



AGATA BORATYŃSKA

Warszawa

Infinitezymalna odporność w bayesowskich modelach statystycznych

(Praca wpłynęła do Redakcji 20.03.1994)

Streszczenie. Praca przedstawia niektóre z problemów dotyczących odporności (badania oscylacji wielkości a posteriori) procedur bayesowskich w wybranych klasach rozkładów a priori. Główna uwaga poświęcona jest odporności infinitesimalnej (ciągłość i różniczkowalność wielkości a posteriori), gdy w przestrzeni rozkładów a priori wprowadzona jest metryka Kołmogorowa. Omówiony jest też istotny problem wyznaczania procedur optymalnych.

1. Wstęp. Termin odporność pojawia się w literaturze statystycznej po raz pierwszy w latach pięćdziesiątych w pracach Boxa (1953) oraz Boxa i Andersena (1955). Od tego czasu powstało szereg publikacji poświęconych badaniu odporności procedur statystycznych. Należy tu wymienić przykładowo prace Hubera (1981), Hampła (1986), Bickela (1976), Zielińskiego (1983).

W pracy rozważamy problem odporności w bayesowskich modelach statystycznych, szczególnie w terminach wrażliwości pewnych wielkości a posteriori na zaburzenia rozkładu a priori. Filozofia i pewne techniczne aspekty tego problemu zostały przedstawione między innymi w pracach Bergera (1980, 1984, 1990) oraz Sivaganesana i Bergera (1989). Dokładne sformułowanie problemu podajemy w rozdziale drugim. Mówiąc najogólniej zadanie polega na:

1. modelowaniu zaburzeń rozkładu a priori za pomocą pewnej rodziny rozkładów Γ ;

2. badaniu oscylacji interesujących nas wielkości a posteriori, gdy rozkład a priori przebiega klasę Γ .

Wybór rodziny rozkładów Γ zależy oczywiście od rozważanego problemu. Nas interesować będą: rodziny otoczeń typu „ ε -zaburzenie”, rodziny roz-

kładów sprzężonych, rodziny rozkładów z ograniczonymi gęstościami (tak zwane „density bounded classes”) i rodziny rozkładów oparte na ilorazie gęstości (tak zwane „density ratio classes”). Zostaną one zaprezentowane w rozdziale trzecim. W rozdziale czwartym pokażemy pewien związek łączący te cztery klasy z otoczeniami w metryce Kołmogorowa; umożliwi to zastosowanie w ich badaniu wyników otrzymanych dla tej ostatniej klasy. Analizę odporności dla pewnych innych rodzin Γ można znaleźć, między innymi, w pracach: Berger (1990), Betro, Męczarski i Ruggeri (1992), O’Hagan i Berger (1988), Ruggeri (1990, 1991) i in.

Wśród interesujących i najczęściej badanych wielkości a posteriori występują: wartość oczekiwana i wariancja rozkładu a posteriori, ryzyko a posteriori oraz prawdopodobieństwo a posteriori ustalonego zbioru. Bada się również tak zwane globalne miary odporności w przestrzeniach rozkładów jakimi są odległości między rozkładami a posteriori odpowiadającymi rozkładom a priori z rodziny Γ , na przykład metryka całkowitego wahania zdefiniowana, dla rozkładów prawdopodobieństwa P i Q , wzorem

$$\sigma(P, Q) = \sup_A |P(A) - Q(A)|,$$

metryka Kołmogorowa, metryka Levy’ego i inne.

W pracy będziemy badać infinitezymalną odporność różnych wielkości a posteriori (dokładne sformułowanie problemu przedstawione jest w rozdziale drugim). W rozważaniach dotyczących odporności procedur statystycznych często odporność utożsamia się z ciągłością danej procedury (por. np. Bednarski (1985)) i badanie odporności sprowadza się do badania odporności infinitezymalnej. Samo stwierdzenie czy procedura jest ciągła nie pozwala jednak na porównywanie różnych procedur i na ocenę, która z nich jest bardziej odporna. Dlatego oprócz stwierdzenia ciągłości badanej wielkości a posteriori będziemy starali się wyznaczyć pochodne i rząd zbieżności funkcji odporności w „zerze”. I tak, w rozdziale 3.1 pokażemy infinitezymalną odporność rozkładu a posteriori w metrykach σ , Levy’ego, Kołmogorowa i Prochorowa. W rozdziale 3.2 przedstawimy rząd infinitezymalnej odporności pewnych estymatorów średniej w modelu normalnym. W rozdziale czwartym udowodnimy, że w pewnych modelach statystycznych operator przyporządkowujący rozkładowi a priori pewien funkcjonal a posteriori jest ciągły i różniczkalny, gdy w przestrzeni rozkładów wprowadzimy metrykę Kołmogorowa, metrykę Levy’ego lub metrykę σ .

