Math-Net.Ru

Общероссийский математический портал

А. А. Боровков, Асимптотически оптимальные решения в задаче о разладке, *Теория вероятн. и ее примен.*, 1998, том 43, выпуск 4, 625–654

DOI: <https://doi.org/10.4213/tvp2013>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.162.27.143

26 апреля 2023 г., 16:47:32

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
И ЕЕ ПРИМЕНЕНИЯ

Основан в январе 1956 г.

Выходит 4 раза в год

Москва • Наука • ТВП

Том 43

Выпуск 4, 1998 год

октябрь, ноябрь, декабрь

© 1998 г.

БОРОВКОВ А. А.*

АСИМПТОТИЧЕСКИ ОПТИМАЛЬНЫЕ РЕШЕНИЯ
В ЗАДАЧЕ О РАЗЛАДКЕ¹⁾

В работе найдены асимптотически оптимальные процедуры в классической задаче о разладке в предположении, что момент разладки, являющийся неизвестным параметром, неограниченно возрастает.

Ключевые слова и фразы: задача о разладке, асимптотически оптимальные правила, асимптотически однородные правила.

Напомним постановку задачи о разладке.

Пусть дана последовательность X независимых наблюдений: $X = (x_1, x_2, \dots, x_{\theta-1}, x_{\theta}, \dots)$, первые $\theta - 1$ из которых имеют распределение G_1 , остальные — распределение $G_2 \neq G_1$. Не ограничивая общности, можно считать, что G_1 и G_2 имеют плотности (абсолютно непрерывны) относительно некоторой меры μ . В качестве μ можно взять, например, $\mu = G_1 + G_2$. Параметр $\theta < \infty$ неизвестен, и задача, которую мы будем рассматривать, состоит в том, чтобы по возможности точно «угадать» или оценить этот параметр по наблюдению X . Двойственный характер этой задачи состоит в том, что, с одной стороны, ее можно рассматривать как задачу о проверке конечного или счетного числа гипотез $H_j = \{\theta = j\}$, $j = 1, 2, \dots, n$; $n \leq \infty$, и тогда естественно рассматривать задачу об отыскании оптимального критерия $\delta(X)$, который минимизировал бы вероятность ошибки, или, что то же, максимизировал бы вероятности $P_t(\delta(X) = H_t)$ «угадывания» неизвестного параметра. Через P_t мы здесь обозначаем распределение в пространстве последовательностей X в случае $\theta = t$.

*Институт математики им. С. Л. Соболева СО РАН, Университетский пр., 4, 630090 Новосибирск, Россия.

¹⁾Работа выполнена при поддержке грантов РФФИ № 96-01-01939, 96-15-96295.

С другой стороны, параметр θ носит количественный характер и, стало быть, естественно рассматривать задачу об *оценке параметра* θ и об отыскании такой оценки $\theta^* = \theta^*(X)$, которая минимизировала бы среднеквадратические отклонения $E_{\theta}(\theta^* - \theta)^2$. Ниже мы рассмотрим оба этих подхода в их асимптотической форме, предполагая параметр θ большим. Различие между этими подходами обусловлено по сути тем, что в них используются разные функции потерь. В первом случае функция потерь является «вырожденной» (принимаяющей лишь значения 1 и 0); во втором — квадратичной.

Задаче о разладке посвящена обширная литература (см., например, [1]–[9], [17], [18]; там же см. более подробную библиографию); некоторые из уже известных результатов мы будем использовать.

Предлагаемая работа состоит из 4 разделов. В первом разделе показано, что в задаче проверки гипотез $H_j = \{\theta = j\}$ критерий (оценка) максимального правдоподобия является асимптотически оптимальным в классе так называемых асимптотически однородных при $\theta \rightarrow \infty$ тестов. Найдено предельное распределение оценки максимального правдоподобия и доказана ее асимптотическая байесовость (и минимаксность) для широкого класса априорных распределений.

Во втором разделе рассматривается задача оценки неизвестного параметра θ . Здесь асимптотически оптимальной (в среднеквадратичном) оказывается так называемая оценка среднего правдоподобия. Найдено ее предельное распределение и доказана асимптотическая байесовость в широком классе априорных распределений.

Третий и четвертый разделы посвящены последовательным процедурам, т.е. отысканию таких моментов остановки τ для последовательности $x_1, x_2, \dots$, которые «как можно скорее» после момента разладки θ «объявляли бы тревогу». Задача обычно состоит в отыскании такого τ , для которого $E_{\theta}(\tau - \theta \mid \tau \geq \theta)$ было бы в том или ином смысле наименьшим в некотором естественном классе марковских моментов τ . В качестве таких классов часто рассматривают совокупности τ , для которых либо вероятность ложной тревоги не превосходит при всех θ некоторого заданного $\varepsilon > 0$:

$$P_{\theta}(\tau < \theta) \leq \varepsilon,$$

либо время до объявления ложной тревоги достаточно велико,

$$E_{\infty} \tau \geq B.$$

В работе выбрано ограничение первого типа

$$\lim_{\theta \rightarrow \infty} P_{\theta}(\tau < \theta) \leq \varepsilon$$

в классе так называемых *асимптотически однородных моментов остановки*. В этих условиях доказана асимптотическая оптимальность (первого порядка) момента остановки, полученного методом, близким к ме-

тоду кумулятивных сумм (CUSUM), предложенному в [1], [2], и момента остановки, близкого к моменту, предложенному в [9]. Для них найдено асимптотическое представление для $E_t(\tau - t \mid \tau \geq t)$ при $t \rightarrow \infty$ с точностью до величины $o(1)$. Оптимальность названных и несколько иных моментов остановки с точки зрения других подходов рассматривалась в [1]–[9], [17], [18]. Например, в [4] в байесовской постановке задачи был найден явный вид момента остановки, минимизирующий усредненное значение $E_\theta(\tau - \theta)^+$ при фиксированном усредненном значении $P_\theta(\tau < \theta)$. В [8], [9] рассматривались асимптотически оптимальные τ в предположении

$$E_\infty \tau \geq B \rightarrow \infty.$$

1. Задача о разладке как задача проверки гипотез. Мы уже отмечали, что θ является количественным параметром. Поэтому в этом разделе наряду с термином «критерий $\delta(X)$ », где $\delta(X)$ принимает значения в множестве $\{H_1, H_2, \dots\}$, нам будет удобно использовать также термин «оценка $\theta^* = \theta^*(X)$ », где θ^* принимает значения $1, 2, \dots$. Функцию $\theta^*(X)$, как и $\delta(X)$, в рамках этого раздела можно также называть критерием.

Обозначим

$$r_i = \frac{g_1(x_i)}{g_2(x_i)}, \quad R(k) = R(k, X) = \prod_{i=1}^{k-1} r_i.$$

Критерием (оценкой) максимального правдоподобия (о.м.п.) мы будем называть такое наименьшее число $\hat{\theta}^* = \theta^*(X)$, что

$$R(\hat{\theta}^*, X) \geq R(t, X)$$

при всех $t \leq n + 1$, где $n \leq \infty$ — число наблюдений.

Если $n < \infty$, то $\theta = \hat{\theta}^*$ максимизирует функцию правдоподобия

$$f_\theta(X) = \prod_{i=1}^{\theta-1} g_1(x_i) \prod_{i=\theta}^n g_2(x_i). \quad (1)$$

О п р е д е л е н и е 1. Критерий $\theta^* = \theta^*(X)$ называется *асимптотически однородным*, если

$$P_t(\theta^* = t) \longrightarrow p \geq 0 \quad (2)$$

при $t \rightarrow \infty$.

Если мы имеем дело с ограниченной последовательностью $n < \infty$ наблюдений, то в (2) предполагается также, что $n \rightarrow \infty$, $n - t \rightarrow \infty$.

Покажем, что о.м.п. $\hat{\theta}^*$ является асимптотически однородной, и найдем соответствующее ей значение $p = \hat{p}$. Нам понадобятся следующие

обозначения. Рассмотрим две независимые друг от друга последовательности $\{\xi_i\}$ и $\{\zeta_i\}$ независимых случайных величин, определенных равенствами

$$\xi_i = -\ln \frac{g_1(x_i)}{g_2(x_i)}, \quad x_i \in G_1;$$

$$\zeta_i = \ln \frac{g_1(x_i)}{g_2(x_i)}, \quad x_i \in G_2.$$

Очевидно, что

$$E\xi_i = -\int g_1(x) \ln \frac{g_1(x)}{g_2(x)} \mu(dx) = -\rho(G_1, G_2) < 0,$$

$$E\zeta_i = \int g_2(x) \ln \frac{g_1(x)}{g_2(x)} \mu(dx) = -\rho(G_2, G_1) < 0,$$

где ρ — расстояние Кульбака–Лейблера между распределениями. Оба расстояния $\rho(G_1, G_2)$ и $\rho(G_2, G_1)$ мы будем предполагать конечными.

Обозначим

$$\varphi(\lambda) = E e^{\lambda \xi_1}, \quad \psi(\lambda) = E e^{\lambda \zeta_1}$$

и заметим, что $\psi(\lambda) = \varphi(1 - \lambda)$ и $\varphi(\lambda) \leq 1$ при $\lambda \in [0, 1]$, так как

$$\begin{aligned} \varphi(\lambda) &= \int g_1(x) \left(\frac{g_1(x)}{g_2(x)} \right)^{-\lambda} \mu(dx) = \int g_1^{1-\lambda}(x) g_2^\lambda(x) \mu(dx) \\ &\leq \left(\int g_1 d\mu \right)^{1/(1-\lambda)} \left(\int g_2 d\mu \right)^{1/\lambda} = 1. \end{aligned}$$

Такое же неравенство справедливо для $\psi(\lambda)$. Кроме того, производные $\varphi'(\lambda) = -\psi'(1 - \lambda)$ в силу сделанных предположений конечны на всем отрезке $[0, 1]$, $\varphi'(0) < 0$, $\varphi'(1) > 0$. Обозначим, далее,

$$S_n = \sum_{i=1}^n \xi_i, \quad Z_n = \sum_{i=1}^n \zeta_i,$$

$$S = \sup_{k \geq 1} S_k, \quad Z = \sup_{k \geq 1} Z_k.$$

Так как $E\xi_i$ и $E\zeta_i$ отрицательны, то в силу усиленного закона больших чисел S и Z — собственные случайные величины,

$$P(S < 0) > 0, \quad P(Z < 0) > 0$$

(см., например, [10], [11]).

