POLITECHNIKA WROCŁAWSKA
WYDZIAŁ PODSTAWOWYCH PROBLEMÓW TECHNIKI

KIERUNEK: Matematyka
SPECJALNOŚĆ: Statystyka matematyczna

PRACA DYPLOMOWA

Estymacja parametrów rozkładu na podstawie danych otrzymywanych w chwilach losowych

Tomasz Suchocki

PROMOTOR: **dr Alicja Jokiel–Rokita**

WROCŁAW, czerwiec 2006

Spis treści

Spis rysunków

Wstęp

W pracy przedstawiono metody wyznaczania planów sekwencyjnych w przypadku estymacji parametru rzeczywistego na podstawie danych otrzymywanych w chwilach losowych.

Praca składa się z 4 rozdziałów.

W rozdziale pierwszym został omówiony i przedstawiony model, którym będziemy posługiwać się w dalszej części pracy. Zostały zaprezentowane również definicje i fakty, które wykorzystywane będą w następnych rozdziałach.

W rozdziale drugim został przedstawiony problem wyznaczania optymalnego planu sekwencyjnego dla pewnej klasy rozkładów pochodzących z wykładniczej rodziny rozkładów oraz dla kwadratowej funkcji straty z wagą. Plany wyznaczyliśmy w przypadkach, gdy

- rozkład chwil pojawiania się danych jest znany,
- rozkład chwil pojawiania się danych jest rozkładem wykładniczym z nieznanym parametrem, o którym zakładamy, że ma rozkład a priori gamma.

Problematyka zawarta w tym rozdziale rozpatrywana była w pracy Magiery (1982).

W rozdziale trzecim omówiliśmy problem wyznaczenia optymalnego planu sekwencyjnego w przypadku estymacji średniej rozkładu normalnego, gdy za funkcję straty związaną z błędem estymacji przyjmujemy asymetryczną funkcję straty LINEX. Ponownie wyznaczyliśmy dwie klasy optymalnych planów sekwencyjnych dla znanego i nieznanego rozkładu zmiennych, będących chwilami pojawiania się danych.

W ostatnim rozdziale wyznaczyliśmy optymalne plany sekwencyjne w estymacji średniej rozkładu normalnego przyjmując ograniczoną funkcję straty („reflected normal”). Przedstawione w tym rozdziale rezultaty uzyskane zostały przy wykorzystaniu tych samych metod co w rozdziale drugim i trzecim.

Rozdział 1

Model oraz ogólne założenia

Niech $(\Omega, \mathcal{F}, P)$ będzie przestrzenią probabilistyczną. Przez $(\mathcal{X}, \mathcal{B})$ oznaczmy przestrzeń mierzalną, gdzie $\mathcal{X} \subseteq \mathbf{R}$, a $\mathcal{B}$ jest borelowskim σ-ciałem podzbiorów $\mathcal{X}$. Rozważmy niezależne zmienne losowe $X_1, X_2, \ldots, X_n$ zdefiniowane na $(\Omega, \mathcal{F}, P)$ o wartościach w $(\mathcal{X}, \mathcal{B})$ o jednakowym rozkładzie $P_\theta \in \mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Załóżmy, że Θ jest przedziałem otwartym, a wszystkie rozkłady $P_\theta \in \mathcal{P} = \{P_\theta : \theta \in \Theta\}$ są absolutnie ciągłe względem σ-skończonej miary określonej na $(\mathcal{X}, \mathcal{B})$ oraz, że $E_\theta(X_i^2) < \infty$ dla każdego $\theta \in \Theta$. Załóżmy, że zmienne losowe X_i są obserwowane w chwiliach t_i, gdzie $t_i = U_{i:n}$ (i-ta statystyka pozycyjna), $i = 1, 2, \ldots, n$. Dodatkowo załóżmy, że $U_1, U_2, \ldots, U_n$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie zadanym dystrybuantą G oraz, że $U_1, U_2, \ldots, U_n$ są niezależne od $X_1, X_2, \ldots, X_n$.

Rozpatrzmy problem sekwencyjnej estymacji bayesowskiej parametru θ, przy założeniu, że θ ma rozkład a priori π, gdzie wszystkie parametry tego rozkładu są znane.

Niech

$$k(t) = \sum_{i=1}^{n} \mathbf{1}_{[0,t]}(U_i), \tag{1.1}$$

oznacza liczbę obserwacji dostępną do chwili t włącznie oraz niech

$$\mathcal{F}_t = \sigma\{k(s), s \leq t, X_1, X_2, \ldots, X_{k(t)}\}, \tag{1.2}$$

oznacza informację dostępną w chwili t. Niech $\mathcal{D}$ będzie klasą funkcji d, $\mathcal{F}_t$ mierzalnych, których wartości, dla konkretnej realizacji będziemy przyjmować za oszacowanie parametru θ.

Przy założeniu, że obserwacje zatrzymamy w chwili t, całkowita strata jaką poniesiemy będzie określona następującym wzorem

$$L_t(\theta, d) = L(\theta, d) + c_A k(t) + c(t), \tag{1.3}$$

gdzie θ jest prawdziwą wartością szacowanego parametru, a d estymatorem tegoż parametru. Ponadto

1. $L(\theta, d)$ jest stratą związaną z błędem estymacji parametru θ,
2. c_A jest kosztem związanym z otrzymaniem pojedynczej obserwacji,
3. $c(t)$ jest kosztem związanym z prowadzeniem obserwacji do chwili t.

Załóżmy, że funkcja $c(t)$ jest różniczkowalna, wypukła oraz $c(0) = 0$.

Za czas zatrzymania, będziemy uważać zmienną losową τ, dla której prawdopodobieństwo $P_\theta(0 \leq \tau < \infty) = 1$ dla każdej wartości parametru $\theta \in \Theta$ oraz $\{\tau < t\} \in \mathcal{F}_t$ dla każdego $t \geq 0$. Parę $\delta = (\tau, d(\tau))$ nazywamy procedurą sekwencyjną, natomiast funkcję $R(\theta, \delta)$, określoną następującym wzorem

$$R(\theta, \delta) = E_\theta[L_\tau(\theta, d(\tau))], \tag{1.4}$$

funkcją ryzyka procedury δ.

Bedziemy rozważać jedynie te procedury sekwencyjne δ, dla których funkcja $R(\theta, \delta) < \infty$ dla każdego $\theta \in \Theta$.

Wprowadźmy zmienną losową Y o wartościach $\theta \in \Theta$. Niech $\mathcal{M}$ oznacza σ-ciało podzbiorów Θ, a π rozkład a priori zmiennej losowej Y na przestrzeni mierzalnej $(\Theta, \mathcal{M})$. Zakładamy, że $Y = \theta$. Ryzykiem bayesowskim procedury sekwencyjnej $\delta = (\tau, d(\tau))$ względem rozkładu a priori π nazywamy

$$r(\pi, \delta) = E^\pi[R(\theta, \delta)] = \int_\Theta R(\theta, \delta) d\pi(\theta). \tag{1.5}$$

Przez π będziemy rozumieć zarówno rozkład a priori parametru θ jak i jego funkcję gęstości prawdopodobieństwa.

Jeżeli dla rozkładu a priori π, rozkład a posteriori $\pi_t = (\pi \mid \mathcal{F}_t)$ względem $\mathcal{F}_t$ jest dobrze określony, to ryzyko a posteriori dla procedury sekwencyjnej $\delta = (\tau, d(\tau))$ ma postać

$$\widetilde{R}(\pi, d(\tau))) = E^{\pi_t}[L_\tau(\theta, d(\tau))] = \int_\Theta L_\tau(\theta, d(\tau)) d\pi_t(\theta \mid x). \tag{1.6}$$

W dalszej części pracy wyznaczymy bayesowską procedurę estymacji parametru θ przy powyższych założeniach, tzn. parę $\delta^* = (\tau^*, d^*(\tau^*))$, która minimalizuje ryzyko bayesowskie postaci (1.5) po wszystkich chwilach zatrzymania τ i estymatorach $d(\tau)$, tzn.

$$r(\pi, \delta^*) = \inf_\delta r(\pi, \delta).$$

Wiadomo, że rozwiązanie tego problemu składa się z dwóch etapów:

1. wyznaczenie estymatora bayesowskiego dla dowolnej chwili zatrzymania τ;

2. wyznaczenie optymalnej chwili zatrzymania τ^*, która minimalizuje (po wszystkich chwilach zatrzymania) wartość oczekiwaną całkowitej straty (ryzyko a posteriori postaci (1.6) związane z wyznaczonym w punkcie 1 estymatorem). (Metody wyznaczania optymalnej chwili zatrzymania przedstawił Ross (1971)).