Drugą drogą prowadzącą do ilościowego określenia odporności są nierówności określające oscylację badanej wielkości a posteriori. W rozdziale 3.2 przedstawimy przykłady procedur optymalnych w pewnych statystycznych problemach decyzyjnych, to znaczy wskażemy reguły decyzyjne, dla których oscylacja ryzyka a posteriori jest najmniejsza. W rozdziale 4.3 przedstawimy

górne ograniczenie na oscylację w terminach odległości Kolmogorowa rozkładów a priori.

2. Sformułowanie problemu i oznaczenia. Rozważmy bayesowski model statystyczny

$$(\mathcal{X}, \mathcal{F}_{\mathcal{X}}, \Theta, \mathcal{F}_{\Theta}, \{P_{\theta} : \theta \in \Theta\}, \mathcal{P}),$$

gdzie $\mathcal{X}$ jest przestrzenią obserwacji („przestrzenią próby”), $\mathcal{F}_{\mathcal{X}} \subset 2^{\mathcal{X}}$ jest σ -ciałem zdarzeń obserwowalnych, $\{P_{\theta} : \theta \in \Theta\}$ jest rodziną rozkładów na przestrzeni mierzalnej $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$, Θ jest przestrzenią parametrów, $\mathcal{F}_{\Theta} \subset 2^{\Theta}$ jest σ -ciałem podzbiorów tej przestrzeni oraz $\mathcal{P}$ jest rodziną rozkładów na przestrzeni mierzalnej $(\Theta, \mathcal{F}_{\Theta})$. Przestrzeń prób $\mathcal{X}$ będziemy interpretować jako przestrzeń obserwacji pewnej zmiennej losowej X o wartościach $x \in \mathcal{X}$. O rodzinie rozkładów $\{P_{\theta} : \theta \in \Theta\}$ zakładamy, że jest to rodzina dominowana przez pewną σ -skończoną miarę η .

Niech $l : \mathcal{X} \times \Theta \rightarrow \mathcal{R}^1$ oznacza funkcję rzeczywistą, mierzalną względem σ -ciała produktowego $\mathcal{F}_{\mathcal{X}} \times \mathcal{F}_{\Theta}$ taką, że dla każdego $\theta \in \Theta$ funkcja $l(\cdot, \theta)$ jest gęstością rozkładu P_{θ} względem miary η . Dla ustalonego $x \in \mathcal{X}$, funkcję $l_x : \Theta \rightarrow \mathcal{R}^1$, określoną wzorem $l_x(\cdot) = l(x, \cdot)$ nazywamy funkcją wiarygodności lub wiarygodnością (parametru θ , gdy zaobserwowano x).

Rozkłady z rodziny $\mathcal{P}$ nazywamy rozkładami a priori i oznaczamy literami $F, G, Q, \dots$. Prawdopodobieństwo (a priori) zbioru $A \in \mathcal{F}_{\Theta}$ przy rozkładzie F oznaczamy przez $F\{A\}$, a tam, gdzie używamy dystrybuantę tego rozkładu, oznaczamy ją tą samą literą F , tak że $F(z)$ oznacza wartość dystrybuanty rozkładu F w punkcie z . O rodzinie $\mathcal{P}$ zakładamy, że jest to rodzina dominowana przez pewną σ -skończoną miarę μ . Gęstości rozkładów $F, G, Q, \dots$ oznaczamy odpowiednimi małymi literami $f, g, q, \dots$, utożsamiając funkcje różniące się na zbiorze miary μ równej 0.