Обозначим $n(S)$ момент *последнего* достижения неотрицательного максимума S , так что $S_{n(S)} \equiv S$, если $S \geq 0$, и $n(Z)$ — момент *первого* достижения $Z \geq 0$, $Z_{n(Z)} \equiv Z$. Если $S < 0$, $Z < 0$, то полагаем $n(S) = 0$, $n(Z) = 0$ соответственно. Определим на одном вероятностном

пространстве две независимые пары $(S, n(S))$ и $(Z, n(Z))$ и случайную величину

$$D = \begin{cases} -n(S) < 0 & \text{на множестве } Z \leq S, S \geq 0; \\ 0 & \text{на множестве } S < 0, Z \leq 0; \\ n(Z) > 0 & \text{на множестве } Z > S, Z > 0. \end{cases}$$

Если $P(S_k = 0) = 0$ при всех k , то моменты первого и последнего достижения максимума S совпадают с вероятностью 1 (см. [10], [11]). То же справедливо для Z , если $P(Z_k = 0) = 0$ при всех k . В этом случае распределение D совпадает с распределением

$$D_0 = \begin{cases} -n(S) < 0 & \text{на множестве } Z^+ \leq S^+, S^+ > 0; \\ 0 & \text{на множестве } S^+ = Z^+ = 0; \\ n(Z) > 0 & \text{на множестве } Z^+ > S^+, Z^+ > 0, \end{cases}$$

где $x^+ = \max(0, x)$. Совместные распределения пар $(S^+, n(S))$ и $(Z^+, n(Z))$ (точнее, двойные преобразования $E z^{n(S)} e^{\lambda S^+}$ и $E z^{n(Z)} e^{\lambda Z^+}$) могут быть найдены в явном виде в терминах компонент факторизации функций $1 - z\varphi(\lambda)$ и $1 - z\psi(\lambda)$ соответственно. Если $w_{z+}(v)$ — положительная компонента канонической факторизации функции $1 - z\varphi(iv)$, т.е. функция, аналитическая в верхней полуплоскости, непрерывная включая границу $\text{Im } v = 0$ и обладающая свойствами

$$\inf_{\text{Im } v > 0} |w_{z+}(v)| > 0, \quad w_{z+}(i\infty) = 1,$$

(подробнее см. [11, § 14, 15]; [10, гл. 11]), то

$$E z^{n(S)} e^{ivS^+} = \frac{w_{1+}(0)}{w_{z+}(v)}.$$

В широких предположениях справедливо асимптотическое представление

$$P(n(S) = k) \sim \frac{c\varphi^k}{k^{3/2}}, \quad c = \text{const} > 0, \quad \varphi = \min_{\lambda} E e^{\lambda \xi} < 1 \quad (3)$$

(см. [11, § 21]). Аналогичные соотношения верны для $E z^{n(Z)} e^{ivZ^+}$ и $P(n(Z) = k)$.

Теорема 1. P_t -распределение $\hat{\theta}^* - t$ сходится к распределению D :

$$P_t(\hat{\theta}^* - t = k) \rightarrow P(D = k) \quad (4)$$

при $t \rightarrow \infty$. Это означает, что оценка максимального правдоподобия $\hat{\theta}^*$ является асимптотически однородной,

$$P_t(\hat{\theta}^* = t) \downarrow \hat{p} = P(S < 0) P(Z \leq 0) \quad (5)$$

при $t \uparrow \infty$ (и $n - t \uparrow \infty$, если $n < \infty$),

$$\begin{aligned} P_t(\hat{\theta}^* > t) &\longrightarrow P(Z > S; Z > 0), \\ P_t(\hat{\theta}^* < t) &\longrightarrow P(Z \leq S; S \geq 0). \end{aligned} \quad (6)$$

Кроме того,

$$P_t(\hat{\theta}^* > t + k) \leq 2\varphi^k, \quad P_t(\hat{\theta}^* < t - k) \leq 2\varphi^k. \quad (7)$$

В силу оценок (7) сходимость (4) будет иметь место вместе со сходимостью моментов любого порядка. Соотношения (3) показывают, что оценки (7) являются правильными с точностью до степенного множителя.

Доказательство. Чтобы не выписывать громоздких соотношений, мы докажем сначала более простые утверждения о сходимости (5), (6). Имеем

$$\{\hat{\theta}^* = t\} = \left\{ \begin{aligned} \prod_{i=1}^{t-1} r_i > \prod_{i=1}^{k-1} r_i & \text{ при всех } k, 1 \leq k < t; \\ \prod_{i=1}^{t-1} r_i \geq \prod_{i=1}^{k-1} r_i & \text{ при всех } k, n+1 \geq k > t \end{aligned} \right\}. \quad (8)$$

Поэтому в силу независимости x_i

$$\begin{aligned} P_t(\hat{\theta}^* = t) &= P(-\xi_k - \dots - \xi_{t-1} > 0 \text{ при всех } k, 1 \leq k < t) \\ &\quad \times P(\zeta_t + \dots + \zeta_{k-1} \leq 0 \text{ при всех } k, t < k \leq n+1) \\ &\quad \uparrow P(S < 0) P(Z \leq 0) \end{aligned}$$

при $t \uparrow \infty$, $n - t \uparrow \infty$. Аналогично находим

$$\{\hat{\theta}^* > t\} = \left\{ \max_{k > t} \prod_{i=1}^{k-1} r_i > \max_{s \leq t} \prod_{i=1}^{s-1} r_i \right\};$$

$$\begin{aligned} P_t(\hat{\theta}^* > t) &= P\left(\max_{t < k \leq n+1} (\zeta_t + \dots + \zeta_{k-1}) > \max_{s \leq t} (\xi_s + \dots + \xi_{t-1}) \right) \\ &= P\left(\max_{1 \leq l \leq n-t} Z_l > \max_{0 \leq l \leq t-1} S_l \right) \longrightarrow P(Z > S; Z > 0). \end{aligned}$$

Доказательство сходимости (4) происходит столь же просто, основано на представлении события $\{\hat{\theta}^* = t + k\}$ в виде, аналогичном (8), и отличается от предыдущего лишь несколько большей громоздкостью. Далее,

$$\{\hat{\theta}^* > t + k\} = \left\{ \max_{j > t+k} \prod_{i=1}^{j-1} r_i > \max_{j \leq t+k} \prod_{i=1}^{j-1} r_i \right\}, \quad (9)$$

$$P_t(\hat{\theta}^* > t + k) = P\left(S_{t-1} + Z_k + \max_{1 \leq l \leq n-t-k} Z_l^* > \max(S_1, \dots, S_{t-1}, S_{t-1} + \zeta_1, \dots, S_{t-1} + Z_k)\right), \quad (10)$$

где последовательность $\{Z_l^*\}$ имеет то же распределение, что и последовательность $\{Z_l\}$ и не зависит от нее. Так как последний $\max$ под знаком вероятности в (10) больше или равен S_{t-1} , то вероятность (10) не превосходит

$$P(Z_k + Z^* > 0) \leq P(Z_k > 0) + \int_{-\infty}^0 P(Z_k \in dv) P(Z^* > -v).$$

Так как $\psi(\lambda) = E e^{\lambda \zeta_1} = 1$ при $\lambda = 1$, то

$$P(Z^* > u) \leq e^{-u}, \quad P(Z_k > 0) \leq \psi^k(\lambda)$$

при любом λ , $1 \geq \lambda \geq 0$ (см., например, [10], [11]). Поэтому

$$P_t(\hat{\theta}^* > t + k) \leq \psi^k(\lambda) + \int_{-\infty}^{\infty} e^{v\lambda} P(Z_k \in dv) = \psi^k(\lambda) + \psi^k(\lambda) = 2\varphi^k$$

при $\lambda = \lambda_0$, $\varphi = \psi(\lambda_0) = \min_{(\lambda)} \psi(\lambda) < 1$. Теорема доказана.

Класс $K_{(0)}$ асимптотически однородных критериев, введенный в определении 1, следует считать весьма естественным, поскольку он состоит из критериев (оценок), ведущих себя «симметричным образом» по отношению к разным большим значениям параметра θ . Аналогичное свойство симметричности в теории оценивания и в теории проверки гипотез вводится, например, понятиями несмещенности или инвариантности. Если мы не имеем дело с байесовской постановкой задачи, то у нас нет оснований выделять те или иные значения θ , так что для разумных оценок θ^* значения $P_t(\theta^* = t)$ должны быть примерно одинаковыми при всех больших значениях t (и $n - t$, если $n < \infty$). Теорема 1 показывает, что класс $K_{(0)}$ не пуст.

Покажем теперь, что о.м.п. $\hat{\theta}^*$ обладает свойством оптимальности в $K_{(0)}$.

Теорема 2. Пусть θ^* — любая оценка из $K_{(0)}$:

$$P_t(\theta^* = t) \rightarrow p \geq 0$$

при $t \rightarrow \infty$ (и $n - t \rightarrow \infty$, если $n < \infty$). Тогда $p \leq \hat{p}$.

Доказательство. Рассмотрим байесовскую постановку задачи, когда на множестве $\{1, \dots, n+1\}$ задано априорное распределение Q с вероятностями $q(k) = 1/(n+1)$. (Точнее, мы будем рассматривать последовательность $Q = Q^{(n)}$ распределений, зависящих от $n \rightarrow \infty$.)

Байесовский критерий θ_Q^* выбирает в качестве θ значение k , при котором максимизируется апостериорная вероятность

$$q(k | X) = \frac{q(k)f_k(X)}{\sum_{i=1}^{n+1} q(i)f_i(X)} = \frac{f_k(X)}{\sum_{i=1}^{n+1} f_i(X)}.$$

Это означает, что $\theta_Q^* = \hat{\theta}^*$. Далее, для вероятности ошибки $\alpha_Q(\theta^*) = \sum_{k=1}^{n+1} q(k) P_k(\theta^* \neq k)$ асимптотически однородного критерия θ^* будем иметь $\alpha_Q(\theta_Q^*) \leq \alpha_Q(\theta^*)$,

$$\begin{aligned} p &= \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=1}^{n+1} P_k(\theta^* = k) = \lim_{n \rightarrow \infty} (1 - \alpha_Q(\theta^*)) \\ &\leq \lim_{n \rightarrow \infty} (1 - \alpha_Q(\hat{\theta}^*)) = \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=1}^{n+1} P_k(\hat{\theta}^* = k) = \hat{p}. \end{aligned}$$

Теорема доказана.

Можно выделить еще один класс критериев (оценок), в котором о.м.п. будет асимптотически оптимальной. Это класс байесовских критериев с «плавно меняющимися» априорными вероятностями $q(k)$.

Чтобы ввести этот класс, рассмотрим произвольную последовательность $\delta \rightarrow 0$. Если $n < \infty$, то мы будем считать, что $n = n(\delta) \rightarrow \infty$ при $\delta \rightarrow 0$ (можно отождествить, например, δ с $1/n$).

Обозначим через $\mathcal{Q}$ класс распределений Q на $\{1, \dots, n+1\}$, $n \leq \infty$, зависящих от параметра δ и обладающих следующим свойством: для любых $k_1 = k_1(\delta) < k_2 = k_2(\delta)$ таких, что $k_1 \rightarrow \infty$, $n - k_2 \rightarrow \infty$ при $\delta \rightarrow 0$, выполняется

$$\begin{aligned} \frac{q(k+1)}{q(k)} &< e^\varepsilon \quad \text{при } k_1 \leq k \leq n, \\ \frac{q(k-1)}{q(k)} &< e^\varepsilon \quad \text{при } 2 \leq k \leq k_2, \end{aligned} \quad (11)$$

где $\varepsilon = \varepsilon(k_1, k_2) \rightarrow 0$ при $\delta \rightarrow 0$.

Очевидно, что для любого фиксированного l

$$\frac{q(k+l)}{q(k)} \rightarrow 1$$

при $\delta \rightarrow 0$, $k \rightarrow \infty$, $n - k \rightarrow \infty$.

Равномерное распределение $q(k) = (n+1)^{-1}$ на множестве $\{1, \dots, n+1\}$ принадлежит $\mathcal{Q}$ при $\varepsilon = 0$, $\delta = n^{-1}$. Геометрическое распределение $q(k) = p(1-p)^{k-1}$, $k = 1, 2, \dots$, также принадлежит $\mathcal{Q}$ при $\delta = p \rightarrow 0$, $\varepsilon = -\ln(1-p)$.

В обоих этих примерах соотношения (11) выполняются и без условия роста k_1 и $n - k_2$.