Niech h oznacza funkcję rzeczywistą określoną na zbiorze $E_n = \{0, 1, \ldots, n\}$, taką że $0 \le h(k) < \infty$ dla każdego $k \in E_n$ oraz niech

$$\mathcal{L}_h(t) = \mathcal{L}_h(k(t), t) = h(k(t)) + c(t),$$

$t \ge 0$, oznacza stratę związaną z zatrzymaniem obserwacji w chwili t. Będziemy rozpatrywać jedynie te procedury sekwencyjne, które spełniają tzw. przypadek monotoniczny, który zakłada, że jeżeli $\mathcal{L}_h$ pierwszy raz zacznie przyjmować niesatysfakcjonujące nas wyniki to już zawsze będzie je przyjmowała. Jezeli $\mathcal{L}_h(t)$ spełnia przypadek monotoniczny, wtedy optymalną chwilę zatrzymania łatwo uzyskać z tożsamości Dynkina

$$E[\mathcal{L}_h(\tau)] - h(k(0)) = E\left\{\int_0^\tau [\mathcal{A}_t h(k(t)) + c^{'}(t)]dt\right\}, \tag{1.7}$$

gdzie $\mathcal{A}_t h(k(t))$ jest operatorem infinitezymalnym, tzn.

$$\mathcal{A}_t h(k(t)) = \lim_{\Delta \to 0^+} \frac{E[h(k(t+\Delta)) - h(k(t)) \mid k(t) = k]}{\Delta}, \tag{1.8}$$

która zachodzi dla chwil zatrzymania τ względem $\mathcal{F}_t$, takich że $E(\tau) < \infty$ oraz funkcji h należących do dziedziny operatora infinitezymalnego (zobacz Dynkin (1965), Aven, Jensen (1999) oraz Shapiro i Wardrop (1980)).

Problematyka ta została zapoczątkowana przez Starra i in. (1976), a następnie rozwinięta i uogólniona przez Magierę (1982) i Jokiel-Rokitę, Magierę (1999).

Rozdział 2

Estymacja parametru θ przy ważonej kwadratowej funkcji straty

W rozdziale tym zakładamy, że rozkłady $P_\theta \in \mathcal{P} = \{P_\theta : \theta \in \Theta\}$ należą do wykładniczej rodziny rozkładów $\mathcal{E}(\theta, \alpha)$, definiowanej następująco:

Za $\mathcal{E}(\theta, \alpha)$ będziemy uważać rodzinę rozkładów $P_\theta \in \mathcal{P} = \{P_\theta : \theta \in \Theta\}$, dla której gęstość prawdopodobieństwa względem σ-skończonej miary v wyraża się wzorem

$$\frac{dP_\theta}{dv}(x) = p(x; \theta, \alpha) = s(x, \alpha) \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right], \tag{2.1}$$

gdzie α jest stałą dodatnią, $s(x, \alpha)$ funkcją nieujemną, mierzalną, niezależną od parametru θ oraz $w_1(\theta)$ i $w_2(\theta)$ są funkcjami określonymi na Θ, dwukrotnie różniczkowalnymi w zbiorze Θ z pierwszymi pochodnymi $w_1'(\theta)$, $w_2'(\theta)$ takimi, że $w_2'(\theta) > 0$ i $\dfrac{w_1'(\theta)}{w_2'(\theta)}$ jest ściśle malejąca na Θ.

Fakt 2.1 *Wartość oczekiwana i wariancja zmiennych losowych X_i, $i = 1, 2, \ldots, n$, o rozkładzie zadanym przez* (2.1) *wyraża się wzorem*

$$E_\theta(X_i) = -\alpha \frac{w_1'(\theta)}{w_2'(\theta)} \tag{2.2}$$

i

$$Var_\theta(X_i) = -\frac{\alpha}{w_2'(\theta)} \frac{d}{d\theta} \left[\frac{w_1'(\theta)}{w_2'(\theta)}\right]. \tag{2.3}$$

Dowód:

Podstawiając $w_2(\theta) := \lambda$ otrzymujemy jednoparametrową rodzinę wykładniczą rozkładów w postaci kanonicznej o gęstości postaci

$$p(x;\theta,\alpha) = s(x,\alpha)\exp\left[\alpha w_1(\theta) + x w_2(\theta)\right] = s(x,\alpha)\exp\left\{\lambda x - \left[-\alpha w_1(w_2^{-1}(\lambda))\right]\right\},$$

gdzie $T(X) = x$, $\Phi(\lambda) = -\alpha w_1(w_2^{-1}(\lambda))$, a nieznanym parametrem jest parametr λ. Korzystając z twierdzenia o postaci wartości oczekiwanej i wariancji zmiennej losowej należącej do naturalnej rodziny wykładniczej rozkładów (zobacz np. Magiera (2002), Twierdzenie 2.16) mamy

$$E_\theta(X_i) = \frac{d}{d\lambda}\Phi(\lambda) = -\frac{d}{d\lambda}\alpha w_1(w_2^{-1}(\lambda)) = -\alpha\frac{w_1'(\theta)}{w_2'(\theta)},$$

oraz

$$\mathrm{Var}_\theta(X_i) = \frac{d^2}{d\lambda^2}\Phi(\lambda) = \frac{d}{d\lambda}\left(-\alpha\frac{w_1'(w_2^{-1}(\lambda))}{w_2'(w_2^{-1}(\lambda))}\right) = -\frac{\alpha}{w_2'(\theta)}\frac{d}{d\theta}\frac{w_1'(\theta)}{w_2'(\theta)}.$$

□

Do rodziny $\mathcal{E}(\theta,\alpha)$ należą między innymi rozkład normalny $\mathcal{N}(\alpha\theta,\alpha)$ dla $\theta \in (-\infty,\infty)$, rozkład gamma $\mathcal{G}(\theta^{-1},\alpha)$, Poissona $\mathcal{P}(\alpha\theta)$ dla $\theta \in (0,\infty)$.

2.1 Estymacja parametru θ w przypadku, gdy dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest znana

Problemy poruszane w tym podrożdziale zostały rozwiązane i opisane w pracy Magiery (1982).

Niech Θ będzie otwartym przedziałem $(a,b) \in \mathbf{R}$. Chcemy znaleźć optymalną procedurę sekwencyjną $\delta^* = (\tau^*, d^*(\tau^*))$ dla $\theta \in \Theta$ bazującą na niezależnych zmiennych losowych $X_1, X_2, \ldots, X_n$ o rozkładzie $P_\theta \in \mathcal{E}(\theta,\alpha)$ spełniającym następujące postulaty:

1. dla każdego $\theta \in \Theta$
$$\theta = -\frac{w_1'(\theta)}{w_2'(\theta)}; \tag{2.4}$$

2. istnieje stała $\beta \geq 0$ taka, że
$$\int_\Theta \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right] d\theta = \frac{1}{(\alpha - \beta)s(x, \alpha)} \tag{2.5}$$
zachodzi dla każdego $\alpha > \beta$ i $x \in \mathcal{X}$ takiego, że $s(x, \alpha) > 0$;

3. dla każdego $\alpha > \beta$ oraz $x \in \mathcal{X}$ poza $x = \inf \mathcal{X}$,
$$\lim_{\theta \to a^+} \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right] = \lim_{\theta \to b^-} \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right] \tag{2.6}$$
i
$$\lim_{\theta \to a^+} \theta \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right] = \lim_{\theta \to b^-} \theta \exp\left[\alpha w_1(\theta) + x w_2(\theta)\right]. \tag{2.7}$$

Przyjmujemy, że $\pi(\theta)$ jest postaci
$$\pi(\theta) = \alpha_0 p(\gamma; \theta, \alpha_0 + \beta) = \alpha_0 s(\gamma, \alpha_0 + \beta) \exp\left[(\alpha_0 + \beta) w_1(\theta) + \gamma w_2(\theta)\right], \tag{2.8}$$
gdzie $\alpha_0 > 0$ i γ są znanymi stałymi, natomiast funkcja s przyjmuje jedynie wartości dodatnie. Zauważmy, że $\pi(\theta)$ jest gęstością prawdopodobieństwa na Θ, gdyż
$$\int_\Theta \pi(\theta) d\theta = 1.$$

Na podstawie (2.5) mamy
$$\begin{aligned} \int_\Theta \pi(\theta) d\theta &= \alpha_0 s(\gamma, \alpha_0 + \beta) \int_\Theta \exp\left[(\alpha_0 + \beta) w_1(\theta) + \gamma w_2(\theta)\right] d\theta \\ &= \alpha_0 s(\gamma, \alpha_0 + \beta) \frac{1}{(\alpha_0 + \beta - \beta) s(\gamma, \alpha_0 + \beta)} = 1. \end{aligned}$$

Niech $\mathcal{E}_0(\alpha_0, \gamma)$ oznacza rodzinę rozkładów prawdopodobieństwa na Θ z funkcją gęstości postaci (2.8), wtedy otrzymujemy następujący lemat.