Niech F będzie ustalonym rozkładem a priori. Rozkład prawdopodobieństwa na Θ postaci

$$F_x(d\theta) = \frac{l_x(\theta)F(d\theta)}{m_x(F)},$$

gdzie

$$m_x(F) = \int_{\Theta} l_x(\theta)F(d\theta),$$

nazywamy rozkładem a posteriori odpowiadającym rozkładowi a priori F przy ustalonej wartości x zmiennej losowej X . Rozkład a posteriori (przy ustalonej wartości $x \in \mathcal{X}$) dla rozkładu a priori $F, G, Q, \dots$ oznaczamy przez $F_x, G_x, Q_x, \dots$

Weźmy ponadto pod uwagę pewną ustaloną przestrzeń mierzalną $(\mathcal{A}, \mathcal{F}_{\mathcal{A}})$, pewien zbiór $\mathcal{D}$ funkcji mierzalnych $\delta : \mathcal{X} \rightarrow \mathcal{A}$ oraz pewną usta-

loną funkcję mierzalną $L : \Theta \times \mathcal{A} \rightarrow \mathcal{R}^1$. Dla ustalonego $a \in \mathcal{A}$ oznaczmy przez $L_a : \Theta \rightarrow \mathcal{R}^1$ funkcję określoną wzorem $L_a(\theta) = L(\theta, a)$.

Przy ustalonych $x \in \mathcal{X}$, L oraz $\delta \in \mathcal{D}$, głównym obiektem naszych zainteresowań będzie funkcjonal $R_x(\cdot, L_{\delta(x)})$, określony na przestrzeni rozkładów a priori $\mathcal{P}$ wzorem

$$R_x(F, L_{\delta(x)}) = \frac{\int_{\Theta} L(\theta, \delta(x)) l_x(\theta) F(d\theta)}{m_x(F)},$$

oraz odwzorowanie $\overline{R_x}$ przestrzeni rozkładów $\mathcal{P}$ w siebie, określone wzorem

$$\overline{R_x}(F) = F_x.$$

(Jest to więc odwzorowanie, które każdemu rozkładowi a priori $F \in \mathcal{P}$ przyporządkowuje odpowiadający mu rozkład a posteriori F_x .) Tam, gdzie to nie powinno budzić niejasności, symbol $R_x(F, L_{\delta(x)})$ będziemy skracali do symbolu $R_x(F, L)$.

W tak ogólnym sformułowaniu obejmujemy, jako przypadki szczególne, takie wielkości a posteriori najczęściej badane w bayesowskiej teorii odporności, jak

1. momenty rozkładu a posteriori (por. np. Sivaganesan i Berger (1989)), przyjmując

$$L(\theta, \delta(x)) = \theta^k, \quad k = 1, 2, \dots;$$

2. prawdopodobieństwo a posteriori danego zbioru $C \in \mathcal{F}_{\Theta}$ (por. np. Berger i Berliner (1986)), przyjmując

$$L(\theta, \delta(x)) = 1_C(\theta),$$

gdzie $1_C(\theta) = 1$, gdy $\theta \in C$, $1_C(\theta) = 0$, gdy $\theta \notin C$;

3. ryzyko a posteriori, interpretując $\mathcal{A}$ jako przestrzeń akcji, $\mathcal{D}$ jako przestrzeń reguł decyzyjnych i $L(\theta, a)$ jako wielkość straty ponoszonej w sytuacji, gdy (nieznana) wartość parametru jest równa $\theta \in \Theta$ i podejmujemy akcję $a \in \mathcal{A}$ (por. np. Betro i Ruggeri (1992)).

Ustalmy pewien rozkład a priori F^0 i rozważmy rodzinę rozkładów a priori $\Gamma = \Gamma(F^0)$, $\Gamma \subset \mathcal{P}$ taką, że $F^0 \in \Gamma$.

Za miarę oscylacji funkcjonalu $R_x(F, L)$, gdy rozkład a priori przebiega rodzinę $\Gamma = \Gamma(F^0)$, przyjmujemy następujące wielkości:

$$\rho_x(\Gamma, R_x) = \sup_{F \in \Gamma} R_x(F, L) - \inf_{F \in \Gamma} R_x(F, L)$$

oraz

$$\rho_x^0(\Gamma, F^0, R_x) = \sup_{F \in \Gamma} |R_x(F, L) - R_x(F^0, L)|.$$

Za miarę oscylacji odwzorowania $\overline{R_x}$ przyjmujemy

$$\rho_x^\nu(\Gamma, F^0) = \sup_{F \in \Gamma} \nu(F_x, F_x^0),$$

gdzie ν oznacza wybraną metrykę w przestrzeni rozkładów $\mathcal{P}$.