Названные два примера можно обобщить следующим образом. Пусть, как и прежде, δ — произвольная последовательность, сходящаяся к 0, и $p(t)$ — плотность, непрерывная и положительная внутри отрезка $[0, a]$ и равная нулю вне его, $a \leq \infty$. Тогда при широких условиях на поведение $p(t)$ на концах интервала $(0, a)$ последовательность

$$q(k) = \int_{(k-1)\delta}^{k\delta} p(u) du, \quad k = 1, \dots, \frac{a}{\delta}$$

(считаем для простоты, что δ нацело делит a), будет принадлежать $\mathcal{Q}$. Можно взять, например,

$$p(u) = \begin{cases} c(\alpha, \beta) u^\alpha (1-u)^\beta, & u \in [0, 1], \\ 0, & u \notin [0, 1], \end{cases}$$

при любых $\alpha \geq 0, \beta \geq 0$, где $c(\alpha, \beta)$ — нормирующая постоянная.

Для определенности будем называть байесовским критерием $\theta_Q^* = \theta_Q^*(X)$ функцию, которая выбирает минимальное значение k , для которого $q(k | X) = \max_j q(j | X)$.

Теорема 3. Если $Q \in \mathcal{Q}$, то

$$P_t(\theta_Q^* - t > k) \leq 2(\varphi e^{\varepsilon \lambda_0})^k, \quad P_t(\theta_Q^* - t < -k) \leq 2(\varphi e^{\varepsilon \lambda_0})^k, \quad (12)$$

где, как и прежде, λ_0 — точка, в которой достигается $\min \psi(\lambda)$, $\varphi = \psi(\lambda_0)$. Если $Q \in \mathcal{Q}$, $P(S_j = 0) = P(Z_j = 0) = 0$ при всех j , то

$$P_t(\theta_Q^* \neq \hat{\theta}^*) \rightarrow 0 \quad (13)$$

при $t \rightarrow \infty, n - t \rightarrow \infty$.

Очевидно, что наряду с (13) имеет место также сходимость $P(\theta_Q^* \neq \hat{\theta}^*) \rightarrow 0$ относительно распределения P пары (Q, X) , $\theta \in Q \in \mathcal{Q}$.

Вместе с теоремой 2 утверждение (13) означает, что все байесовские критерии, соответствующие $Q \in \mathcal{Q}$, асимптотически эквивалентны о.м.п. $\hat{\theta}^*$ и имеют в пределе одну и ту же вероятность ошибки, равную $1 - \hat{p}$.

Так как о.м.п. $\hat{\theta}^*$ является байесовским критерием (для равномерного априорного распределения), а $P_t(\hat{\theta}^* = t)$ асимптотически не зависит от t , то нетрудно установить также, что $\hat{\theta}^*$ является асимптотически минимаксным критерием при соответствующем естественном определении этого понятия (ср. с [16]).

Доказательство теоремы 3. Доказательство неравенств (12) вполне аналогично доказательству неравенств (7) в теореме 1. Имеем

$$\{\theta_Q^* > t + k\} = \left\{ \max_{j > t+k} q(j) \prod_{i=1}^{j-1} r_i > \max_{j \leq t+k} q(j) \prod_{i=1}^{j-1} r_i \right\}.$$

Совершенно аналогично (10) имеем

$$\begin{aligned}
 P_t(\theta_Q^* > t+k) &= P\left(S_{t-1} + Z_k + \max_{l \geq 1} (Z_l^* + \ln q(t+k+l)) \right. \\
 &> \max(S_1 + \ln q(2), \dots, S_{t-1} + \ln q(t); \\
 &\quad \left. S_{t-1} + Z_1 + \ln q(t+1), \dots, S_{t-1} + Z_k + \ln q(t+k))\right) \\
 &\leq P(S_{t-1} + Z_k + \varepsilon k + \max_{l \geq 1} (Z_l^* + \varepsilon l) > S_{t-1}) \\
 &\leq P(Z_k + \varepsilon k > 0) + \int_{-\infty}^0 P(Z_k + \varepsilon k \in dv) P\left(\max_{l \geq 1} (Z_l^* + \varepsilon l) > -v\right).
 \end{aligned}$$

Пусть $s > 0$ — решение уравнения $E e^{\lambda \zeta_1 + \varepsilon \lambda} = \psi(\lambda) e^{\varepsilon \lambda} = 1$. Очевидно, что $s < 1$, $s \rightarrow 1$ при $\varepsilon \rightarrow 0$. Поэтому при любом фиксированном $\lambda \in (0, 1)$

$$P_t(\theta_Q^* > t+k) < (\psi(\lambda) e^{\varepsilon \lambda})^k + \int_{-\infty}^0 e^{s v} P(Z_k + \varepsilon k \in dv) < 2(\psi(\lambda) e^{\varepsilon \lambda})^k.$$

Отсюда следует первое утверждение теоремы. Второе неравенство в (12) получается аналогично.

Докажем (13). Событие $\theta_Q^* \neq \hat{\theta}^*$ наступает в том случае, когда максимум последовательности $q(k) \prod_{i=1}^{k-1} r_i$ достигается в точке k , отличной от точки максимума $\prod_{i=1}^{k-1} r_i$. В силу неравенств (7), (12) мы можем ограничиться рассмотрением значений k , $|t-k| \leq N$, при достаточно большом, но фиксированном N . Но в этих пределах

$$m_1 = \min_{|k-t| \leq N} \frac{q(k)}{q(t)} \rightarrow 1, \quad m_2 = \max_{|k-t| \leq N} \frac{q(k)}{q(t)} \rightarrow 1$$

при $t \rightarrow \infty$, $n-t \rightarrow \infty$. Стало быть, $P_t(\theta_Q^* \neq \hat{\theta}^*)$ оценивается P_t -вероятностью того, что S_r и Z_l при некоторых $r \leq N$, $l \leq N$ попадут в интервал в окрестности нуля, ширина которого не превышает $\ln m_2 - \ln m_1 \rightarrow 0$. Так как распределения S_r и Z_l непрерывны в точке 0, то P_t -вероятности этих событий сходятся к нулю. Теорема доказана.

2. Задача о разладке как задача оценки параметра. Если число $\hat{p}$ в теореме 1 невелико, то шансы точно угадать параметр θ тоже невелики. Поэтому подходы предыдущего раздела, максимизирующие вероятность точного угадывания, не всегда являются наиболее естественными. Как правило, целью является не точное угадывание θ , а определение значения θ^* , которое было бы по возможности близко к θ , например, в среднеквадратичном смысле. Отметим при этом, что мы оцениваем *целочисленный* параметр *случайного процесса* $x_1, x_2, \dots$ по *одному наблюдению* X . Поэтому невозможно рассчитывать получить оценки, которые обладали бы свойством состоятельности (это видно и

из содержания предыдущего раздела). Мы опять будем рассматривать асимптотическую постановку задачи.

О п р е д е л е н и е 2. Оценка θ^* называется *асимптотически однородной* (в среднеквадратичном), если

$$E_t(\theta^* - t)^2 \longrightarrow \sigma^2 \leq \infty \quad (14)$$

при $t \rightarrow \infty$ (и $n - t \rightarrow \infty$, если $n < \infty$). Класс таких оценок обозначим $K_{(2)}$. Из теоремы 1 следует, что о.м.п. $\hat{\theta}^* \in K_{(2)}$, при этом $\sigma^2 = \hat{\sigma}^2 = ED^2$.

О п р е д е л е н и е 3. Оценка $\tilde{\theta}^*$, определенная равенством

$$\tilde{\theta}^* = \sum_{k \geq 1} k R(k) \left(\sum_{k \geq 1} R(k) \right)^{-1}, \quad (15)$$

называется *оценкой среднего правдоподобия*.

Ясно, что эта оценка не обязательно целочисленна. Наряду с $\tilde{\theta}^*$ можно рассматривать эквивалентную ей в известном смысле рандомизированную целочисленную оценку $\tilde{\theta}^{**}$, принимающую значение $[\tilde{\theta}^*]$ с вероятностью $1 - \{\tilde{\theta}^*\}$ и значение $[\tilde{\theta}^*] + 1$ с вероятностью $\{\tilde{\theta}^*\}$, где $[x]$ и $\{x\}$ — целая и дробная части x соответственно.

Введем в рассмотрение случайные величины

$$\begin{aligned} E_1 &= \sum_{k=1}^{\infty} e^{S_k}, & E_2 &= \sum_{k=1}^{\infty} e^{Z_k}, \\ E'_1 &= \sum_{k=1}^{\infty} k e^{S_k}, & E'_2 &= \sum_{k=1}^{\infty} k e^{Z_k}. \end{aligned} \quad (16)$$

Так как по усиленному закону больших чисел $S_k < -ck < 0$, $Z_k < -ck < 0$ с вероятностью 1 при некотором $c > 0$ ($c < -E\xi_1$, $c < -E\xi_1$) и при всех k , начиная с некоторого, то рассматриваемые ряды сходятся, а введенные случайные величины являются собственными (заметим, что при этом $E e^{S_k} = E e^{Z_k} = 1$, $E E_i = \infty$).

Аналогичным образом определяются величины $E_{i,s}$, $E'_{i,s}$, $i = 1, 2$, с заменой суммирования в (16) по всем k на суммирование по $k \leq s$, так что $E_{i,s} \rightarrow E_{i,\infty} = E_i$, $E'_{i,s} \rightarrow E'_{i,\infty} = E'_i$ при $s \rightarrow \infty$. В дальнейшем важную роль будет играть случайная величина

$$H = \frac{E'_2 - E'_1}{1 + E_1 + E_2}.$$

Для изучения ее свойств нам понадобится следующее вспомогательное утверждение. Обозначим

$$b^{-1} = 3 + \sum_{k=1}^{\infty} k^{-1} \ln^{-2} k.$$

Лемма 1. При всех достаточно больших x справедливо неравенство

$$P\left(\frac{E'_1}{E_1} > x\right) \leq \frac{2(bx)^\lambda}{1-\varphi} \exp\left(\frac{\sqrt{bx}}{\ln \sqrt{bx}} \ln \varphi\right),$$

где $\lambda > 0$ есть значение аргумента, при котором достигается $\min \varphi(\lambda) = \varphi < 1$.

Такое же неравенство справедливо и для $P(E'_{i,s}/E_{i,s} > x)$ при всех $s > 3; i = 1, 2$.

Таким образом, случайные величины E'_i/E_i имеют моменты любого порядка. Так как $|H| \leq E'_1/E_1 + E'_2/E_2$, то такое же утверждение справедливо и для H . Из доказательства леммы будет видно, что полученные оценки могут быть улучшены. Однако истинная асимптотика $P(E'_i/E_i > x)$ при $x \rightarrow \infty$ остается неясной. Возможно, она имеет вид (3).