Lemat 2.1 *Niech $P_\theta \in \mathcal{E}(\theta, \alpha)$ oraz niech zachodzi (2.5). Jeżeli $\pi \in \mathcal{E}_0(\alpha_0, \gamma)$, wtedy*
$$\pi_t \in \mathcal{E}_0(\alpha_0 + \alpha k(t), \gamma + \sum_{i=1}^{k(t)} X_i). \tag{2.9}$$

Dowód:

Wiemy, że

$$\pi_t(\theta) = \frac{\prod_{i=1}^{k(t)} p(x;\theta,\lambda)\cdot\pi(\theta)}{m(x)},$$

gdzie

$$m(x) = \int_\Theta \prod_{i=1}^{k(t)} p(x;\theta,\lambda)\cdot\pi(\theta)d\theta$$

Ze wzorów (2.1) i (2.8) mamy

$$\pi_t(\theta) = \frac{\exp\left[(\alpha_0+\beta+\alpha k(t))w_1(\theta)+(\gamma+\sum_{i=1}^{k(t)} X_i)w_2(\theta)\right]}{\int_\Theta \exp\left[(\alpha_0+\beta+\alpha k(t))w_1(\theta)+(\gamma+\sum_{i=1}^{k(t)} X_i)w_2(\theta)\right]d\theta}.$$

Wykorzystując (2.5) otrzymujemy

$$\pi_t(\theta) = [\alpha_0+\alpha k(t)]s(\gamma+\sum_{i=1}^{k(t)} X_i, \alpha_0+\beta+\alpha k(t))\exp\Bigg[(\alpha_0+\beta+\alpha k(t))w_1(\theta)$$
$$+(\gamma+\sum_{i=1}^{k(t)} X_i)w_2(\theta)\Bigg],$$

co daje tezę lematu.

□

Wniosek 2.1 *Dla dowolnej chwili zatrzymania τ*

$$\pi_\tau \in \mathcal{E}_0(\alpha_0+\alpha k(\tau), \gamma+\sum_{i=1}^{k(\tau)} X_i). \tag{2.10}$$

Wniosek wynika z Lematu 2.1 oraz mocnego prawa Markowa.

Na podstawie (2.4) i (2.6) otrzymujemy następujące zależności

$$\begin{aligned}&\alpha\int_\Theta \theta w_2'(\theta)\exp\left[\alpha w_1(\theta)+xw_2(\theta)\right]d\theta\\ &= x\int_\Theta w_2'(\theta)\exp\left[\alpha w_1(\theta)+xw_2(\theta)\right]d\theta\end{aligned} \tag{2.11}$$

oraz

$$\begin{aligned}\int_{\Theta} \theta(x-\alpha\theta)w_2'(\theta)\exp\left[\alpha w_1(\theta)+xw_2(\theta)\right]d\theta \\ = -\int_{\Theta} \exp\left[\alpha w_1(\theta)+xw_2(\theta)\right]d\theta. \qquad (2.12)\end{aligned}$$

Z (2.5), (2.11) oraz (2.12) mamy

$$\int_{\Theta}(x-\alpha\theta)^2 w_2'(\theta)\exp\left[\alpha w_1(\theta)+xw_2(\theta)\right]d\theta = \frac{\alpha}{(\alpha-\beta)s(x,\alpha)}. \qquad (2.13)$$

Załóżmy, że strata $L(\theta,d)$ związana z błędem estymacji parametru θ jest postaci

$$L(\theta,d) = w_2'(\theta)(\theta-d)^2. \qquad (2.14)$$

Lemat 2.2 *Niech $P_\theta \in \mathcal{E}(\theta,\alpha)$ oraz niech zachodzą warunki* (2.4), (2.5) *oraz* (2.6). *Wtedy dla funkcji straty określonej przez* (2.14) *i dowolnej chwili zatrzymania τ estymator bayesowski $d^*(\tau)$ parametru θ względem rozkładu a priori π ma postać*

$$d^*(\tau) = \frac{\gamma+\sum_{i=1}^{k(\tau)} X_i}{\alpha_0+\beta+\alpha k(\tau)}, \qquad (2.15)$$

a ryzyko a posteriori jest postaci

$$\widetilde{R}(\pi, d(\tau)) = \frac{1}{\alpha_0+\alpha k(\tau)+\beta}. \qquad (2.16)$$

Dowód:

Wiemy, że estymator bayesowski parametru θ przy kwadratowej funkcji straty z wagą wyznacza się ze wzoru (zobacz np. Krzyśko (1998), Twierdzenie 5.9).

$$d^*(\tau) = \frac{\int_{\Theta}\theta w_2'(\theta)\pi_t d\theta}{\int_{\Theta} w_2'(\theta)\pi_t d\theta}$$

Oznaczmy przez

$$\begin{aligned}\alpha_t &= \alpha_0+\alpha k(t),\\ \gamma_t &= \gamma + \sum_{i=1}^{k(t)} X_i.\end{aligned}$$

Korzystając z tego, że $\pi_t \in \mathcal{E}_0(\alpha_t, \gamma_t)$ mamy

$$d^*(\tau) = \frac{\displaystyle\int_\Theta \theta w_2'(\theta) \exp\left[(\alpha_t + \beta) w_1(\theta) + \gamma_t w_2(\theta)\right] d\theta}{\displaystyle\int_\Theta w_2'(\theta) \exp\left[(\alpha_t + \beta) w_1(\theta) + \gamma_t w_2(\theta)\right] d\theta}.$$

Następnie korzystając ze wzoru (2.11) otrzymujemy postać (2.15) estymatora $d^*(\tau)$. Wiemy, że

$$\widetilde{R}(\pi, d^*(\tau)) = E^{\pi_t}[L_\tau(\theta, d^*(\tau))] = E^{\pi_t}[w_2'(\theta)(d^*(\tau) - \theta)^2]$$

$$= \int_\Theta \left(\frac{\gamma_t}{\alpha_t + \beta} - \theta\right)^2 w_2'(\theta)[\alpha_t s(\gamma_t, \alpha_t + \beta)] \exp\left[(\alpha_t + \beta) w_1(\theta) + \gamma_t w_2(\theta)\right] d\theta$$

$$= \frac{\alpha_t s(\gamma_t, \alpha_t + \beta)}{(\alpha_t + \beta)^2} \int_\Theta (\gamma_t - (\alpha_t + \beta)\theta)^2 w_2'(\theta) \exp\left[(\alpha_t + \beta) w_1(\theta) + \gamma_t w_2(\theta)\right] d\theta.$$

Na mocy (2.13) mamy

$$\widetilde{R}(\pi, d^*(\tau)) = \frac{\alpha_t s(\gamma_t, \alpha_t + \beta)(\alpha_t + \beta)}{(\alpha_t + \beta)^2 s(\gamma_t, \alpha_t + \beta)\alpha_t} = \frac{1}{\alpha_t + \beta},$$

co kończy dowód.

□

Z powyższego lematu wynika, że problem sprowadza się wyznaczenia takiej chwili zatrzymania τ^*, która będzie minimalizowała wartość oczekiwaną całkowitej straty

$$E\left[\frac{1}{\alpha_0 + \alpha k(\tau) + \beta} + c_A k(\tau) + c(\tau)\right], \tag{2.17}$$

po wszystkich chwilach zatrzymania τ.

Niech G będzie dystrybuantą niezależnych zmiennych losowych $U_1, U_2, \ldots, U_n$. Zakładamy, że $G(0) = 0$, $G(t) > 0$ dla $t > 0$ oraz, że G jest absolutnie ciągła oraz posiada funkcję gęstości g, która jest prawostronnie różniczkowalna na przedziale $(0, \infty)$. Klasę takich dystrybuant oznaczmy przez $\mathcal{G}$. Niech $\xi = \sup\{t : G(t) < 1\}$ oraz niech $\rho(z) = g(z)[1 - G(z)]^{-1}$, $0 \le z < \xi$ oznacza funkcję intensywności awarii.

Fakt 2.2 *Niech* $G \in \mathcal{G}$, $\xi = \sup\{t : G(t) < 1\}$ *oraz niech* $\rho(z) = g(z)[1 - G(z)]^{-1}$, $0 \leq z < \xi$ *oznacza funkcję intensywnoiści awarii. Wtedy* $k(t)$, $0 \leq t < \xi$, *jest niestacjonarnym łańcuchem Markowa ze względu na* $\mathcal{F}_t$, $0 \leq t < \xi$, *z operatorem infinitezymalnym*

$$\mathcal{A}_t(h(k)) = (n-k)\rho(t)[h(k+1) - h(k)], \tag{2.18}$$

dla $k \in E_n = \{0, 1, \ldots, n\}$ *i wszystkich rzeczywistych funkcji* h *na* E_n.