Wyznaczenie oscylacji funkcjonału $R_x(F, L)$ dla różnych procedur statystycznych umożliwia ich porównywanie i wybór procedury optymalnej w tym sensie, że jest to procedura o minimalnej oscylacji. Przy badaniu odporności reguł decyzyjnych wykorzystuje się następujące definicje reguł optymalnych.

DEFINICJA 1. Reguła decyzyjna $\tilde{\delta}$ jest optymalna przy ustalonej wartości x zmiennej losowej X wtedy i tylko wtedy, gdy zachodzi związek:

$$\begin{aligned} \inf_{a \in \mathcal{A}} [\sup_{F \in \Gamma} R_x(F, L_a) - \inf_{F \in \Gamma} R_x(F, L_a)] \\ = \sup_{F \in \Gamma} R_x(F, L_{\tilde{\delta}(x)}) - \inf_{F \in \Gamma} R_x(F, L_{\tilde{\delta}(x)}). \end{aligned}$$

DEFINICJA 2. Reguła decyzyjna $\tilde{\delta}$ jest optymalna w zbiorze reguł $\mathcal{D}$ wtedy i tylko wtedy, gdy dla każdego $x \in \mathcal{X}$ zachodzi równość:

$$\begin{aligned} \inf_{\delta \in \mathcal{D}} [\sup_{F \in \Gamma} R_x(F, L_{\delta(x)}) - \inf_{F \in \Gamma} R_x(F, L_{\delta(x)})] \\ = \sup_{F \in \Gamma} R_x(F, L_{\tilde{\delta}(x)}) - \inf_{F \in \Gamma} R_x(F, L_{\tilde{\delta}(x)}). \end{aligned}$$

W statystyce bayesowskiej wykorzystuje się jeszcze jedno kryterium, oparte na porównywaniu kresu górnego ryzyka a posteriori dla różnych reguł decyzyjnych. W zastosowaniach na ogół chodzi nie tylko o to, aby oscylacja była mała, ale również o to, aby ryzyko nie przyjmowało zbyt dużych wartości (por. rozdział 3.2).

DEFINICJA 3. Reguła decyzyjna $\tilde{\delta}$ jest regułą warunkowo Γ -minimaksową przy ustalonej wartości x zmiennej losowej X wtedy i tylko wtedy, gdy zachodzi związek:

$$\inf_{a \in \mathcal{A}} \sup_{F \in \Gamma} R_x(F, L_a) = \sup_{F \in \Gamma} R_x(F, L_{\tilde{\delta}(x)}).$$

Betro i Ruggeri (1992) dowodzą istnienia reguły warunkowo Γ -minimaksowej dla dowolnej rodziny rozkładów a priori Γ na przestrzeni $\Theta \subseteq \mathcal{R}^1$, przy pewnych założeniach dotyczących przestrzeni akcji $\mathcal{A}$ i funkcji straty L . Ich główne wyniki przedstawiają się następująco.

TWIERDZENIE 1 (Betro i Ruggeri (1992)). *Niech Γ będzie daną rodziną rozkładów a priori na przestrzeni $\Theta \subseteq \mathcal{R}^1$, $\Gamma \subseteq \mathcal{P}$. Niech przestrzeń akcji $\mathcal{A}$ będzie przedziałem domkniętym i ograniczonym w $\mathcal{R}^1$ oraz niech funkcja straty $L : \Theta \times \mathcal{A} \rightarrow \mathcal{R}^1$ będzie funkcją wypukłą drugiej zmiennej.*

Wtedy, dla każdego ustalonego $x \in X$, zachodzi:

a) $\sup_{F \in \Gamma} R_x(F, L_a)$ ma przynajmniej jedno minimum w $\mathcal{A}$ i jest ono jedyne, jeżeli funkcja L jest ściśle wypukłą funkcją zmiennej a oraz dla każdego $a \in \mathcal{A}$ zbiór

$$\pi_a = \{F \in \Gamma : R_x(F, L_a) = \sup_{G \in \Gamma} R_x(G, L_a)\}$$

jest niepusty.

b) Jeżeli L jest ściśle wypukłą funkcją zmiennej a i dla pewnego $\bar{a} \in \mathcal{A}$ istnieją rozkłady F_1, F_2 takie, że $F_1, F_2 \in \pi_{\bar{a}}$ oraz

$$a_{Bay}^{F_1} \leq \bar{a} \leq a_{Bay}^{F_2},$$

gdzie a_{Bay}^F oznacza akcję bayesowską przy rozkładzie a priori F , to reguła warunkowo Γ -minimaksowa ma wartość $\bar{\delta}(x) = \bar{a}$.