Доказательство. Обозначим через η значение k , при котором достигается максимум последовательности S_k^* , $S_k^* = S_k + 2 \ln k + 2 \ln \ln k$, при $k \geq 3$; $S_0^* = 0, S_1^* = S_1, S_2^* = S_2$. Обозначим $A_{n,dv} = \{\eta = n, S_\eta^* \in dv\}$, $B = \{E'_1/E_1 > x\}$. Тогда

$$P(B) = \sum_{n=0}^{\infty} \int_0^{\infty} P(BA_{n,dv}). \quad (17)$$

На событии $A_{n,dv}$ выполняется

$$\begin{aligned} E'_1 &= \sum_{k=1}^{\infty} k e^{S_k} \leq 3e^v + \sum_{k=3}^{\infty} k e^{v-2 \ln k - 2 \ln \ln k} \\ &\leq e^v \left(3 + \sum_{k=3}^{\infty} k^{-1} \ln^{-2} k \right) = b^{-1} e^v; \\ E_1 &\geq e^{v-2 \ln n - 2 \ln \ln n} = e^v (n \ln n)^{-2}; \\ \frac{E'_1}{E_1} &\leq b^{-1} (n \ln n)^2. \end{aligned}$$

Поэтому произведение $BA_{n,dv}$ не пусто лишь при n таких, что $n \ln n \geq \sqrt{bx}$,

$$BA_{n,dv} \subset A_{n,dv} I \left\{ n \geq \frac{\sqrt{bx}}{\ln \sqrt{bx}} \right\}.$$

Следовательно, в силу (17)

$$P(B) \leq P \left(\eta \geq \frac{\sqrt{bx}}{\ln \sqrt{bx}} \right). \quad (18)$$

Но $\{\eta = k\} \subset \{S_k + 2 \ln k + 2 \ln \ln k \geq 0\}$,

$$\begin{aligned} P(\eta = k) &\leq P(S_k \geq -2 \ln k - 2 \ln \ln k) \\ &\leq \varphi^k(\lambda) e^{\lambda(2 \ln k + 2 \ln \ln k)} = \varphi^k(k \ln k)^{2\lambda}, \end{aligned}$$

где под $\lambda > 0$ мы понимаем значение, при котором достигается $\min \varphi(\lambda) = \varphi < 1$. Следовательно, при достаточно больших k

$$P(\eta \geq k) \leq \frac{2\varphi^k}{1-\varphi} (k \ln k)^{2\lambda},$$

и, в силу (18),

$$P(B) \leq \frac{2}{1-\varphi} \exp \left(\frac{\sqrt{bx}}{\ln \sqrt{bx}} \ln \varphi \right) \left(\frac{\sqrt{bx}}{\ln \sqrt{bx}} \ln \frac{\sqrt{bx}}{\ln \sqrt{bx}} \right)^{2\lambda}.$$

Очевидно, что при $bx \geq e$ последний множитель не превосходит $(\sqrt{bx})^{2\lambda}$. Это доказывает первое утверждение леммы.

Второе утверждение доказывается точно так же. Надо лишь вместо η рассмотреть значение $\eta_{(s)} = \min(s, \eta)$, при котором достигается максимум S_k^* , $k \leq s$. Лемма доказана.

Мы можем сформулировать теперь основное утверждение этого раздела.

Теорема 4. P_θ -распределение $\tilde{\theta}^* - \theta$ слабо сходится при $\theta \rightarrow \infty$ (и $n - \theta \rightarrow \infty$, если $n < \infty$ и $n \rightarrow \infty$) к распределению H . Эта сходимость имеет место вместе со сходимостью моментов любого порядка, $\tilde{\sigma}^2 = EH^2 < \infty$. Оценка среднего правдоподобия $\tilde{\theta}^*$ является асимптотически оптимальной в $K_{(2)}$: для любой оценки $\theta^* \in K_{(2)}$ с предельным «разбросом» σ^2 (см. (14)) выполняется $\sigma^2 \geq \tilde{\sigma}^2$. Кроме того, для любого $Q \in \mathcal{Q}$ байесовская оценка θ_Q^* и оценка $\tilde{\theta}^*$ асимптотически эквивалентны:

$$P_t(|\theta_Q^* - \tilde{\theta}^*| > \varepsilon) \rightarrow 0 \quad (19)$$

при $t \rightarrow \infty$, $n - t \rightarrow \infty$.

Во избежание недоразумений напомним, что байесовский критерий θ_Q^* , фигурирующий в предыдущем разделе, и байесовская оценка θ_Q^* , обозначаемая тем же символом, определяются по-разному.

Доказательство. Докажем сначала, что $\tilde{\theta}^* \in K_{(2)}$. Имеем

$$R(k) = \begin{cases} e^{-S_{k-1}} & \text{при } k \leq \theta, \\ e^{-S_{\theta-1} + Z_{k-\theta}} & \text{при } k \geq \theta. \end{cases}$$

Умножим числитель и знаменатель в (15) на $e^{S_{\theta-1}}$. Тогда

$$\begin{aligned} \tilde{\theta}^* - \theta &= \left(\sum_{k \leq \theta} (k - \theta) e^{S_{\theta-1} - S_{k-1}} + \sum_{k > \theta} (k - \theta) e^{Z_{k-\theta}} \right) \\ &\times \left(\sum_{k \leq \theta} e^{S_{\theta-1} - S_{k-1}} + \sum_{k > \theta} e^{Z_{k-\theta}} \right)^{-1} = \frac{E'_{2,n-\theta} - E'_{1,\theta}}{d \cdot 1 + E_{2,n-\theta} + E_{1,\theta}}. \end{aligned}$$

Очевидно, далее, что

$$\left| \frac{E'_{1,\theta}}{1 + E_{1,\theta} + E_{2,n-\theta}} - \frac{E'_1}{1 + E_1 + E_2} \right| \xrightarrow{p} 0$$

при $\theta \rightarrow \infty$, $n - \theta \rightarrow \infty$. Кроме того, согласно лемме 1, момент любого фиксированного порядка величин

$$\frac{E'_{1,\theta}}{E_{1,\theta}} \geq \frac{E'_{1,\theta}}{1 + E_{1,\theta} + E_{2,n-\theta}}$$

равномерно по θ ограничен. Отсюда следует сходимость распределений $E'_{1,\theta}/(1 + E_{1,\theta} + E_{2,n-\theta})$ с моментами любого порядка. Такое же утверждение справедливо для $E'_{2,n-\theta}/(1 + E_{1,\theta} + E_{2,n-\theta})$. Это доказывает сходимость P_{θ} -распределения $\theta^* - \theta$ к распределению H вместе с моментами любого порядка.

Рассмотрим теперь байесовскую постановку задачи и равномерное на $\{1, \dots, n+1\}$ априорное распределение Q . Тогда очевидно, что $\tilde{\theta}^*$ совпадает с байесовской оценкой θ_Q^* . Обозначая $E = E_Q$ математическое ожидание относительно распределения (θ, X) , $\theta \in Q$, получим, что для любой оценки $\theta^* \in K_{(2)}$ с предельным разбросом σ^2 (см. (14)) справедливо

$$\begin{aligned} \sigma^2 &= \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{t=1}^{n+1} E_t(\theta^* - t)^2 = \lim_{n \rightarrow \infty} E(\theta^* - \theta)^2 \\ &\geq \lim_{n \rightarrow \infty} E(\theta_Q^* - \theta)^2 = \lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{t=1}^{n+1} E_t(\tilde{\theta}^2 - t)^2 = \tilde{\sigma}^2. \end{aligned}$$

Нам осталось убедиться, что байесовская оценка $\theta_Q^* = E(\theta | X)$ для любого априорного распределения $Q \in \mathcal{Q}$ асимптотически эквивалентна $\tilde{\theta}^*$, т.е., что выполнено (19). Имеем

$$\theta_Q^* - \tilde{\theta}^* = \frac{\sum_k (k - \theta) R(k) q(k)}{\sum_k q(k) R(k)} - \frac{\sum_k (k - \theta) R(k)}{\sum_k R(k)}. \quad (20)$$

Из доказательства первой части теоремы мы видели, что второе слагаемое в (20) будет сколь угодно близко к $(E'_{2,N} - E'_{1,N})/(1 + E_{1,N} + E_{2,N})$

при достаточно большом N , т.е. близким к слагаемому того же вида, в котором, однако, суммирование в числителе и знаменателе ведется лишь по значениям k , $|k - \theta| \leq N$.

То же самое справедливо, очевидно, и для первого слагаемого в (20), и для их разности. Но для $Q \in \mathcal{Q}$ выполняется $q(k)/q(\theta) \rightarrow 1$ при $|k - \theta| \leq N$, $\theta \rightarrow \infty$, $n - \theta \rightarrow \infty$. Поэтому разность (20), в которой суммирование ведется по k , $|k - \theta| \leq N$, сходится к 0 при $\theta \rightarrow \infty$, $n - \theta \rightarrow \infty$. Другими словами, при $\theta = t$ $P_t(|\theta_Q^* - \tilde{\theta}^*| > \varepsilon) \rightarrow 0$ при $t \rightarrow \infty$, $n - t \rightarrow \infty$. Теорема доказана.

Утверждение, аналогичное теореме 4, можно получить и для целочисленной рандомизированной оценки $\tilde{\theta}^{**}$. При этом под классом $K_{(2)}$ надо будет понимать класс асимптотически однородных рандомизированных целочисленных оценок;

$$E_t(\tilde{\theta}^{**} - \theta)^2 = E_t(\tilde{\theta}^* - \theta)^2 + E_t(\tilde{\theta}^* - \tilde{\theta}^{**})^2$$

будет отличаться от $E_t(\tilde{\theta}^* - \theta)^2$ на величину

$$E_t(\tilde{\theta}^* - \tilde{\theta}^{**})^2 = 2E_t\{\tilde{\theta}^*\}(1 - \{\tilde{\theta}^*\}) \leq \frac{1}{2}.$$

Отметим также, что оценка среднего правдоподобия, являясь асимптотически байесовской и однородной ($E_t(\tilde{\theta}^* - t)^2$ асимптотически от t не зависит), будет также *асимптотически минимаксной оценкой* при соответствующем естественном определении этого понятия (ср. с [16]).

3. Последовательные процедуры. Последовательные процедуры представляют интерес в тех случаях, когда существует необходимость принять решение о наличии разладки и ее локализации в возможно короткие сроки (по минимальному числу наблюдений). Сокращение числа наблюдений естественно проводить с помощью следующей процедуры. Обозначим $\hat{\theta}_m^*$ наименьшее значение k , при котором $R(k) \geq R(j)$ при всех $j \leq m + 1$, так что $\hat{\theta}_n^* = \hat{\theta}^*$, если мы располагаем n наблюдениями. В дальнейшем для определенности будем считать $n = \infty$. В качестве момента ν остановки наблюдений естественно выбирать такое m , при котором разность $\ln R(\hat{\theta}_m^*) - \ln R(m)$ достигнет достаточно большого значения, которое мы обозначим $N + b(m)$ (нетрудно видеть, что если $b(m) = o(m)$, то при $m - \theta \rightarrow \infty$ это произойдет с вероятностью 1; если же $m < \theta$, то $\hat{\theta}_m^*$ будет расти вместе с m , и рассматриваемая разность будет оставаться «ограниченной»). Итак, мы положим

$$\nu = \min \{m: \ln R(\hat{\theta}_m^*) - \ln R(m) \geq N + b(m)\}. \quad (21)$$

Технически, чтобы упростить вычисления и не запоминать все наблюдения, можно воспользоваться следующей эквивалентной процедурой.

Рассмотрим последовательность $\{w_n\}$, определенную рекурсивными соотношениями

$$\begin{aligned} w_k &= (w_{k-1} + \gamma_k)^+, \quad k \geq 1, \\ x^+ &= \max(0, x), \quad \gamma_n = -\ln r_n = -\ln \frac{g_1(x_n)}{g_2(x_n)}. \end{aligned} \quad (22)$$

Основные свойства последовательности w_n мы сформулируем в виде леммы.