Dowód:
Stąd, że $\forall t \geq 0$ oraz $\forall i \in \mathbb{N}$ mamy

$$\mathbf{1}(U_i \leq t) = \begin{cases} 0 \text{ z prawdopodobieństwem } 1 - G(t), \\ 1 \text{ z prawdopodobieństwem } G(t), \end{cases}$$

to zmienna losowa $k(t) = \sum_{i=1}^{n} \mathbf{1}(U_i \leq t)$ ma rozkład Bernoulliego $\mathcal{B}(n, G(t))$. $\forall \epsilon > 0$ zmienna losowa $k(t+\epsilon)$ ma rozkład Bernoulliego $\mathcal{B}(n, G(t+\epsilon))$, tak więc proces $k(t)$ jest niestacjonarny, gdyż jest „czuły" na przesunięcie ($k(t)$ ma inny rozkład niż $k(t+\epsilon)$).
Aby udowodnić, że $k(t)$ jest łańcuchem Markowa musimy pokazać, że dla $\forall l$ oraz $s_1 \leq s_2 \leq \ldots, \leq s_l \leq t$ oraz $k_1 \leq k_2 \leq \ldots, \leq k_l \leq n$ zachodzi

$$P(k(t) = k \mid k(s_l) = k_l, k(s_{l-1}) = k_{l-1}, \ldots, k(s_1) = k_1) = P(k(t) = k | k(s_l) = k_l).$$

Korzystając ze wzoru Bayesa oraz z postaci rozkładu łącznego $(k(t), k(s_1), \ldots, k(s_l))$ oraz $(k(s_1), \ldots, k(s_l))$ mamy

$$\begin{aligned}
&P(k(t) = k \mid k(s_l) = k_l, k(s_{l-1}) = k_{l-1}, \ldots, k(s_1) = k_1) \\
&= \frac{P(k(t) = k, k(s_1) = k_1, \ldots, k(s_l) = k_l)}{P(k(s_1) = k_1, \ldots, k(s_l) = k_l)} = \frac{\binom{n}{k_1}\binom{n-k_1}{k_2-k_1}\cdots\binom{n-k_l}{k-k_l}}{\binom{n}{k_1}\binom{n-k_1}{k_2-k_1}\cdots\binom{n-k_{l-1}}{k_l-k_{l-1}}} \\
&\times \frac{[G(s_1)]^{k_1}\ldots[G(s_l) - G(s_{l-1})]^{k_l-k_{l-1}}[G(t) - G(s_l)]^{k-k_l}[1 - G(t)]^{n-k}}{[G(s_1)]^{k_1}\ldots[G(s_l) - G(s_{l-1})]^{k_l-k_{l-1}}[1 - G(s_l)]^{n-k_l}} \\
&= \frac{\binom{n-k_l}{k-k_l}[G(t) - G(s_l)]^{k-k_l}[1 - G(t)]^{n-k}}{[1 - G(s_l)]^{n-k_l}} \\
&= \frac{\binom{n}{k_l}\binom{n-k_l}{k-k_l}[G(s_l)]^{k_l}[G(t) - G(s_l)]^{k-k_l}[1 - G(t)]^{n-k}}{\binom{n}{k_l}[G(s_l)]^{k_l}[1 - G(s_l)]^{n-k_l}}
\end{aligned}$$

$$= \frac{P(k(t) = k, k(s_l) = k_l)}{P(k(s_l) = k_l)} = P(k(t) = k \mid k(s_l) = k_l).$$

Tak więc pokazaliśmy, że proces $k(t)$ jest niestacjonarnym łańcuchem Markowa. Wyprowadzimy teraz wzór na postać operatora infinitezymalnego procesu $k(t)$. W tym celu ustalmy $k \in \{0, 1, \ldots, n\}$. Jest oczywiste, że

$$E\left[h(k(t+s)) - h(k) \mid k(t) = k\right] = \sum_{i=k+1}^{n} \left[h(i) - h(k)\right] P\big(k(t+s) = i \mid k(t) = k\big)$$

$$= \left[h(k+1) - h(k)\right] P\big(k(t+s) = k+1 \mid k(t) = k\big)$$

$$+ \sum_{i=k+2}^{n} \left[h(i) - h(k)\right] P\big(k(t+s) = i \mid k(t) = k\big)$$

$$= \left[h(k+1) - h(k)\right](n-k)\frac{G(t+s) - G(t)}{1 - G(t)} \left[\frac{1 - G(t+s)}{1 - G(t)}\right]^{n-k-1}$$

$$+ \sum_{i=k+2}^{n} \left[h(i) - h(k)\right] P\big(k(t+s) = i \mid k(t) = k\big)$$

$$\leq \left[h(k+1) - h(k)\right](n-k)\frac{G(t+s) - G(t)}{1 - G(t)} \left[\frac{1 - G(t+s)}{1 - G(t)}\right]^{n-k-1}$$

$$+2 \sup_{i \leq n} \mid h(i) \mid P\big(k(t+s) \geq k+2 \mid k(t) = k\big)$$

$$= \left[h(k+1) - h(k)\right](n-k)\frac{G(t+s) - G(t)}{1 - G(t)} \left[\frac{1 - G(t+s)}{1 - G(t)}\right]^{n-k-1}$$

$$+2 \sup_{i \leq n} \mid h(i) \mid \left\{1 - \left[\frac{1 - G(t+s)}{1 - G(t)}\right]^{n-k} \left[1 - (n-k)\frac{G(t+s) - G(t)}{[1 - G(t+s)]}\right]\right\}.$$

Łatwo zauważyć, że

$$\lim_{s \to 0^+} \frac{E\left[h(k(t+s)) - h(k) \mid k(t) = k\right]}{s} = (n-k)\left[h(k+1) - h(k)\right]\rho(t),$$

co dowodzi tezę lematu.

□

Problem sprowadził się zatem do znalezienia optymalnej chwili zatrzymania pewnej funkcji niestacjonarnego łańcucha Markowa. W naszym przypadku funkcją tą jest

$$h(k(t)) = \frac{1}{\alpha_0 + \alpha k(t) + \beta} + c_A k(t).$$

Twierdzenie 2.1 *Załóżmy, że* $G \in \mathcal{G}$ *oraz* G *ma nierosnącą funkcję intensywności awarii* ρ. *Jeśli* $\pi \in \mathcal{E}(\alpha_0, \gamma)$, $0 < \alpha_0 < \infty$, *wtedy* $\delta^* = (\tau^*, d^*(\tau^*))$, *gdzie*

$$\begin{aligned}\tau^* = \inf\{t \geq 0 \ : \ & [n - k(t)]\rho(t)\{\alpha - c_A\{\alpha_0 + \beta + \alpha[k(t) + 1]\}[\alpha_0 + \beta + \alpha k(t)]\} \\ & \leq c^{'}(t)\{\alpha_0 + \beta + \alpha[k(t) + 1]\}[\alpha_0 + \beta + \alpha k(t)]\},\end{aligned} \tag{2.19}$$

a $d^*(\tau^*)$ *ma postać* (2.15) *jest bayesowskim planem sekwencyjnym.*

Dowód:

Pokażemy, że przy powyższych założeniach

$$\tau^* = \{t \geq 0 : \mathcal{A}_t h[k(t)] + c^{'}(t) \geq 0\}$$

jest optymalną chwilą zatrzymania, co przy operatorze infinitezymalnym postaci (2.18) da nam tezę twierdzenia. Korzystając z tożsamości Dynkina wiemy, że dla dowolnej chwili zatrzymania τ zachodzi

$$E(h[k(\tau)] + c(\tau)) - h(0) = E\left\{\int_0^{\tau} (\mathcal{A}_t h[k(t)] + c^{'}(t))dt\right\}.$$

Przy założeniu, że $\rho(t)$ jest nierosnąca, operator infinitezymalny $\mathcal{A}_t h(k(t)) = (n - k(t))\,[h(k(t)+1) - h(k(t))]\,\rho(t)$ jest nierosnącą funkcją t. Zatem dla dowolnej chwili zatrzymania τ mamy

$$\begin{aligned}&E(h[k(\tau^*)] + c(\tau^*)) - E(h[k(\tau)] + c(\tau)) \\ &= \int_{\{\tau^* > \tau\}} \left\{\int_{\tau}^{\tau^*} (\mathcal{A}_t h[k(t)] + c^{'}(t))dt\right\} dP - \int_{\{\tau > \tau^*\}} \left\{\int_{\tau^*}^{\tau} (\mathcal{A}_t h[k(t)] + c^{'}(t))dt\right\} dP \leq 0,\end{aligned}$$

co daje nam optymalność chwili zatrzymania τ^*.

