаний может колебаться в пределах от 500 до 100 000 ч, а объем N от 1 до 10 изделий, в зависимости от сложности и надежности испытуемого объекта. Именно этот фиксированный интервал следует рассматривать в качестве эталона при вычислении (минимизации) функционалов $A_{10}(\hat{\theta}(R, N, \tau))$ и $B_{10}(\hat{\theta}(R, N, \tau))$.

На практике шаг суммирования δ_τ следует выбирать конечной величиной, достаточной для получения величин функционалов с приемлемой точностью. Заметим, что при вычислениях варьирование шагом суммирования приводит к изменению результата функционала, но не меняет сути вещей – результат сравнения оценок не меняется.

Идеальным вариантом в задачах оценивания является использование несмещенной оценки с минимальным уклонением, если такая оценка существует. В противном случае следует искать оценки с минимальным смещением, а среди них – с минимальным уклонением [8–11]. Такой процесс поиска гарантирует получение оценок с хорошими точностными характеристиками. Поэтому, в общем случае, не следует ориентироваться на смещенные оценки, построенные минимизацией только функционала вида $B_{10}(\hat{\theta}(R, N, \tau))$ (см. формулу (1.16)). Заметим, что опыт построения эффективных оценок показывает, что полученная несмещенная эффективная оценка не всегда будет обладать минимальным уклонением [14–18]. Скорее наоборот, всегда найдется оценка, обладающая минимальным уклонением в сравнении с несмещенной оценкой. Этот факт свидетельствует в пользу смещения как первичного фактора при построении критерия эффективности оценок.

1.5. Понятие центрируемой оценки и ее определение

На практике, как уже отмечалось выше, в качестве оценок выбирается результат, полученный методами: моментов, максимального правдоподобия и минимума расстояний [1, 2]. В соответствии с определением оценка – это статистика, используемая для оценивания параметра совокупности. Статистика – это функция от выборочных значений [24]. Поэтому когда говорят об оценке показателя надежности, то понимают, что будет предложена некоторая функция от выборочных значений, с помощью которой имеется возможность получать значения, по которым оценивается истин-