Лемма 2. Уравнение (22) имеет решение

$$w_n = \max \left(w_0 + \Gamma_n, \max_{k \leq n} (\Gamma_n - \Gamma_k) \right) = \Gamma_n + \max(w_0, \bar{\Gamma}_n^-), \quad (23)$$

где $\Gamma_n = \sum_{k=1}^n \gamma_k$, $\bar{\Gamma}_n^- = \max_{k \leq n} (-\Gamma_k)$.

Если γ_k — независимы и одинаково распределены с тем же распределением, что и γ , $E\gamma < 0$, то распределение цепи Маркова w_n сходится к стационарному распределению величины $w = \bar{\Gamma}_\infty^- = \sup_{k \geq 0} (-\Gamma_k)$.

Если уравнение $\varphi_\gamma(\lambda) = Ee^{\lambda\gamma} = 1$ имеет решение $\mu > 0$, то

$$P(w > x) < e^{-\mu x}.$$

Если дополнительно $\varphi'_\gamma(\mu) < \infty$, то

$$P(w > x) \sim ce^{-\mu x}, \quad 0 < c < 1.$$

Утверждение леммы вытекает, например, из результатов глав 1, 4 в [11].

Из (23) при $w_0 = 0$ получаем, что $w_m = \ln R(\hat{\theta}_m^*) - \ln R(m)$, а $\hat{\theta}_m^*$ совпадает со значением $m - s(m)$, где $s(m)$ — наименьшее значение s , при котором $w_{m-s} = 0$.

Итак, наряду с (21) мы можем положить также

$$\nu = \nu(N) = \min \{m: w_m \geq N + b(m)\}. \quad (24)$$

Такой момент остановки наблюдений называют моментом, полученным методом кумулятивных сумм (CUSUM). Этот метод в случае $b(m) = 0$ был введен в [1], [2]. Ясно, что $P(\nu < \infty) = 1$, так как при $m > \theta$ в силу усиленного закона больших чисел $w_{m+k} > ck$, $0 < c < -E\zeta_1$, при всех достаточно больших k .

Если $k < \theta$, то распределение w_k совпадает с распределением $\bar{S}_{k-1} = \max_{0 \leq j \leq k-1} S_j$, которое монотонно сходится к распределению $S^+ = \sup_{j \geq 0} S_j$.

Поэтому если $N + b(m) = \text{const}$, а θ велико, то событие $\{w_m \geq N + b(m)\}$ рано или поздно произойдет, так что вероятность $P_\theta(\nu < \theta)$

близка к 1 и момент остановки ν с высокой вероятностью будет «ложным». Стало быть, чтобы избежать этого, мы должны выбирать растущие последовательности $b(m) \uparrow \infty$ (по возможности медленно растущие, чтобы $\nu - \theta$ было не очень велико). Пусть $b(m) \uparrow \infty$ так, что $b = \sum_{k=1}^{\infty} e^{-b(k)} \leq \infty$ (можно взять, например, $b(k) = \ln k + (1 + \beta) \ln \ln k$, $\beta > 0$ при $k \geq 3$; $b(1) = b(2) = 0$; любая такая последовательность будет верхней для $\{w_k\}$ при $\theta = \infty$, а последовательность $b(k) = \ln k + (1 - \beta) \ln \ln k$, $\beta > 0$, — нижней последовательностью).

Свойства момента остановки $\nu = \nu(N) (\{\nu \leq n\} \in \sigma(x_1, \dots, x_n))$, определенного в (24), описываются следующим утверждением. Чтобы упростить рассуждения, мы будем предполагать, что распределение $-\xi_i$ непрерывно в области $(0, \infty)$. Каким образом обойти технические трудности, возникающие при нарушении этого условия, сказано в замечании 2 в разделе 3.

Обозначим $P(t) = P(t, N) = P_t(\nu(N) < t)$.

Теорема 5. Пусть $\beta > 0$,

$$b(k) = \ln k + (1 + \beta) \ln \ln(k), \quad b = 2 + \sum_3^{\infty} k^{-1} (\ln(k))^{-(1+\beta)}.$$

Тогда при всех t

$$P(t) \leq be^{-N}. \quad (25)$$

Если распределение $-\xi_i$ непрерывно в $(0, \infty)$, то $P(t)$ и $P(\infty)$ непрерывно зависят от N и для любого $\varepsilon < \varepsilon_0$ существует $N = N(\varepsilon)$, являющееся решением уравнения $P(\infty, N) = \varepsilon$, так что $P(t) \uparrow P(\infty) = \varepsilon$ при $t \rightarrow \infty$, где $\varepsilon_0 = P_{\infty}(\cup_{k=1}^{\infty} \{w_k \geq b(k)\})$. Кроме того, для нерешетчатых ζ_i

$$\begin{aligned} E_t(\nu - t \mid \nu \geq t) &= \frac{E_t(\nu - t)^+}{1 - P(t)} \\ &= \frac{1}{\rho} (b(t) + N - E \max(S, Z) + E\chi) + o(1), \end{aligned} \quad (26)$$

где $\rho = \rho(G_2, G_1)$, S и Z независимы, χ есть величина перескока через бесконечно удаленный барьер последовательности сумм $-Z_k$. Если ζ_i решетчатые, то слагаемые $E\chi/\rho + o(1)$ в правой части (26) следует заменить на функцию, ограниченную известной постоянной.

Доказательство. Оценка (25) вытекает из очевидного соотношения

$$\begin{aligned} P(t) &= P_t(\nu < t) = P_t\left(\bigcup_{k=1}^{t-1} \{w_k \geq N + b(k)\}\right) \\ &\leq \sum_{k=1}^{t-1} P(S^+ \geq N + b(k)) \leq e^{-N} \sum_{k=1}^{\infty} e^{-b(k)} = be^{-N}. \end{aligned} \quad (27)$$

Далее, если распределение $-\xi_i$ непрерывно на $(0, \infty)$, то непрерывными будут и распределения w_k на $(0, \infty)$. Стало быть, $P(t)$ и $P(\infty)$ в силу (27) будут непрерывно зависеть от N , так как вероятность скачка в точке N оценивается суммой $\sum_{k=1}^{\infty} P(w_k = N + b(k))$. Далее,

$$P(t) = P_{\infty} \left(\bigcup_{k=1}^{t-1} \{w_k \geq N + b(k)\} \right) \uparrow P(\infty).$$

Таким образом, функция $P(\infty) = P(\infty, N)$ с ростом $N > 0$ монотонно и непрерывно убывает от ε_0 до 0. Положив $N(\varepsilon)$ равным решению уравнения $P(\infty, N) = \varepsilon$, мы получим утверждение теоремы, касающееся $N(\varepsilon)$.

Докажем теперь (26). В силу (22), (24), (27) и леммы 2 мы имеем представление

$$E_t(\nu - t | > t) = E\eta,$$

где η есть марковский момент:

$$\begin{aligned} \eta &= \min \{k: -Z_k \geq B(t, k)\}, \\ B(t, k) &= b(t+k) + N - \max(w_t, \bar{Z}_k), \end{aligned} \quad (28)$$

$\bar{Z}_k = \max_{0 < j < k} Z_j$; $w_t = {}_d S_{t-1}$ и $\{Z_k\}$ независимы.

Пусть для простоты $\beta \leq 1$. Тогда при $t > \max(3, 6/\rho)$ выполняется $b(t+k) \leq 3 \ln(t+k) \leq 3 \ln t + 3k/t \leq 3 \ln t + k\rho/2$. Поэтому

$$\eta \leq \min \{k: -Z_k \geq b(t+k) + N\} \leq \eta_0 = \min \left\{ k: -Z_k \geq 3 \ln t + N + \frac{k\rho}{2} \right\},$$

$$E\eta \leq E\eta_0 \leq \frac{6 \ln t}{\rho} + \text{const.} \quad (29)$$

Кроме того, $E|\zeta_k| < \infty$. Поэтому к мартингалу $-Z_k^0 = -Z_k - \rho k$ применима теорема Вальда-Дуба, в силу которой

$$\begin{aligned} E(-Z_{\eta} - \rho\eta) &= 0, \\ E\eta &= \frac{1}{\rho} E(b(t+\eta) + N - \max(w_t, \bar{Z}_{\eta}) + \chi_t), \end{aligned} \quad (30)$$

где $\chi_t = -Z_{\eta} - B(t, \eta)$. Мы опускаем здесь доказательство того, что $\sup_t E\chi_t < \infty$ (для ограниченных ζ_i это очевидно), и того, что для нерешетчатых ζ_i выполняется $E\chi_t \rightarrow E\chi$. Доказательство этих утверждений можно найти в [14]. Далее, распределения w_t и $\bar{Z}_{\eta}$ сходятся вместе с экспоненциальными моментами к распределениям S^+ и Z^+ соответственно. Кроме того, в силу (29)

$$Eb(t+\eta) = \ln t + O\left(\frac{\ln t}{t}\right) = \ln t + O(1).$$

Сказанное вместе с (30) доказывает (26). Теорема доказана.

Если вернуться к исходной задаче оценивания θ с помощью последовательной процедуры, то из сказанного выше следует, что в качестве приближений для $\hat{\theta}^*$ естественно взять значение

$$\hat{\theta}^{**} = \nu - s(\nu),$$

а в качестве приближения для $\tilde{\theta}^*$ — значение

$$\tilde{\theta}^{**} = \sum_{k=\nu-2s(\nu)}^{\nu} kR(k) \left(\sum_{k=\nu-2s(\nu)}^{\nu} R(k) \right)^{-1}.$$

Теорема 6.

$$\begin{aligned} \hat{\theta}^{**} &\leq \hat{\theta}^*, \\ P_t(\hat{\theta}^* > \hat{\theta}^{**}) &\leq (b + e^{-b(t)})e^{-N}. \end{aligned}$$

Кроме того, для любого $\varepsilon > 0$

$$P_t(|\tilde{\theta}^* - \tilde{\theta}^{**}| > \varepsilon) \rightarrow 0 \quad (31)$$

при $N \rightarrow \infty, t \rightarrow \infty$.

Доказательство. Первое утверждение теоремы очевидно. Далее,

$$P_t(\hat{\theta}^{**} < \hat{\theta}^*) = P_t(\hat{\theta}^{**} < \hat{\theta}^*; \nu \geq t) + P_t(\hat{\theta}^{**} < \hat{\theta}^*; \nu < t). \quad (32)$$

Событие $\{\hat{\theta}^{**} < \hat{\theta}^*; \nu \geq t\}$ означает, что в области $k \geq t$ траектория $\ln R(k)$ будет иметь по крайней мере еще один максимум, превосходящий $\hat{\theta}^{**}$. Стало быть, P_t -вероятность этого события не превосходит $P(Z > N + b(t)) \leq e^{-N-b(t)}$.

Вторая вероятность в (32) не превосходит (см. (27))

$$P_t(\nu < t) \leq be^{-N}.$$

Доказательство последнего утверждения теоремы (31) мы предоставляем читателю. Все необходимое для этого уже подготовлено.

Вернемся к теореме 5. Из нее следует, что при фиксированной вероятности ложной тревоги $P_t(\nu < t) \approx \varepsilon$ среднее значение «задержки» с объявлением тревоги $E_t(\nu - \theta \mid \nu \geq \theta)$ имеет при больших θ порядок $(\ln t)/\rho$. Оказывается, что в широких предположениях улучшить (уменьшить) величину этой задержки в классе произвольных моментов остановки $\tau(\{\tau \leq n\} \in \sigma(x, \dots, x_n))$ с той же вероятностью ложной тревоги $P_t(\tau < t)$ невозможно.