c) Punkt $\bar{a}$ jest jedynym punktem w zbiorze $\mathcal{A}$ takim, że

$$\forall a < \bar{a} \quad \exists F \in \pi_a \quad a \leq a_{Bay}^F$$

oraz

$$\forall a > \bar{a} \quad \exists F \in \pi_a \quad a \geq a_{Bay}^F \quad \blacksquare$$

Pewne warunki, przy których reguła warunkowo Γ -minimaksowa pokrywa się z regułą optymalną podaje Męczarski (1993).

Przejdźmy teraz do pojęcia odporności infinitesimalnej.

Załóżmy, że na przestrzeni $(\Theta, \mathcal{F}_\Theta)$ mamy określoną rodzinę $\{\Gamma_\varepsilon(F^0) : \varepsilon > 0\}$ klas $\Gamma_\varepsilon(F^0)$ rozkładów a priori, spełniającą dwa następujące warunki:

- 1) $(\forall \varepsilon_1, \varepsilon_2 > 0) \quad \varepsilon_1 < \varepsilon_2 \Rightarrow \Gamma_{\varepsilon_1}(F^0) \subseteq \Gamma_{\varepsilon_2}(F^0);$
- 2) $\bigcap_\varepsilon \Gamma_\varepsilon(F^0) = \{F^0\}.$

Rodzina $\{\Gamma_\varepsilon(F^0) : \varepsilon > 0\}$ reprezentuje lokalne (w punkcie $F^0 \in \mathcal{P}$) zaburzenia modelu bayesowskiego, znikające, gdy $\varepsilon \rightarrow 0$.

DEFINICJA 4. Wielkość a posteriori $R_x(\cdot, L)$ jest infinitesimalnie ρ_x -odporna (ρ_x^0 -odporna) w rodzinie zaburzeń $\{\Gamma_\varepsilon(F^0) : \varepsilon > 0\}$ wtedy i tylko wtedy, gdy

$$\lim_{\varepsilon \rightarrow 0} \rho_x(\Gamma_\varepsilon(F^0), R_x) = 0 \quad (\lim_{\varepsilon \rightarrow 0} \rho_x^0(\Gamma_\varepsilon(F^0), F^0, R_x) = 0).$$

DEFINITION 5. Rozkład a posteriori jest infinitesimalnie ρ_x^ν -odporny (infinitesimalnie odporny w metryce ν) wtedy i tylko wtedy, gdy

$$\lim_{\varepsilon \rightarrow 0} \rho_x^\nu(\Gamma_\varepsilon(F^0), F^0) = 0.$$

Zachodzi następujący lemat.

LEMAT 1.

$$\lim_{\varepsilon \rightarrow 0} \rho_x(\Gamma_\varepsilon(F^0), R_x) = 0 \quad \Leftrightarrow \quad \lim_{\varepsilon \rightarrow 0} \rho_x^0(\Gamma_\varepsilon(F^0), F^0, R_x) = 0.$$

D o w ó d. ($\Rightarrow$) Dla każdego $F \in \Gamma_\varepsilon(F^0)$ zachodzi:

$$\begin{aligned} -\rho_x(\Gamma_\varepsilon, R_x) &\leq \inf_{F \in \Gamma_\varepsilon} R_x(F, L) - R_x(F^0, L) \\ &\leq R_x(F, L) - R_x(F^0, L) \leq \sup_{F \in \Gamma_\varepsilon} R_x(F, L) - R_x(F^0, L) \\ &\leq \rho_x(\Gamma_\varepsilon, R_x), \end{aligned}$$

stąd implikacja „w prawo”.