□

2.2 Estymacja parametru θ w przypadku, gdy dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ nie jest znana

Rozważmy problem wyznaczenia optymalnej procedury sekwencyjnej $\delta^* = (\tau^*, d^*(\tau^*))$, przy założeniu, że $U_1, U_2, \ldots, U_n$ tworzą ciąg niezależnych zmiennych losowych o jednakowym rozkładzie zadanym dystrybuantą G pochodzącą z rozkładu wykładniczego zależnego od nieznanego parametru w, gdzie $W = w$ jest zmienną losową posiadającą rozkład a priori gamma $\mathcal{G}(\nu, \lambda)$. Dla $w > 0$ gęstość zmiennej losowej W ma postać

$$\zeta(w) = \frac{\lambda^\nu}{\Gamma(\nu)} w^{\nu-1} \exp(-\lambda w). \tag{2.20}$$

Fakt 2.3 *Rozkład a posteriori zmiennej losowej W względem $\mathcal{F}_t$ jest rozkładem gamma $\mathcal{G}(\nu_t, \lambda_t)$, gdzie*

$$\nu_t = \nu + k(t) \quad i \quad \lambda_t = \lambda + \sum_{j=1}^{k(t)} t_j + [n - k(t)]t. \tag{2.21}$$

Dowód:

Niech g_w i G_w oznaczają odpowiednio funkcję gęstości i dystrybuantę zmiennych losowych $U_1, U_2, \ldots, U_n$, pod warunkiem, że w jest znanym parametrem. Przy założeniu, że $U_i \sim \mathcal{E}(w)$ mamy

$$g_w(t) = w \exp(-wt) \mathbf{1}_{(0,\infty)}(t),$$

$$G_w(t) = 1 - \exp(-wt), \quad t \geq 0.$$

Zatem rozkład a posteriori $\pi_t(W \mid \mathcal{F}_t)$ zmiennej losowej $W = w$ ma gęstość postaci

$$\begin{aligned}
&\pi(w \mid \mathcal{F}_t) = \pi(w \mid t_1, \ldots, t_n, k(t) = k) \\
&= \frac{ng_w(t_1)(n-1)g_w(t_2)\ldots(n-k+1)g_w(t_k)[1-G_w(t)]^{n-k}\zeta(w)}{\int_0^\infty ng_w(t_1)(n-1)g_w(t_2)\ldots(n-k+1)g_w(t_k)[1-G_w(t)]^{n-k}\zeta(w)dw} \\
&= \frac{w^k \exp\left[-w(\sum_{i=1}^k t_i + (n-k)t)\right] w^{\nu-1}\exp(-\lambda w)}{\int_0^\infty w^k \exp\left[-w(\sum_{i=1}^k t_i + (n-k)t)\right] w^{\nu-1}\exp(-\lambda w)dw} \\
&= \frac{w^{\nu+k-1}\exp\left[-(\sum_{i=1}^k t_i + (n-k)t + \lambda)w\right]}{\int_0^\infty w^{\nu+k-1}\exp\left[-(\sum_{i=1}^k t_i + (n-k)t + \lambda)w\right]} \\
&= \frac{(\sum_{i=1}^k t_i + (n-k)t + \lambda)^{k+\nu}}{\int_0^\infty w^{k+v-1}\exp(-w)dw} w^{\nu+k-1}\exp\left[-(\sum_{i=1}^k t_i + (n-k)t + \lambda)w\right] \\
&= \frac{(\sum_{i=1}^k t_i + (n-k)t + \lambda)^{k+\nu}}{\Gamma(k+\nu)} w^{\nu+k-1}\exp\left[-(\sum_{i=1}^k t_i + (n-k)t + \lambda)w\right].
\end{aligned}$$

Tak więc rozkład a posteriori π_t jest rozkładem gamma z parametrami $\nu_t = \nu + k(t)$ i $\lambda_t = \lambda + \sum_{i=1}^{k(t)} t_i + (n - k(t))t$, co kończy dowód.

□

Oznaczmy parametry rozkładu a priori (2.20) odpowiednio przez ν_0 i λ_0.

Fakt 2.4 *Proces stochastyczny (ν_t, λ_t), $t \geq 0$ o wartościach w zbiorze $\{\nu_0, \nu_0 + 1, \ldots, \nu_0 + n\} \times (0, \infty)$ jest stacjonarnym łańcuchem Markowa ze względu na $\mathcal{F}_t$ oraz jego operator infinitezymalny jest postaci*

$$\mathcal{A}H(\nu, \lambda) = [H(\nu+1, \lambda) - H(\nu, \lambda)](m-\nu)\nu\lambda^{-1} + (m-\nu)H_\lambda'(\nu, \lambda), \quad (2.22)$$

gdzie $m = \nu_0 + n$. Dziedzina operatora $\mathcal{A}$ zawiera wszystkie funkcje H, które są różniczkowalne w sposób ciągły względem λ dla każdego ν.

Dowód:
Dowód faktu, który można znaleźć w pracy Stadje (1990) pomijamy.

□

Niech strata związana z obserwacją procesu (ν_t, λ_t), $t \geq 0$, do chwili t będzie postaci

$$L(t) = L(\nu_t, \lambda_t, t) = h(\nu_t) + c(t), \tag{2.23}$$

gdzie $h(\nu)$ jest funkcją określoną na $\{\nu_0, \nu_0 + 1, \ldots, \nu_0 + n\}$, taką że $0 \leq h(\nu) < \infty$ dla $\nu \geq \nu_0$. Jeżeli funkcja $[h(\nu) - h(\nu+1)](m-\nu)\nu$ jest nierosnąca dla $\nu = \nu_0, \nu_0 + 1, \ldots, \nu_0 + n - 1$, wtedy używając tych samych metod co w Rozdziale 2.1 otrzymujemy, że chwila zatrzymania

$$\tau^* = \inf\{t \geq 0 : \mathcal{A}h(\nu_t) + c^{'}(t) \geq 0\}, \tag{2.24}$$

jest optymalna. W szczególności, dla

$$h(\nu_t) = \frac{1}{\alpha_0 + \alpha k(t) + \beta} + c_A k(t) = \frac{1}{\alpha_0 + \alpha(\nu_t - \nu_0) + \beta} + c_A(\nu_t - \nu_0) \tag{2.25}$$

mamy następujące twierdzenie.

Twierdzenie 2.2 *Niech rozkład prawdobodobieństwa zmiennych losowych $U_1, U_2, \ldots,$ U_n spełnia warunki zapisane powyżej oraz niech $\nu_0 \leq n$. Wtedy $\delta^* = (\tau^*, d^*(\tau^*))$, gdzie*

$$\begin{aligned} \tau^* = \inf\{t \geq 0 \ : \ &(m-\nu_t)\nu_t\lambda_t^{-1}\{\alpha - c_A\{\alpha+\beta+\alpha[k(t)+1]\}[\alpha+\beta+\alpha k(t)]\} \\ &\leq c^{'}(t)\{\alpha+\beta+\alpha[k(t)+1]\}[\alpha+\beta+\alpha k(t)]\}, \end{aligned} \tag{2.26}$$

a estymator $d^*(\tau^*)$ ma postać (2.15), jest bayesowskim planem sekwencyjnym.

2.3 Przykład

Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi zmiennymi losowymi mającymi jednakowy rozkład Poissona z parametrem $\alpha\theta$, wtedy

$$p(x; \theta, \alpha) = \frac{(\alpha\theta)^x}{x!} \exp(-\alpha\theta) = \frac{\alpha^x}{x!} \exp[\alpha(-\theta) + x\ln(\theta)],$$

gdzie $x \in \mathbb{N}$.

Niech α=1. Wtedy $s(x,1) = \dfrac{1}{x!}$, $w_1(\theta) = -\theta$ oraz $w_2(\theta) = \ln(\theta)$. Widać, że funkcja $s(x,1)$ jest funkcją nieujemną, mierzalną i niezależną od parametru θ oraz, że funkcje

$w_1(\theta)$ i $w_2(\theta)$ są funkcjami dwukrotnie różniczkowalnymi w zbiorze $\Theta = (0, \infty)$ z pierwszymi pochodnymi $w_1(\theta)'$, $w_2(\theta)'$ takimi, że $w_2(\theta)' > 0$ oraz $\frac{w_1(\theta)'}{w_2(\theta)'}$ jest ściśle malejącą funkcją θ.