Чтобы сформулировать более точное утверждение, введем в рассмотрение классы асимптотически «однородных» (ср. с определениями 1, 2) марковских моментов.

О п р е д е л е н и е 4. Момент остановки τ называется *асимптотически однородным*, если существует

$$\lim_{t \rightarrow \infty} \frac{E_t(\tau - t \mid \tau \geq t)}{\ln t} = v, \quad 0 \leq v \leq \infty. \tag{33}$$

Класс таких моментов, для которых $P_t(\nu < t) \rightarrow \varepsilon$ при $t \rightarrow \infty$, обозначим $K(\varepsilon)$.

Теорема 7. Для любого момента остановки $\tau \in K(\varepsilon)$ выполняется $v \geq 1/\rho$.

Если не требовать существования предела (33), то из доказательства этой теоремы, приведенного ниже, будет следовать, что

$$\limsup_{t \rightarrow \infty} \frac{E_t(\tau - t \mid \tau \geq t)}{\ln t} \geq \frac{1}{\rho}.$$

Из теорем 5 и 7 вытекает, что для момента остановки ν , определенного в (24), нижняя граница $1/\rho$ возможных значений v достигается. В этом смысле момент ν является *асимптотически оптимальным*. Ниже мы построим еще один асимптотически оптимальный момент остановки, который в ряде случаев может быть более предпочтительным.

Приведем теперь утверждение об оптимальном байесовском моменте остановки в задаче о разладке, на которое мы будем опираться при доказательстве теоремы 7.

Итак, пусть задано априорное распределение $Q = \{q(n) = P(\theta = n)\}$. Обозначим $X_n = (x_1, \dots, x_n)$. Тогда определено апостериорное распределение

$$q(k \mid X_n) = \frac{q(k) f_k(X_n)}{\sum_{i=1}^\infty q(i) f_i(X_n)} = \frac{q(k) R(k)}{\sum_{i=1}^n q(i) R(i) + R(n+1) Q_n}, \tag{34}$$

где $f_k(X_n)$ и $R(k)$ определяются так же, как в (1), (2), $Q_n = \sum_{i=n+1}^\infty q(i)$. Положим

$$\begin{aligned} Q(n \mid X_n) &= P(\theta \leq n \mid X_n) = \sum_{k=1}^n q(k \mid X_n) \\ &= \left(1 + \frac{R(n+1) Q_n}{\sum_{i=1}^n q(i) R(i)} \right)^{-1} \end{aligned}$$

и определим момент остановки

$$\begin{aligned} \tau(A) &= \min \{ n \geq 1: Q(n/X_n) \geq A \} \\ &= \min \left\{ n \geq 1: \frac{\sum_{i=1}^n q(i) R(i)}{R(n+1) Q_n} \geq \frac{A}{1-A} \right\}. \end{aligned}$$

Вероятность ложной тревоги $P_Q(\tau(A) < \theta)$ равна

$$\alpha(A) = P_Q(\tau(A) < \theta) = \sum_{n=1}^\infty q(n) P_n(\tau(A) < n). \tag{35}$$

Обозначим $K_{(\varepsilon, Q)}$ класс всех моментов остановки τ , для которых $P_Q(\tau < \theta) \leq \varepsilon$.

Теорема 8. Пусть A_ε — решение уравнения $\alpha(A) = \varepsilon$. Тогда момент остановки $\tau(A_\varepsilon) \in K_{(\varepsilon, Q)}$ является оптимальным в $K_{(\varepsilon, Q)}$ в следующем смысле: для любого $\tau \in K_{(\varepsilon, Q)}$ выполняется

$$E_Q(\tau - \theta)^+ \geq E_Q(\tau(A_\varepsilon) - \theta)^+.$$

Если $P_Q(\tau < \theta) = \varepsilon$, то это эквивалентно неравенству

$$E_Q(\tau - \theta \mid \tau \geq \theta) \geq E_Q(\tau(A_\varepsilon) - \theta \mid \tau(A_\varepsilon) \geq \theta).$$

З а м е ч а н и е 1. Нетрудно видеть (ср. с теоремой 5), что

$$P_t(\tau(A) < t) = P_t\left(\bigcup_{k=1}^{t-1} \{Q(k \mid X_k) \geq A\}\right)$$

будет непрерывно зависеть от A , если распределения ξ_i (и, стало быть, $R(i)$) непрерывны. В этом случае $\alpha(A)$ в (35) также непрерывно зависит от A . Так как, кроме того, $\alpha(A)$ монотонно убывает по A ,

$$\begin{aligned} \alpha(A) &= P(\tau(A) < \theta) = 1 - P(\theta \leq \tau(A)) \\ &= 1 - EP(\theta \leq \tau(A) \mid X_{\tau(A)}) \leq 1 - A, \end{aligned} \quad (36)$$

то для любого $\varepsilon \in (0, 1)$ существует единственное решение A_ε уравнения $\alpha(A) = \varepsilon$. Отметим также, что если все $q(n)$ малы, то «перескок» через уровень A будет тоже мал и наряду с (36) будет справедливо приближенное равенство $\alpha(A) \approx 1 - A$. Например, если в (37) $p \rightarrow 0$, то $\alpha(A) \rightarrow 1 - A$, $A_\varepsilon \rightarrow 1 - \varepsilon$.

Д о к а з а т е л ь с т в о т е о р е м ы 8 для геометрического априорного распределения

$$q(n) = (1 - p)^{n-1} p, \quad n = 1, 2, \dots, \quad (37)$$

содержится в [4]. В этом случае мы и будем пользоваться этой теоремой для доказательства теоремы 7, хотя ее утверждение остается справедливым для любого априорного распределения. Для доказательства теоремы 7 нам понадобится также несколько вспомогательных утверждений. Первое из них касается поведения «хвостов» распределения случайной величины

$$E_1 = \sum_{k=1}^{\infty} e^{S_k}, \quad S_k = \sum_{i=1}^k \xi_i.$$

В рамках следующей леммы ξ_k будут предполагаться произвольными независимыми одинаково распределенными случайными величинами, $E\xi_i < 0$, $\varphi(\lambda) = E e^{\lambda \xi_1} < \infty$ при некотором $\lambda > 0$.

Лемма 3. Пусть $\mu = \sup\{\lambda: \varphi(\lambda) \leq 1\}$. Тогда

$$\ln \mathbf{P}(\ln E_1 > x) = -\mu x(1 + o(1)). \quad (38)$$

Если

$$\varphi(\mu) = 1, \quad \varphi'(\mu) < \infty, \quad (39)$$

то

$$\mathbf{P}(\ln E_1 > x) = ce^{-\mu x}(1 + o(1)), \quad 0 < c < \infty. \quad (40)$$

Доказательство. Рассмотрим последовательность

$$W_n = \ln(e^{\xi_n} + e^{\xi_n + \xi_{n-1}} + \dots + e^{\xi_n + \dots + \xi_1}). \quad (41)$$

Очевидно, что она удовлетворяет рекурсивным соотношениям

$$\begin{aligned} e^{W_n} &= e^{\xi_n} + e^{\xi_n} e^{W_{n-1}}, \\ W_n &= W_{n-1} + \xi_n + \ln(1 + e^{-W_{n-1}}) \end{aligned} \quad (42)$$

и является цепью Маркова. Распределение W_n при любом начальном значении W_1 (не обязательно $W_1 = \xi_1$) сходится к распределению $W = \ln E_1$, которое будет стационарным распределением цепи, так что цепь $\{W_n\}$ является эргодической. Приращение этой цепи $\xi(x) = W_n - W_{n-1}$ при условии $W_{n-1} = x$ имеет вид $\xi(x) = \xi_n + \ln(1 + e^{-x})$. Это означает, что цепь $\{W_n\}$ является асимптотически однородной (в пространстве; см. [12], [13]) и к ней применима теорема 3 из [13], в силу которой выполнено (38). Чтобы применить теорему 4 из [13] о точной асимптотике $\mathbf{P}(W > x)$ (из которой будет следовать (40) при $0 \leq c < \infty$), наряду с условиями (39) требуется также выполнение условий

$$\int_{-\infty}^{\infty} e^{\mu t} |\mathbf{P}(\xi + \ln(1 + e^{-y}) > t) - \mathbf{P}(\xi > t)| dt \leq l(y), \quad (43)$$

где $l(y)$ — правильно меняющаяся функция такая, что

$$\int_0^{\infty} l(y) dy < \infty. \quad (44)$$

Разность под знаком интеграла в (43) неотрицательна, и поэтому нам достаточно оценить разность интегралов

$$\int_{-\infty}^{\infty} e^{\mu t} \mathbf{P}(\xi > t) dt = \frac{1}{\mu} \int_{-\infty}^{\infty} e^{\mu t} d\mathbf{P}(\xi < t) = \frac{\varphi(\mu)}{\mu}$$

и

$$\int_{-\infty}^{\infty} e^{\mu t} \mathbf{P}(\xi + h(y) > t) dt = \frac{\varphi(\mu) e^{\mu h(y)}}{\mu}, \quad h(y) = \ln(1 + e^{-y}),$$

равную

$$\frac{1}{\mu} \varphi(\mu) (1 - e^{\mu h(y)}) \leq \frac{1}{\mu} \varphi(\mu) (1 - (1 + e^{-y})^\mu) \leq \varphi(\mu) e^{-y}.$$

Это гарантирует выполнение условий (43), (44). Положительность c в (40) следует из неравенства $\ln E_1 \geq S$ и асимптотических представлений в лемме. Лемма доказана.

Лемма 4. Если $p > 0$, $p \rightarrow 0$, то

$$p \sum_{k=1}^{\infty} (1-p)^{k-1} \ln k = -\ln p + c + o(1),$$

где $c = \int_0^\infty e^{-y} \ln y dy$.

Доказательство носит элементарный характер и немедленно получается, если воспользоваться представлением

$$\int_k^{k+1} e^{-x\Delta} \ln x dx = e^{-k\Delta} \ln k \left(1 + O(\Delta) + O\left(\frac{1}{k}\right) \right),$$

где $\Delta = -\ln(1-p) = p + O(p^2)$.

Рассмотрим теперь последовательность геометрических априорных распределений Q (37) при $p \rightarrow 0$. Рассмотрим байесовский момент остановки $\tau(A)$, определенный в теореме 8, и для заданного $\varepsilon > 0$ выберем A так, чтобы $P(\tau(A) < \theta) = P_Q(\tau(A) < \theta) = \varepsilon$. Это будет означать, что $\tau(A) \in K_{(\varepsilon, Q)}$. Как уже отмечалось, при малых $q(n) = p(1-p)^{n-1}$ выполняется $A \approx 1 - \varepsilon$.

Нам понадобится асимптотика $E(\tau(A) - \theta)^*$ при $p \rightarrow 0$.

Лемма 5.

$$E(\tau(A) - \theta)^+ = \frac{(1-\varepsilon)(-\ln p - 1 + V)}{\rho} + u_p + o(1), \quad (45)$$

где

$$V = \ln \frac{A}{1-A} - E \ln(1 + E_1 + E_2), \quad \sup_p u_p < c < \infty.$$

Если распределение ζ_i нерешетчато, то $u_p \rightarrow (1-\varepsilon)E\chi < \infty$, где χ — перескок через бесконечно удаленный барьер в блуждании $\{-Z_k\}$.