($\Leftarrow$) Mamy:

$$\begin{aligned} \sup_{F \in \Gamma_\varepsilon} R_x(F, L) - \inf_{F \in \Gamma_\varepsilon} R_x(F, L) \\ = \sup_{F \in \Gamma_\varepsilon} R_x(F, L) - R_x(F^0, L) + R_x(F^0, L) - \inf_{F \in \Gamma_\varepsilon} R_x(F, L) \\ \leq 2 \sup_{F \in \Gamma_\varepsilon} |R_x(F, L) - R_x(F^0, L)|, \end{aligned}$$

stąd implikacja „w lewo”. ■

W pracy będziemy rozważać infinitymalną odporność funkcjonału $R_x(\cdot, L)$ i rozkładu a posteriori dla rodzin $\Gamma_\varepsilon(F^0)$ postaci

$$\Gamma_\varepsilon(F^0) = \{F \in \mathcal{P} : \rho(F, F^0) \leq \varepsilon\},$$

gdzie

$$\rho(F, F^0) = \sup_{\theta \in \Theta} |F(\theta) - F^0(\theta)|.$$

Rodzina $\Gamma_\varepsilon(F^0)$ jest więc ε -otoczeniem rozkładu F^0 w metryce Kołmogorowa.

Rodzina $\Gamma_\varepsilon(F^0)$ jest najmocniejsza z rodzin wymienionych we wstępie, a mianowicie rodziny typu „ ε -zaburzenie”, rodziny rozkładów sprzężonych, rodziny rozkładów z ograniczonymi gęstościami oraz rodziny rozkładów opartej na ilorazie gęstości, w tym sensie, że jeżeli pewna wielkość a posteriori jest infinitymalnie odporna w rodzinie $\Gamma_\varepsilon(F^0)$, to jest infinitymalnie odporna w wymienionych rodzinach. Powiemy o tym dokładniej na początku rozdziału czwartego, gdy wszystkie te rodziny będą już dokładnie zdefiniowane (por. str. 28). Przejdźmy teraz do dokładniejszego omówienia wymienionych rodzin.

3. Analiza odporności funkcjonału $R_x(\cdot, L)$ oraz operatora $\overline{R}_x$ w wybranych rodzinach rozkładów a priori. Przy omawianiu poszczególnych rodzin Γ podamy wyniki dotyczące oscylacji $\rho_x(\Gamma, R_x)$ lub $\rho_x^0(\Gamma, F^0, R_x)$ funkcjonału $R_x(F, L)$. Przedstawimy pewne własności infinitymalne oraz wskażemy przykłady procedur optymalnych. W wyznaczaniu oscylacji funkcjonału $R_x(F, L)$ wykorzystuje się następujące lematy.

LEMAT 2 (Sivaganesan i Berger (1989)). Dany jest zbiór

$$\Phi_I = \{\mu_\tau : \tau \in I\}$$

miar probabilistycznych μ_τ na prostej rzeczywistej $\mathcal{R}^1$, gdzie $I \subseteq \mathcal{R}^k$ jest pewnym zbiorem indeksów. Niech Ψ będzie powłoką wypukłą zbioru Φ_I . Przypuśćmy, że h, g są funkcjami rzeczywistymi na $\mathcal{R}^1$ takimi, że $\int_{\mathcal{R}^1} |h(y)| dF(y) < \infty$ dla każdego $F \in \Phi_I$ oraz istnieje stała B taka, że $B + g(y) > 0$ dla każdego $y \in \mathcal{R}^1$.

Wtedy dla każdego $A \in \mathcal{R}^1$

$$\sup_{F \in \Psi} \frac{A + \int_{\mathcal{R}^1} h(y) dF(y)}{B + \int_{\mathcal{R}^1} g(y) dF(y)} = \sup_{\tau \in I} \frac{A + \int_{\mathcal{R}^1} h(y) \mu_\tau(dy)}{B + \int_{\mathcal{R}^1} g(y) \mu_\tau(dy)}.$$

Teza jest prawdziwa również przy zastąpieniu „sup” przez „inf”.

Dowód. Zauważmy, że jeżeli $F \in \Psi$, to $dF(y) = \int_I \mu_\tau(dy) \gamma(d\tau)$ dla pewnej miary probabilistycznej γ na zbiorze I indeksów τ . Zatem

$$A + \int_{\mathcal{R}^1} h(y) dF(y) = \int_{\mathcal{R}^1} (A + h(y)) \int_I \mu_\tau(dy) \gamma(d\tau).$$