Pokażemy, że spełnione są postulaty dotyczące rozkładu $P_\theta \in \mathcal{E}(\theta, \alpha)$. Mianowicie

1. $-\frac{w_1'(\theta)}{w_2'(\theta)} = \theta$,
2. skoro
$$\int_\Theta \frac{1}{x!} \exp[-\theta + x\ln(\theta)] = 1,$$
to
$$\int_\Theta \exp[-\theta + x\ln(\theta)] = x! = \frac{1}{s(x,1)(1-0)},$$
czyli parametr $\beta = 0$,
3. $$\lim_{\theta \to 0^+} \exp[-\theta + x\ln(\theta)] = 0$$
oraz
$$\lim_{\theta \to \infty} \exp[-\theta + x\ln(\theta)] = \lim_{\theta \to \infty} \frac{\theta^x}{\exp(\theta)} = 0,$$
czyli
$$\lim_{\theta \to 0^+} \exp\left[-\theta + x\ln(\theta)\right] = \lim_{\theta \to \infty} \exp\left[-\theta + x\ln(\theta)\right],$$
4. $$\lim_{\theta \to 0^+} \theta \exp[-\theta + x\ln(\theta)] = \lim_{\theta \to 0^+} \exp[-\theta + (x+1)\ln(\theta)] = 0$$
oraz
$$\lim_{\theta \to \infty} \theta \exp[-\theta + x\ln(\theta)] = \lim_{\theta \to \infty} \frac{\theta^{x+1}}{\exp(\theta)} = 0,$$
czyli
$$\lim_{\theta \to 0^+} \theta \exp\left[-\theta + x\ln(\theta)\right] = \lim_{\theta \to \infty} \theta \exp\left[-\theta + x\ln(\theta)\right].$$

Za rozkład a priori π parametru θ przyjmujemy rozkład, którego funkcja gęstości prawdopodobieństwa wyraża się wzorem $\alpha_0 p(\gamma, \theta, \alpha_0 + \beta)$, gdzie $\gamma = 4$, $\alpha_0 = 2$ oraz

$\beta = 0$. Tak więc, funkcja gęstości prawdopodobieństwa dla rozkładu a priori π parametru θ ma postać

$$\alpha_0 p(\gamma, \theta, \alpha_0 + \beta) = \frac{\alpha_0^{\gamma+1}}{\gamma!} \exp[\alpha_0(-\theta) + \gamma \ln(\theta)] = \frac{4}{3} \exp[-2\theta + 4\ln(\theta)].$$

Na podstawie Lematu 2.2 wyznaczamy postać estymatora $d^*(\tau)$ okreslonego wzorem (2.15) i ryzyka a posteriori $\widetilde{R}(\pi, d^*(\tau)))$ określonego wzorem (2.16), mianowicie

$$d^*(\tau) = \frac{4 + \sum_{i=1}^{k(\tau)} X_i}{2 + k(\tau)}$$

i

$$\widetilde{R}(\pi, d^*(\tau))) = \frac{1}{2 + k(\tau)}.$$

Czyli musimy wyznaczyć taką optymalną chwilę zatrzymania τ^*, która będzie minimalizowała wartość oczekiwaną całkowitej straty

$$E\left[\frac{1}{2 + k(\tau)} + c_A k(\tau) + c(\tau)\right],$$

po wszystkich chwilach zatrzymania τ, przy założeniu, że $c_A = \dfrac{1}{100}$, a $c(t) = \dfrac{t^2}{2}$.

1. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest znana i pochodzi z rozkładu wykładniczego z parametrem $\lambda = 1$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 2.1 ma postać

$$\tau^* = \inf\left\{t \geq 0 : [n - k(t)]\left[1 - \frac{1}{100}(3 + k(t))(2 + k(t))\right] \leq t(3 + k(t))(2 + k(t))\right\}.$$

2. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest zadana rozkładem wykładniczym z nieznanym parametrem $\lambda = w$, gdzie zmienna losowa $W = w$ ma rozkład a priori gamma $\mathcal{G}(2, 3)$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 2.2 ma postać

$$\tau^* = \inf\left\{t \geq 0 \;:\; [n - k(t)]\frac{2 + k(t)}{3 + \sum_{j=1}^{k(t)} U_j + [n - k(t)]t}\left[1 - \frac{1}{100}(3 + k(t))(2 + k(t))\right]\right.$$
$$\left. \leq t(3 + k(t))(2 + k(t))\right\}.$$

Rozdział 3

Estymacja średniej z rozkładu normalnego przy funkcji straty LINEX

3.1 Estymacja średniej z rozkładu normalnego w przypadku, gdy dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest znana

Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym o nieznanej średniej θ i znanej wariancji σ^2. Zmienna losowa θ posiada rozkład a priori π, bedący rozkładem normalnym o znanej średniej μ i znanej wariancji η^2.

Za funkcję straty związaną z błędem estymacji nieznanego parametru θ przyjmiemy funkcję straty LINEX, wyrażającą się wzorem

$$L(\theta, d) = b\{\exp\left[a(\theta - d)\right] - a(\theta - d) - 1\}, \tag{3.1}$$

gdzie $a \neq 0$, a $b > 0$. Funkcja straty LINEX jest asymetryczną funkcją różnicy $\theta - d$. Rysunek 3.1 przedstawia wykresy funkcji LINEX dla róznych wartości parametru a i dla $b = 1$. W dalszej części tego rozdziału bez straty ogólności będziemy przyjmować, że $b = 1$.

Rysunek 3.1: Wykres funkcji straty LINEX

Oznaczmy

$$Y_t = \sum_{i=1}^{k(t)} X_i \quad . \tag{3.2}$$

Lemat 3.1 *Dla dowolnej chwili zatrzymania τ, bayesowski estymator $d^*(\tau)$ parametru θ ze względu na rozkład a priori π pod warunkiem $\mathcal{F}_\tau$ jest postaci*

$$d^*(\tau) = \frac{1}{\eta^2 k(\tau) + \sigma^2}(\sigma^2\mu + \eta^2 Y_\tau + \frac{1}{2}a\sigma^2\eta^2), \tag{3.3}$$

a jego ryzyko a posteriori wynosi

$$E[L(\theta, d^*(\tau)) \mid \mathcal{F}_\tau] = \frac{1}{2}a^2\frac{\sigma^2\eta^2}{\eta^2 k(\tau) + \sigma^2}. \tag{3.4}$$

Dowód:

Niech t będzie ustaloną chwilą zatrzymania obserwacji. Wówczas rozkład a posteriori π_t parametru θ przy danym $\mathcal{F}_t$ jest rozkładem normalnym $\mathcal{N}(\mu_t, \eta_t^2)$, gdzie

$$\mu_t = \frac{\sigma^2}{\eta^2 k(t) + \sigma^2}\mu + \frac{\eta^2}{\eta^2 k(t) + \sigma^2}Y_t$$

i

$$\eta_t^2 = \frac{\sigma^2\eta^2}{\eta^2 k(t) + \sigma^2}.$$

Postać estymatora $d^*(\tau)$ otrzymamy z formuły Zellnera. Mianowicie

$$\begin{aligned} d^*(t) &= \frac{1}{a}\ln\left\{E^{\pi_t}[\exp(a\theta)]\right\} = \frac{1}{a}\ln\left\{\int_\Theta \exp(a\theta)\frac{1}{\sqrt{2\pi\eta_t^2}}\exp\left(-\frac{(\theta-\mu_t)^2}{2\eta_t^2}\right)d\theta\right\} \\ &= \frac{1}{a}\ln\left\{\frac{1}{\sqrt{2\pi\eta_t^2}}\int_\Theta \exp\left(-\frac{(\theta-(\mu_t+a\eta_t^2))^2 - a^2\eta_t^4 - 2\mu_t a\eta_t^2}{2\eta_t^2}\right)\right\} \\ &= \frac{1}{a}\ln\left\{\exp(\frac{1}{2}a^2\eta_t^2) + a\mu_t\right\} = \frac{1}{2}a\eta_t^2 + \mu_t. \end{aligned}$$

Ryzyko a posteriori wynosi

$$\begin{aligned} E[L(\theta, d^*(t)) \mid \mathcal{F}_t] &= E^{\pi_t}\{\exp[a(\theta - d^*(t))] - a(\theta - d^*(t)) - 1\} \\ &= \exp(-ad^*(t))E^{\pi_t}(\exp(a\theta)) - aE^{\pi_t}(\theta) + ad^*(t) - 1 \\ &= \exp(-a\mu_t - \frac{1}{2}a^2\eta_t^2)\exp(a\mu_t + \frac{1}{2}a^2\eta_t^2) - a\mu_t + a\mu_t + \frac{1}{2}a^2\eta_t^2 - 1 \\ &= \frac{1}{2}a^2\eta_t^2 = \frac{1}{2}a^2\frac{\sigma^2\eta^2}{\eta^2 k(t) + \sigma^2}. \end{aligned}$$

□

Oznaczmy przez $\epsilon = \frac{\sigma^2}{\eta^2}$. Z powyższego lematu wynika, że problem sprowadza się do wyznaczenia takiej chwili zatrzymania τ^*, która będzie minimalizowała wartość oczekiwaną całkowitej straty

$$E\left[\frac{1}{2}a^2\frac{\sigma^2}{k(\tau) + \epsilon} + c_A k(\tau) + c(\tau)\right], \tag{3.5}$$

po wszystkich chwilach zatrzymania τ.