Доказательство. Имеем

$$\begin{aligned} E(\tau(A) - \theta)^+ &= p \sum_{t=1}^{\infty} (1-p)^{t-1} E_t(\tau(A) - t)^+ \\ &= p \sum_{t=1}^{\infty} (1-p)^{t-1} P_t(\tau(A) \geq t) E_t(\tau(A) - t | \tau(A) \geq t). \end{aligned} \quad (46)$$

Наша ближайшая цель — найти асимптотику $E_t(\tau(A) - t \mid \tau(A) \geq t)$ при $t \rightarrow \infty$. Рассмотрим более внимательно момент остановки

$$\begin{aligned} \tau(A) &= \min \left\{ n \geq 1: \frac{\sum_{k=1}^n (1-p)^{k-1} R(k)}{(1-p)^n R(n+1)} \geq \frac{A}{p(1-A)} \right\} \\ &= \min \{ n \geq 1: M_n \geq e^h \}, \end{aligned} \quad (47)$$

где

$$h = \ln \frac{A}{p(1-A)},$$

$$M_n = (1-p)r_n^{-1} + (1-p)^2(r_n r_{n-1})^{-1} + \dots + (1-p)^n(r_n \dots r_1)^{-1}. \quad (48)$$

В нашем случае

$$\begin{aligned} r_k^{-1} &= e^{\xi_k} \quad \text{при } k < t, \\ r_k^{-1} &= e^{-\zeta_k} \quad \text{при } k \geq t. \end{aligned}$$

Поэтому, полагая $\Delta = -\ln(1-p) = p + O(p^2)$, получим

$$\begin{aligned} M_n &= e^{-\zeta_n - \Delta} + e^{-\zeta_n - \zeta_{n-1} - 2\Delta} + \dots + e^{-\zeta_n - \dots - \zeta_t - (n-t+1)\Delta} \\ &\quad + e^{-\zeta_n - \dots - \zeta_t - (n-t+1)\Delta + \xi_{t-1} - \Delta} + \dots \\ &\quad + e^{-\zeta_n - \dots - \zeta_t - (n-t+1)\Delta + \xi_{t-1} + \dots + \xi_1 - (t-1)\Delta} \\ &= e^{-\zeta_n - \dots - \zeta_t - (n-t+1)\Delta} [1 + e^{\zeta_t + \Delta} + \dots + e^{\zeta_t + \dots + \zeta_{n-1} + (n-t)\Delta} \\ &\quad + e^{\xi_{t-1} - \Delta} + \dots + e^{\xi_{t-1} + \dots + \xi_1 - (t-1)\Delta}]. \end{aligned}$$

Пользуясь независимостью $\{\xi_i\}$ и $\{\zeta_i\}$ и сдвигая индексы у ζ_i на $t-1$, получим

$$\tau(A) - t + 1 \stackrel{d}{=} \tau \equiv \min \{ k \geq 1: -Z_k^\Delta + \ln[1 + E_{1,t-1}^\Delta + E_{2,k-1}^\Delta] \geq h \},$$

где

$$\begin{aligned} Z_k^\Delta &= Z_k + k\Delta, & S_k^\Delta &= S_k - k\Delta, \\ E_{1,n}^\Delta &= \sum_{k=1}^n e^{S_k^\Delta}, & E_{2,n}^\Delta &= \sum_{k=1}^n e^{Z_k^\Delta}. \end{aligned}$$

Таким образом, величина $\tau(A) - t + 1$ при условии $\tau(A) \geq t$ имеет такое же распределение, как момент остановки τ для мартингала $\{-Z_k^\Delta - k(\rho - \Delta)\}$. Нетрудно видеть, что здесь, как и для величины η в доказательстве теоремы 5, применимо тождество Вальда-Дуба, в силу которого $-EZ_\tau^\Delta = E\tau(\rho - \Delta)$. Отсюда вытекает $E\tau = -EZ_\tau^\Delta / (\rho - \Delta)$, где $-Z_\tau^\Delta = h - \ln[1 + E_{1,t}^\Delta + E_{2,\tau-1}^\Delta] + \chi_t$.

Здесь «перескок» χ_t имеет тот же смысл, что в доказательстве теоремы 5, и для $E\chi_t$ справедливо то же утверждение, что в теореме 5.

Кроме того, $E_{i,k}^{\Delta} \uparrow E_{i,\infty}^{\Delta} < \infty$ и в силу леммы 2: $E \ln[1 + E_{1,\infty}^{\Delta} + E_{2,\infty}^{\Delta}] < \infty$ при достаточно малых Δ ; $E_{i,\infty}^{\Delta} \rightarrow E_i$ при $\Delta \rightarrow 0$, $i = 1, 2$. Поэтому

$$E\tau = \frac{h - E \ln[1 + E_1 + E_2] + E\chi_t}{\rho} + o(1),$$

где $\sup_{t,\Delta} E\chi_t < c < \infty$ и для нерешетчатых ζ_k выполняется $E\chi_t = E\chi + o(1)$ при $t \rightarrow \infty$, $\Delta \rightarrow 0$ (см. [14]). Остается заметить, что в силу (46)

$$\begin{aligned} E(\tau(A) - \theta)^+ &= p \sum_{t=1}^{\infty} (1-p)^{t-1} P_t(\tau(A) \geq t) E\tau \\ &= \frac{h - E \ln[1 + E_1 + E_2]}{\rho} P_t(\tau(A) \geq t) + u_p + o(1) \\ &= \frac{-\ln p + \ln A/(1-A) - E \ln[1 + E_1 + E_2]}{\rho} (1 - \varepsilon) \\ &\quad + u_p + o(1), \end{aligned}$$

где $\sup_p u_p < \infty$, $u_p = (1 - \varepsilon)E\chi + o(1)$ для нерешетчатых ζ_k . Лемма доказана.

Мы можем обратиться теперь непосредственно к доказательству теоремы 7.

Доказательство теоремы 7. Возьмем произвольный асимптотически однородный момент остановки $\tau \in K_{(\varepsilon_1)}$, $\varepsilon_1 < \varepsilon$. Тогда $P_t(\tau < t) \rightarrow \varepsilon_1$ при $t \rightarrow \infty$, и, стало быть, $P_Q(\tau < \theta) \rightarrow \varepsilon_1$ при $p \rightarrow 0$. Это означает, что $\tau \in K_{(\varepsilon, Q)}$ при всех достаточно малых p . Имеем в силу леммы 3

$$\begin{aligned} E(\tau - Q)^+ &= p \sum_{t=1}^{\infty} (1-p)^{t-1} E_t(\tau - t)^+ \\ &= p \sum_{t=1}^{\infty} (1-p)^{t-1} P_t(\tau \geq t) v \ln t (1 + o(1)) \\ &= -v(\ln p)(1 - \varepsilon_1)(1 + o(1)). \end{aligned}$$

Следовательно, в силу теоремы 8 при $A = A_\varepsilon$ и леммы 4

$$\begin{aligned} v(1 - \varepsilon_1) &= \lim_{p \rightarrow 0} \frac{E_Q(\tau - \theta)^+}{-\ln p} \geq \lim_{p \rightarrow 0} \frac{E_Q(\tau(A) - \theta)^+}{-\ln p} = \frac{(1 - \varepsilon)}{\rho}, \\ v &\geq \frac{(1 - \varepsilon)}{(1 - \varepsilon_1)\rho}. \end{aligned}$$

Так как ε_1 — произвольное число, меньшее ε , то мы получаем $v \geq 1/\rho$. Теорема доказана.

Из теоремы 5 видно, что полученная нижняя граница $1/\rho$ достигается для момента остановки ν , построенного в (24). Сказанное выше означает также, что ν асимптотически эквивалентен (в смысле приведенных утверждений) байесовскому моменту $\tau(A)$ относительно априорного распределения Q (37) при $p \rightarrow 0$. Нетрудно убедиться, что это останется справедливым и для других распределений Q из класса $\mathcal{Q}$, описанного в п. 2.

З а м е ч а н и е 2. Предположение о непрерывности распределения ξ_i (в области $(-\infty, 0)$; см., например, теорему 5 и замечание 1) делалось нами, чтобы обеспечить непрерывность функции $P(\infty, N)$. Эти предположения о непрерывности ξ_i в теоремах 5, 8 не носят принципиального характера и могут быть обойдены путем следующей рандомизации процедуры принятия решения. Пусть, например, мы находимся в условиях теоремы 5 и монотонно убывающая по N функция

$$P(\infty, N) = P_{\infty}(\nu(N) < \infty) = P\left(\bigcup_k \{w_k \geq b(k) + N\}\right)$$

такова, что уравнение $P(\infty, N) = \varepsilon$ для заданного фиксированного $\varepsilon > 0$ не имеет решения. Это означает, что найдется точка $N(\varepsilon)$, в которой $P(\infty, N)$ имеет скачок,

$$P(\infty, N(\varepsilon) - 0) = \varepsilon_1 > \varepsilon, \quad P(\infty, N(\varepsilon) + 0) = \varepsilon_2 < \varepsilon.$$

Введем случайную величину $\nu(N(\varepsilon), p)$, которая при первом попадании w_k в область $(b(k) + N(\varepsilon), \infty)$ ведет себя как $\nu(N)$. Если же $w_k = b(k) + N(\varepsilon)$, то «тревога» о разладке объявляется с вероятностью p независимо от X ; с дополнительной вероятностью $1 - p$ испытания продолжают по той же схеме: либо до первого попадания w_k в $(b(k) + N(\varepsilon), \infty)$ (и тогда объявляется тревога), либо до следующего попадания w_k в точку $b(k) + N(\varepsilon)$, и тогда повторяется процедура «случайного розыгрыша» решения, описанная выше. Очевидно, $P_{\infty}(\nu(N(\varepsilon), 1) < \infty) = P_{\infty}(\nu(N(\varepsilon)) < \infty) = \varepsilon_1 > \varepsilon$, $P_{\infty}(\nu(N(\varepsilon), 0) < \infty) = P_{\infty}(\nu(N(\varepsilon) + 0) < \infty) = \varepsilon_2 > \varepsilon$. Кроме того, $P_{\infty}(\nu(N(\varepsilon), p) < \infty)$ непрерывным и монотонным образом зависит от p . Поэтому найдется $p(\varepsilon) \in (0, 1)$ такое, что $P_{\infty}(\nu(N(\varepsilon), p(\varepsilon)) < \infty) = \varepsilon$.

Далее, $\nu(N, p)$ есть вновь момент остановки, но уже относительно последовательности $(x_1, \omega_1), \dots, (x_n, \omega_n)$, где ω_i принимают значения 1 с вероятностью p и 0 с вероятностью $1 - p$ независимо от X и других значений ω_k . Нетрудно проверить, что все остальные рассуждения в доказательстве теоремы 5 сохраняют свою силу применительно к $\nu(N, p)$.

Аналогичные модификации можно осуществить и относительно момента остановки $\tau(A)$ в теореме 8, после которых уравнение $\alpha(A) = \varepsilon$ всегда будет иметь решение.

Таким образом, мы можем считать, что предположения о непрерывности ξ_i в теоремах 5, 8 не ограничивают общности.

4. Еще одна конструкция асимптотически оптимального момента остановки. Эта конструкция возникает естественным образом из приведенных в доказательстве теоремы 7 (см. также лемму 4) рассуждений, из которых видно, что оптимальное правило строится на последовательности статистик M_n (см. (47), (48)), которые асимптотически близки при $p \rightarrow 0$ к последовательности

$$M_n^0 = r_n^{-1} + (r_n r_{n-1})^{-1} + \dots + (r_n \dots r_1)^{-1}.$$

Из доказательства леммы 2 видно, что статистики $W_n = \ln M_n^0$, допускающие рекурсивное представление

$$W_{n+1} = -\ln r_{n+1} + \ln(1 + e^{W_n}) = W_n + \gamma_{n+1} + \ln(1 + e^{-W_n}), \quad (49)$$

очень близки по своим свойствам к последовательности $\{w_n\}$, определенной в (22). Аналогично тому, как это делалось в п. 3, мы можем определить момент остановки

$$\nu_1 = \nu_1(N) = \min \{k \geq 1: W_k \geq b(k) + N\}, \quad (50)$$

где функция $b(k)$ имеет тот же вид, что и в теореме 5.