Wykorzystując twierdzenie Fubniego otrzymujemy:

$$\begin{aligned} A + \int_{\mathcal{R}^1} h(y) dF(y) &= \int_I \left(\int_{\mathcal{R}^1} (A + h(y)) \mu_\tau(dy) \right) \gamma(d\tau) \\ &= \int_I \left[\frac{\int_{\mathcal{R}^1} (A + h(y)) \mu_\tau(dy)}{\int_{\mathcal{R}^1} (B + g(y)) \mu_\tau(dy)} \cdot \int_{\mathcal{R}^1} (B + g(y)) \mu_\tau(dy) \right] \gamma(d\tau) \\ &\leq \sup_{\tau \in I} \left(\frac{\int_{\mathcal{R}^1} (A + h(y)) \mu_\tau(dy)}{\int_{\mathcal{R}^1} (B + g(y)) \mu_\tau(dy)} \right) \cdot \left(B + \int_{\mathcal{R}^1} g(y) dF(y) \right). \end{aligned}$$

Zatem

$$\sup_{F \in \Psi} \frac{A + \int_{\mathcal{R}^1} h(y) dF(y)}{B + \int_{\mathcal{R}^1} g(y) dF(y)} \leq \sup_{\tau \in I} \frac{A + \int_{\mathcal{R}^1} h(y) \mu_\tau(dy)}{B + \int_{\mathcal{R}^1} g(y) \mu_\tau(dy)}.$$

Ponieważ $\Psi \supseteq \Phi_I$, więc w powyższej nierówności mamy równość. Dowód dla „inf” jest analogiczny. ■

LEMAT 3. Niech $\mathcal{P}$ będzie rodziną wszystkich miar probabilistycznych na $(\Theta, \mathcal{F}_\Theta)$ oraz niech h, g będą funkcjami rzeczywistymi na Θ takimi, że $g(\theta) > 0$ dla każdego $\theta \in \Theta$. Wtedy

$$\sup_{F \in \mathcal{P}} \frac{\int_{\Theta} h(\theta) F(d\theta)}{\int_{\Theta} g(\theta) F(d\theta)} = \sup_{\theta \in \Theta} \frac{h(\theta)}{g(\theta)}.$$

Teza jest prawdziwa również przy zastąpieniu „sup” przez „inf”.

Dowód. Wystarczy zauważyć, że

$$\frac{\int_{\Theta} h(\theta) F(d\theta)}{\int_{\Theta} g(\theta) F(d\theta)} = \frac{\int_{\Theta} \frac{h(\theta)}{g(\theta)} g(\theta) F(d\theta)}{\int_{\Theta} g(\theta) F(d\theta)} \leq \sup_{\theta \in \Theta} \frac{h(\theta)}{g(\theta)}. \blacksquare$$

Rozważmy problem wyznaczenia $\sup_{F \in \Gamma} R_x(F, L)$, gdzie $\Gamma \subseteq \mathcal{P}$ i założmy, że $\sup_{F \in \Gamma} m_x(F) < \infty$. Wtedy dla każdego ustalonego $q \in \mathcal{R}^1$ zachodzi

$$\sup_{F \in \Gamma} R_x(F, L) < q \Leftrightarrow \sup_{F \in \Gamma} \int_{\Theta} (L(\theta, \delta(x)) - q) l_x(\theta) F(d\theta) < 0.$$

Oznaczmy

$$\bar{g}(q) = \sup_{F \in \Gamma} \int_{\Theta} (L(\theta, \delta(x)) - q) l_x(\theta) F(d\theta).$$

LEMAT 4 (Lemat o linearyzacji. Lavine (1991a,b), Lavine, Wasserman, Wolpert (1991)). *Jeżeli $\int_{\Theta} l_x(\theta) F(d\theta) > 0$ dla każdego $F \in \Gamma$, to funkcja $\bar{g}$ jest nierosnąca i wypukła oraz*

$$\sup_{F \in \Gamma} R_x(F, L) = \inf\{q : \bar{g}(q) \leq 0\}.$$

Dowód. Funkcja $\bar{g}$ jest nierosnąca i wypukła jako supremum funkcji nierosnącej i liniowej. Pokażemy najpierw, że

$$\sup_{F \in \Gamma} R_x(F, L) \leq \inf\{q : \bar{g}(q) \leq 0\}.$$

Warunek ten równoważny jest implikacji:

$$q < \sup_{F \in \Gamma} R_x(F, L) \Rightarrow \bar{g}(q) > 0.$$

Ale jeżeli $q < \sup_{F \in \Gamma} R_x(F, L)$, to istnieje $\tilde{F} \in \Gamma$ takie, że $q < R_x(\tilde{F}, L)$, więc mamy $R_x(\tilde{F}, L) - q > 0$, a stąd $\bar{g}(q) > 0$.