Ponownie niech G będzie dystrybuantą niezależnych zmiennych losowych U_1, $U_2, \ldots, U_n$. Zakładamy, że $G(0) = 0$, $G(t) > 0$ dla $t > 0$ oraz, że G jest absolutnie ciągła oraz posiada funkcję gęstości g, która jest prawostronnie różniczkowalna na przedziale $(0, \infty)$. Klasę takich dystrybuant będziemy oznaczać przez $\mathcal{G}$. Niech $\xi = \sup\{t : G(t) < 1\}$ oraz niech $\rho(z) = g(z)[1 - G(z)]^{-1}$, $0 \le z < \xi$ oznacza funkcję intensywności awarii.

Analogicznie do Rozdziału 2 stwierdzamy, że $k(t)$ jest niestacjonarnym łańcuchem Markowa ze względu na $\mathcal{F}_t$ z operatorem infinitezymalnym postaci (2.18).

Twierdzenie 3.1 *Załóżmy, że $G \in \mathcal{G}$ ma nierosnącą funkcję intensywności awarii $\rho(t)$. Wtedy $\delta^*(\tau^*, d^*(\tau^*))$, gdzie*

$$\tau^* = \inf\left\{t : [n - k(t)]\rho(t)\left\{\frac{a^2\sigma^2}{2[(k(t) + 1) + \epsilon][k(t) + \epsilon]} - c_A\right\} \le c'(t)\right\}, \tag{3.6}$$

a estymator $d^*(\tau^*)$ ma postać (3.3)*, jest bayesowskim planem sekwencyjnym.*

Dowód:

Dowód twierdzenia, który można przeprowadzić w sposób analogiczny do dowodu Twierdzeniu 2.1 pomijamy.

□

3.2 Estymacja średniej z rozkładu normalnego w przypadku, gdy dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ nie jest znana

Twierdzenie 3.2 *Niech rozkład prawdobodobieństwa zmiennych losowych* $U_1, U_2, \ldots, U_n$ *spełnia warunki zapisane w Podrozdziale 2.2 oraz niech* $\nu \leq n$. *Wtedy* $\delta^* = (\tau^*, d^*(\tau^*))$, *gdzie*

$$\tau^* = inf\left\{t : [n - k(t)]\nu_t \lambda_t^{-1}\left\{\frac{a^2\sigma^2}{2[(k(t)+1)+\epsilon][k(t)+\epsilon]} - c_A\right\} \leq c^{'}(t)\right\}, \quad (3.7)$$

a estymator $d^*(\tau^*)$ *ma postać* (3.3) *jest bayesowskim planem sekwencyjnym.*

Dowód:
Dowód twierdzenia, który można przeprowadzić w sposób analogiczny do dowodu Twierdzeniu 2.2 pomijamy.

□

3.3 Przykład

Będziemy zakładać, że $X_1, X_2, \ldots, X_n$ mają jednakowy rozkład normalny o nieznanej średniej θ i znanej wariancji 1 oraz, że zmienna losowa θ ma rozkład normalny o średniej 0 i wariancji 1.
Za funkcję straty związaną z błędem estymacji parametru θ przyjmiemy funkcję straty LINEX z parametrami $a = 2$ i $b = 1$. Rysunek 3.2 przedstawia wykres tej funkcji jako funkcji różnicy $\theta - d$.

Rysunek 3.2: Wykres funkcji straty LINEX dla $a = 2$ i $b = 1$

Na podstawie **Lematu 3.1** wyznaczamy postać estymatora bayesowskiego $d^*(\tau)$ oraz jego ryzyko a posteriori $E[L(\theta, d^*(\tau)/\mathcal{F}_t)]$, mianowicie

$$d^*(\tau) = \frac{\sum_{i=1}^{k(\tau)} X_i + 1}{k(\tau) + 1}$$

oraz

$$E[L(\theta, d^*(\tau)/\mathcal{F}_t)] = \frac{2}{k(\tau) + 1}.$$

Czyli musimy wyznaczyć taką optymalną chwilę zatrzymania τ^*, która będzie minimalizowała wartość oczekiwaną całkowitej straty

$$E\left[\frac{2}{k(\tau) + 1} + c_A k(\tau) + c(\tau)\right],$$

po wszystkich chwilach zatrzymania τ. W naszym przypadku $c_A = \frac{1}{100}$, a $c(t) = \frac{t^2}{2}$.

1. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest znana i pochodzi z rozkładu wykładniczego z parametrem $\lambda = 1$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 3.1 ma postać

$$\tau^* = \inf\left\{t \geq 0 : [n - k(t)]\left\{\frac{2}{(k(t) + 2)(k(t) + 1)} - \frac{1}{100}\right\} \leq t\right\}.$$

2. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest zadana rozkładem wykładniczym z nieznanym parametrem $\lambda = w$, gdzie zmienna losowa $W = w$ ma rozkład a priori gamma $\mathcal{G}(2, 3)$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 3.2 ma postać

$$\tau^* = \inf\left\{t \geq 0 : \frac{[n - k(t)][2 + k(t)]}{3 + \sum_{j=1}^{k(t)} U_j + [n - k(t)]t}\left[\frac{2}{(k(t) + 2)(k(t) + 1)} - \frac{1}{100}\right] \leq t\right\}.$$

Rozdział 4

Estymacja średniej z rozkładu normalnego przy funkcji straty „reflected normal"

Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym o nieznanej średniej θ i znanej wariancji σ^2. Zmienna losowa θ posiada rozkład a priori π, będący rozkładem normalnym o znanej średniej μ i znanej wariancji η^2.

Analogicznie do Rozdziału 2 sformuujemy twierdzenia o postaci planu sekwencyjnego. Tym razem za funkcję straty związaną z błędem estymacji nieznanego parametru θ przyjmiemy funkcję straty „reflected normal", wyrażającą się wzorem

$$L(\theta, d) = K\left\{1 - \exp\left[-\frac{(d-\theta)^2}{2\gamma^2}\right]\right\}, \tag{4.1}$$

gdzie $\gamma > 0$ i $K > 0$. Współczynnik K nazywany jest parametrem maksymalnej straty. Rysunek 4.1 przedstawia wykres funkcji straty „reflected normal" jako funkcji różnicy $\theta - d$ dla przykładowych wartości parametrów γ i K.

Rysunek 4.1: Wykres funkcji straty „reflected normal”

Lemat 4.1 *Dla dowolnej chwili zatrzymania τ, bayesowski estymator $d^*(\tau)$ parametru θ ze względu na rozkład a priori π pod warunkiem $\mathcal{F}_\tau$ jest postaci*

$$d^*(\tau) = E^{\pi_t}(\theta) = \mu_\tau = \frac{\sigma^2}{\eta^2 k(\tau) + \sigma^2}\mu + \frac{\eta^2}{\eta^2 k(\tau) + \sigma^2} Y_\tau, \tag{4.2}$$

gdzie Y_τ jest postaci (3.2)*, a jego ryzyko a posteriori wynosi*

$$E[L(\theta, d^*(\tau)) \mid \mathcal{F}_\tau] = K\left(1 - \frac{\gamma}{\sqrt{\frac{\eta^2\sigma^2}{k(\tau)\eta^2+\sigma^2} + \gamma^2}}\right). \tag{4.3}$$

Dowód:
Niech t będzie ustaloną chwilą zatrzymania. Wówczas rozkład a posteriori π_t parametru θ przy danym $\mathcal{F}_t$ jest rozkładem normalnym $\mathcal{N}(\mu_t, \eta_t^2)$, gdzie

$$\mu_t = \frac{\sigma^2}{\eta^2 k(t) + \sigma^2}\mu + \frac{\eta^2}{\eta^2 k(t) + \sigma^2} Y_t$$

i

$$\eta_t^2 = \frac{\sigma^2\eta^2}{\eta^2 k(t) + \sigma^2}.$$

Musimy najpierw wyznaczyć postać ryzyka a posteriori dla dowolnego estymatora d, wtedy estymatorem optymalnym d^* będzie ten estymator d, który minimalizuje to ryzyko. Mianowicie

$$\begin{aligned}
E^{\pi_t} L(\theta, d) &= K - E^{\pi_t}\left[\exp\left(-\frac{(\theta-d)^2}{2\gamma^2}\right)\right] \\
&= K - \frac{1}{\sqrt{2\pi}\eta_t}\int_{\infty}^{\infty} \exp\left(-\frac{(\theta-d)^2}{2\gamma^2}\right)\exp\left(-\frac{(\theta-\mu_t)^2}{2\eta_t^2}\right) d\theta \\
&= K - \frac{1}{\sqrt{2\pi}\eta_t}\int_{\infty}^{\infty} \exp\left(-\frac{\theta^2\eta_t^2 - 2\theta d\eta_t^2 + d^2\eta_t^2 + \theta^2\gamma^2 - 2\theta\mu_t\gamma^2 + \mu_t\gamma^2}{2\gamma^2\eta_t^2}\right) d\theta \\
&= K - \frac{1}{\sqrt{2\pi}\eta_t}\int_{\infty}^{\infty} \exp\left(-\frac{\theta^2(\eta_t^2+\gamma^2) - 2\theta(d\eta_t^2 + \mu_t\gamma^2) + d^2\eta_t^2 + \mu_t^2\gamma^2}{2\gamma^2\eta_t^2}\right) d\theta \\
&= K - \gamma\sqrt{\eta_t^2+\gamma^2}\exp\left(-\frac{(d-\mu_t)^2}{2(\eta_t^2+\gamma^2)}\right).
\end{aligned}$$