Момент остановки ν_1 при $b(k) \equiv 0$ был предложен в [7]. Относительно него в [8], [9] установлено, что он является асимптотически оптимальным в следующем смысле: он минимизирует главную часть асимптотики $\sup_t \mathbf{E}_t(\nu_1 - t \mid \nu_1 \geq t)$ при $B \rightarrow \infty$ в классе моментов τ , для которых $\mathbf{E}_\infty \tau \geq B$.

Для ν_1 справедлив следующий аналог теоремы 5.

Теорема 9. При всех t

$$\mathbf{P}_t(\nu_1 < t) \leq c_1 b e^{-N}, \quad c_1 = \text{const} < \infty. \quad (51)$$

Если распределения ξ_k непрерывны, то $P_1(t) = P_1(t, N) = \mathbf{P}_t(\nu_1 < t)$ и $P_1(\infty)$ непрерывно зависят от N и для любого $\varepsilon < \varepsilon_0 = \mathbf{P}(\bigcup_{k=1}^{\infty} \{W_k \geq b(k)\})$ существует $N_1 = N_1(\varepsilon)$ такое, что $P_1(t) \uparrow P_1(\infty) = \varepsilon$ при $t \rightarrow \infty$. Кроме того,

$$\mathbf{E}_t(\nu_1 - t \mid \nu_1 \geq t) = \frac{b(t) + N - \mathbf{E} \ln(1 + E_1 + E_2) + \mathbf{E} \chi}{\rho} + o(1), \quad (52)$$

где χ для нерешетчатых ξ_k имеет тот же смысл, что и в теореме 5. Для нерешетчатых ξ_k значение $\mathbf{E} \chi$ следует заменить функцией, ограниченной известной постоянной (той же, что и в теореме 5).

Как и в теоремах 5, 8, условие о непрерывности распределения ξ_k здесь не является существенным.

Доказательство теоремы 9. Если $\theta = \infty$, то $\gamma_k = \xi_k$ и последовательность W_n образует эргодическую цепь Маркова, рассмотренную в доказательстве леммы 2. Распределение W_n монотонно сходится к распределению $W = \ln M_\infty^0 =_d \ln E_1$.

Очевидно, что в нашем случае $\gamma = \xi$ выполнены условия второй части леммы 2 при $\mu = 1$, и, стало быть,

$$\begin{aligned} P_1(t, N) &= P_t(\nu_1 < t) = P\left(\bigcup_{k=1}^t \{W_k \geq b(k) + N\}\right) \\ &\leq \sum_{k=1}^{\infty} P(W_k \geq b(k) + N) \leq \sum_{k=1}^{\infty} c_1 e^{-b(k)-N} (1 + o(1)) \\ &= c_1 b e^{-N} (1 + o(1)). \end{aligned}$$

Это доказывает (51) при, возможно, несколько ином значении постоянной c_1 . Непрерывная зависимость $P_1(t, N)$ от N доказывается так же, как в теореме 5.

Если $\theta = t$, то $\nu_1 - t$ при условии $\nu_1 \geq t$ распределено так же, как время τ первого прохождения уровня $b(t+k) + N$ блужданием W_k при $\gamma_j = -\zeta_j$ и с начальным значением W_0 , распределение которого монотонно сходится при $t \rightarrow \infty$ к распределению $\ln E_1$. В этом случае приращения $\xi(x) = W_{k+1} - W_k$ при условии $W_k = x$ превосходят $-\zeta_{k+1}$, а в области $W_k \geq M$ допускают мажоранту $\xi(x) \leq -\zeta_{k+1} + e^{-M}$. Отсюда и из рассмотрений [14] нетрудно получить существование предельного значения $E\chi$ в нерешетчатом случае и ограниченность соответствующей функции в общем случае. Далее,

$$\begin{aligned} e^{W_n} &= e^{Z_n + W_0} + e^{Z_n} + e^{Z_n - Z_1} + \dots + e^{Z_n - Z_{n-1}} \\ &= e^{Z_n} (1 + E_1 + E_2 + \varepsilon_n), \end{aligned}$$

где $\varepsilon_n \rightarrow_p 0$. Таким образом, W_n ведет себя асимптотически как $Z_n + \ln(1 + E_1 + E_2)$, где E_1 и E_2 независимы. С помощью тождества Дуба-Вальда находим, как и прежде,

$$\begin{aligned} E\tau &= \frac{EZ_\tau}{\rho} = \rho^{-1} E[b(t + \tau) + N_1 - \ln(1 + E_1 + E_2) + E\chi_t] + o(1) \\ &= \rho^{-1} [b(t) + N_1 - E \ln(1 + E_1 + E_2) + E\chi_t + o(1)], \end{aligned}$$

где χ_t имеет прежний смысл величины перескока, $E\chi_t \rightarrow E\chi$ в нерешетчатом случае. Теорема доказана.

Представляет интерес сравнение средних значений задержек моментов остановки ν и ν_1 в условиях теорем 5 и 9 (см. правые части в (26) и (52)) при фиксированной вероятности ложной тревоги ε . Однако сделать это довольно трудно, так как вид функций $N(\varepsilon)$ и $N_1(\varepsilon)$ в этих теоремах нам неизвестен. Мы можем лишь высказать некоторые соображения, которые позволят хотя бы приближенно понять природу этих функций. Заметим, что для функции $P(\infty, N)$ в теореме 5 вместо (25) при больших N можно использовать в силу леммы 2 более точную оценку

$$P(\infty, N) \approx c b e^{-N}, \quad (53)$$

где постоянная c — из леммы 2. Слабая зависимость событий $A_n = \{w_n < b(n) + N\}$ и A_{n+k} при больших k и сравнение с некоторыми близкими результатами о распределении $\max_{j \leq n} w_j$ (см. [15]) позволяют сделать предположение, что асимптотическое поведение $P(\infty, N)$ отличается от (53) лишь постоянным множителем K , так что

$$P(\infty, N) \approx Kcbe^{-N}. \quad (54)$$

Такое же замечание справедливо относительно функции

$$P_1(\infty, N) \approx K_1c_1be^{-N}, \quad (55)$$

где c_1 — из леммы 2, при этом в силу большого сходства в устройстве последовательностей $\{w_k\}$ и $\{W_k\}$ значения множителей K и K_1 близки друг к другу. Из соотношений (54), (55) следует, что при малых ε выполняется

$$N_1(\varepsilon) - N(\varepsilon) \approx \ln c_1 - \ln c.$$

Взяв разность между правыми частями в (26) и (52), мы получим, что «выигрыш» при использовании ν_1 по сравнению с ν будет измеряться величиной

$$\mathbf{E} \ln(1 + E_1 + E_2) - \mathbf{E} \max(S^+, Z^+) - \ln \frac{c_1}{c}. \quad (56)$$

Если эта величина окажется значимой положительной, то момент остановки ν_1 следует считать более предпочтительным. Относительно разности (56) заметим, что в ней, очевидно,

$$\ln(1 + E_1) > S^+, \quad \ln(1 + E_1 + E_2) > \max(S^+, Z^+), \quad c_1 > c.$$

Теоремы 2–9 дают нам вид асимптотически оптимальных статистик, которые следует использовать в задачах о разладке при больших θ . Однако теоремы 5, 9 оставляют существенный произвол в выборе статистической процедуры. Речь идет о выборе взаимосвязанных компонент N и $b(k)$ криволинейной границы $b(k) + N$. Если θ очень велико, то, по-видимому, предпочтение следует отдавать «пологим» границам с функцией $b(k)$, рассмотренной в теореме 5. Если же θ не очень велико, то для таких границ слагаемое $N(\varepsilon)$ в оценке (26) может оказаться существенно больше $b(\theta)$. Поэтому для уменьшения $N(\varepsilon)$, возможно, следует выбирать более быстро растущую функцию $b(t)$, например, $(1 + \beta) \ln t$ или даже βt . Отыскание оптимального соотношения между $b(t)$ и N представляет собой трудную задачу, так как, как уже отмечалось, явный вид зависимости $N(\varepsilon)$ от ε и функции $b(t)$ нам неизвестен.

Автор признателен В. И. Лотову за информацию, касающуюся библиографии.

СПИСОК ЛИТЕРАТУРЫ

1. Page E. S. Continuous inspection schemes. — *Biometrika*, 1954, v. 41, № 1, p. 100–115.
2. Lorden G. Procedures for reacting to a change in distribution. — *Ann. Math. Statist.*, 1971, v. 42, № 6, p. 1897–1908.
3. Moustakides G. V. Optimal stopping times for detecting changes in distributions. — *Ann. Statist.*, 1986, v. 14, № 4, p. 1379–1387.
4. Ширяев А. Н. Статистический последовательный анализ. М.: Наука, 1976, 272 с.
5. Дарховский Б. С., Бродский Б. Е. Непараметрический метод скорейшего обнаружения изменений среднего случайной последовательности. — *Теория вероятн. и ее примен.*, 1987, т. XXXII, в. 4, с. 703–711.
6. Ключене Н., Телькснис Л. Методы обнаружения момента изменения свойств случайных последовательностей. — *Автоматика и телемеханика*, 1983, № 10, с. 5–56.
7. Roberts S. W. A comparison of some control chart procedures. — *Technometrics*, 1966, v. 8, p. 411–430.
8. Pollak M. Optimal detection of a change in distribution. — *Ann. Statist.*, 1985, v. 13, № 1, p. 206–227.
9. Pollak M. Average run length of an optimal method of detecting a change in distribution. — *Ann. Statist.*, 1987, v. 15, № 2, p. 749–779.
10. Боровков А. А. Теория вероятностей. М.: Наука, 1986.
11. Боровков А. А. Вероятностные процессы в теории массового обслуживания. М.: Наука, 1972, 368 с.
12. Borovkov A. A. Ergodicity and Stability of Stochastic Processes. New York: John Wiley, 1998.
13. Боровков А. А., Коршунов Д. А. Вероятности больших отклонений одномерных цепей Маркова. Часть I. Стационарные распределения. — *Теория вероятн. и ее примен.*, 1996, т. 41, в. 1, с. 3–30.
14. Боровков А. А., Фосс С. Г. Оценки для эксцесса (перескока) случайного блуждания через произвольную границу и их применения. — *Теория вероятн. и ее примен.* (в печати).
15. Iglehart D. Extreme values in the GI/G/1 queue. — *Ann. Math. Statist.*, 1972, v. 43, № 2, p. 627–635.
16. Боровков А. А. Математическая статистика. М.: Наука, 1984, 143 с.
17. Ширяев А. Н. Минимаксная оптимальность метода кумулятивных сумм (CUSUM) в случае непрерывного времени. — *Успехи матем. наук*, 1996, т. 51, № 4, с. 173–174.
18. Carlstein E., Muller H.-G., Sigmund D. (ed.). Change-Point Problems. Institute of Math. Statist. Lecture Notes — Monograph ser., V. 23, 1994.

Поступила в редакцию
27.X.1997