Łatwo zauważyć, że powyższe wyrażenie jest minimalizowane dla $d = \mu_t$, co potwierdza tezę lematu dotyczącą postaci estymatora d^*. Ryzyko a posteriori wynosi

$$\begin{aligned}
E[L(\theta, d^*(t)) \mid \mathcal{F}_t] &= KE^{\pi_t}\left\{1 - \exp\left(-\frac{(\mu_t-\theta)^2}{2\gamma^2}\right)\right\} \\
&= K\left\{1 - \int_{-\infty}^{\infty} \exp\left(-\frac{(\mu_t-\theta)^2}{2\gamma^2}\right)\frac{1}{\sqrt{2\pi}\eta_t}\exp\left(-\frac{(\theta-\mu_t)^2}{2\eta_t^2}\right)d\theta\right\} \\
&= K\left\{1 - \frac{1}{\sqrt{2\pi}\eta_t}\int_{-\infty}^{\infty} \exp\left(-\frac{\eta_t^2(\theta-\mu_t)^2+\gamma^2(\theta-\mu_t)^2}{2\eta_t^2\gamma^2}\right)d\theta\right\} \\
&= K - \frac{K\gamma}{\sqrt{\eta_t^2+\gamma^2}} = K\left(1 - \frac{\gamma}{\sqrt{\frac{\eta^2\sigma^2}{k(t)\eta^2+\sigma^2}+\gamma^2}}\right).
\end{aligned}$$

□

Twierdzenie 4.1 *Załóżmy, że $G \in \mathcal{G}$ ma nierosnącą funkcję intensywności awarii $\rho(t)$. Wtedy $\delta^*(\tau^*, d^*(\tau^*))$, gdzie*

$$\tau^* = \inf\left\{t \geq 0 : \frac{1}{\sqrt{\frac{\eta^2\sigma^2}{\eta^2(k(t)+1)+\sigma^2} + \gamma^2}} - \frac{1}{\sqrt{\frac{\eta^2\sigma^2}{\eta^2 k(t)+\sigma^2} + \gamma^2}} - c_A \leq \frac{c^{'}(t)}{K\gamma\rho(t)[n-k(t)]}\right\},$$

a estymator $d^*(\tau^*)$ ma postać (4.2), jest bayesowskim planem sekwencyjnym.

Dowód:

Dowód twierdzenia, który można przeprowadzić w sposób analogiczny do dowodu Twierdzeniu 2.1 pomijamy.

□

Twierdzenie 4.2 *Niech rozkład prawdobodobieństwa zmiennych losowych $U_1, U_2, \ldots, U_n$ spełnia warunki zapisane w Podrozdziale 2.2 oraz niech $\nu \leq n$. Wtedy $\delta^* = (\tau^*, d^*(\tau^*))$, gdzie*

$$\tau^* = \inf\left\{t \geq 0 : \frac{1}{\sqrt{\frac{\eta^2\sigma^2}{\eta^2(k(t)+1)+\sigma^2} + \gamma^2}} - \frac{1}{\sqrt{\frac{\eta^2\sigma^2}{\eta^2 k(t)+\sigma^2} + \gamma^2}} - c_A \leq \frac{c^{'}(t)\lambda_t}{K\gamma\nu_t[n-k(t)]}\right\}$$

a estymator $d^*(\tau^*)$ ma postać (4.2), jest bayesowskim planem sekwencyjnym.

Dowód:

Dowód twierdzenia, który można przeprowadzić w sposób analogiczny do dowodu Twierdzeniu 2.2 pomijamy.

□

4.1 Przykład

Będziemy zakładać, że $X_1, X_2, \ldots, X_n$ mają jednakowy rozkład normalny o nieznanej średniej θ i znanej wariancji 1 oraz, że zmienna losowa θ ma rozkład normalny o średniej 0 i wariancji 1.

Za funkcję straty związaną z błędem estymacji parametru θ przyjmiemy symetryczną funkcje straty „reflected normal" z parametrami $K = 2$ i $\gamma = 1$. Rysunek 4.2 przedstawia wykres tej funkcji jako funkcji różnicy $\theta - d$.

Rysunek 4.2: Wykres funkcji straty „reflected normal” dla $K = 2$ i $\gamma = 1$

Na podstawie **Lematu 4.1** wyznaczamy postać estymatora bayesowskiego $d^*(\tau)$ oraz jego ryzyko a posteriori $E[L(\theta, d^*(\tau) \mid \mathcal{F}_t)]$, mianowicie

$$d^*(\tau) = \frac{\sum_{i=1}^{k(\tau)} x_i}{k(\tau) + 1}$$

oraz

$$E[L(\theta, d^*(\tau) \mid \mathcal{F}_t)] = 2\left(1 - \sqrt{\frac{k(\tau)}{k(\tau)+1}}\right).$$

Czyli musimy wyznaczyć taką optymalną chwilę zatrzymania τ^*, która będzie minimalizowała wartość oczekiwaną całkowitej straty

$$E\left[2\left(1 - \sqrt{\frac{k(\tau)}{k(\tau)+1}}\right) + c_A k(\tau) + c(\tau)\right],$$

po wszystkich chwilach zatrzymania τ. W naszym przypadku $c_A = \frac{1}{100}$, a $c(t) = \frac{t^2}{2}$.

1. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest znana i pochodzi z rozkładu wykładniczego z parametrem $\lambda = 1$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 4.1 ma postać

$$\tau^* = \inf\left\{t \geq 0 : [n - k(t)]\left[\sqrt{\frac{k(\tau)+2}{k(\tau)+3}} - \sqrt{\frac{k(\tau)+1}{k(\tau)+2}} - \frac{1}{100}\right] \leq \frac{t}{2}\right\}.$$

2. Zakładamy, że dystrybuanta G zmiennych losowych $U_1, U_2, \ldots, U_n$ jest zadana rozkładem wykładniczym z nieznanym parametrem $\lambda = w$, gdzie zmienna losowa $W = w$ ma rozkład a priori gamma $\mathcal{G}(2, 3)$. Wtedy optymalna chwila zatrzymania τ^* wyznaczona na podstawie Twierdzenia 4.2 ma postać

$$\tau^* = \inf\left\{t \geq 0 : \frac{[n - k(t)][2 + k(t)]}{3 + \sum_{j=1}^{k(t)} U_j + [n - k(t)]t}\left[\sqrt{\frac{k(\tau)+2}{k(\tau)+3}} - \sqrt{\frac{k(\tau)+1}{k(\tau)+2}} - \frac{1}{100}\right] \leq \frac{t}{2}\right\}.$$

Bibliografia

[1] Aven, T. i Jensen, U. (1999). Stochastic models in reliability. *Springer.*

[2] Dynkin, E. B. (1965). Markov Processes. *Springer.*

[3] Fisz, M. (1967). Rachunek prawdopodobieństwa i statystyka matematyczna. *Państwowe Wydawnictwo Naukowe.*

[4] Jokiel-Rokita, A. i Magiera, R. (1999). Estimation with delayed observations for multinomial distribution. *Statistics* **32**, 353–367.

[5] Krzyśko, M. (1996). Statystyka Matematyczna. *Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.*

[6] Krzyśko, M. (1998). Statystyka Matematyczna, Tom II, Statystyczne funkcje decyzyjne. *Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.*

[7] Magiera, R. (1982). Estimation with delayed observations. *Zastosowania Matematyki* **17**, 249–258.

[8] Magiera, R. (2002). Modele i metody statystyki matematycznej. *Oficyna Wydawnicza GiS, Wrocław.*

[9] Ross, S. M. (1971). Infintesimal look-ahead stopping rules. *The Annals of Statistics* **42**, 297–303.

[10] Shapiro, C. P. i Wardrop, R. L. (1980). Dynkin's identity applied to Bayes sequential estimation of a Poisson process rate. *The Annals of Statistics* **8**, 171–182.

[11] Stadje, W. (1990). A Sequential Estimation Procedure for the parameter of an Exponential Distribution. *Statistics* **21**, 239–250.

[12] Starr, N., Wardrop, R. i Woodroofe, M. (1976). Estimating a mean from delayed observations. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.* **35**, 103–113.