Należy jeszcze pokazać, że

$$\sup_{F \in \Gamma} R_x(F, L) \leq \inf\{q : \bar{g}(q) \leq 0\},$$

co jest równoważne następującej implikacji:

$$q < \inf\{q : \bar{g}(q) \leq 0\} \Rightarrow \sup_{F \in \Gamma} R_x(F, L) \geq q.$$

Ale jeżeli $q < \inf\{q : \bar{g}(q) \leq 0\}$, to zachodzi ciąg implikacji:

$$\begin{aligned} q < \inf\{q : \bar{g}(q) \leq 0\} &\Rightarrow \bar{g}(q) > 0 \Rightarrow \\ \exists \tilde{F} \in \Gamma \int_{\Theta} (L(\theta, \delta(x)) - q) l_x(\theta) \tilde{F}(d\theta) &> 0 \Rightarrow \\ R_x(\tilde{F}, L) > q &\Rightarrow \sup_{F \in \Gamma} R_x(F, L) > q. \blacksquare \end{aligned}$$

Podobnie postępujemy przy wyznaczaniu $\inf_{F \in \Gamma} R_x(F, L)$ i otrzymujemy następującą równość:

$$\inf_{F \in \Gamma} R_x(F, L) = \sup\{q : \underline{g}(q) \geq 0\},$$

gdzie

$$\underline{g}(q) = \inf_{F \in \Gamma} \int_{\Theta} (L(\theta, \delta(x)) - q) l_x(\theta) F(d\theta).$$

Przykłady ilustrujące działanie tego lematu przy różnych klasach Γ rozkładów a priori przedstawione są w pracach Lavine (1991a) oraz Basu i DasGupta (1990). Wykorzystamy go przy badaniu odporności w rodzinach rozkładów z ograniczonymi gęstościami w rozdziale 3.3.

Przed przystąpieniem do omawiania poszczególnych rodzin rozkładów a priori chciałam wspomnieć jeszcze jedno podejście, a mianowicie podejście oparte na rozkładzie a priori typu II ML („Type II maximum likelihood prior”). Przy ustalonej rodzinie rozkładów Γ i wartości obserwacji x polega ono na wyznaczeniu rozkładu $\bar{F} \in \Gamma$ takiego, że

$$m_x(\bar{F}) = \sup_{F \in \Gamma} m_x(F),$$

gdzie, przypomnijmy, $m_x(F) = \int_{\Theta} l_x(\theta) F(d\theta)$. Następnie zastępuje się rodzinę Γ tylko tym jednym rozkładem $\bar{F}$ i przeprowadza analizę interesujących nas wielkości korzystając tylko z rozkładu $\bar{F}$. Otrzymane wyniki przyjmuje się jako optymalne. Więcej nie będę wracać do tego problemu; przedstawiony jest on w pracach Bergera i Berlinera (1986) oraz DasGupta i Studden (1989).

3.1. Rodziny typu „ ε -zaburzenie”. Niech F^0 będzie ustalonym rozkładem a priori. Rodziną typu „ ε -zaburzenie” rozkładu F^0 nazywamy rodzinę rozkładów postaci

$$\Gamma_{\varepsilon}(F^0, \mathcal{Q}) = \{F^Q = (1 - \varepsilon)F^0 + \varepsilon Q : Q \in \mathcal{Q}\},$$

gdzie $\mathcal{Q} \subseteq \mathcal{P}$ jest pewną rodziną rozkładów prawdopodobieństwa na przestrzeni $(\Theta, \mathcal{F}_{\Theta})$ i ε jest ustaloną liczbą z przedziału $[0, 1)$. Będziemy mówili, że rozkład F^Q jest rozkładem F^0 , ε -zaburzonym przez rozkład Q .

Od razu zauważmy, że

$$m_x(F^Q) = \int_{\Theta} l_x(\theta) F^Q(d\theta) = (1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)$$

oraz że rozkład a posteriori F_x^Q jest rozkładem a posteriori F_x^0 , η -zaburzonym przez rozkład a posteriori Q_x , gdzie

$$\eta = \frac{\varepsilon m_x(Q)}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)}.$$

W badaniach odporności bayesowskiej najczęściej rozważa się następujące rodziny $\mathcal{Q}$ (θ_0 oznacza modę rozkładu F^0):

$\mathcal{P}$ (wszystkie możliwe rozkłady prawdopodobieństwa na Θ);

$\mathcal{Q}_u = \{ \text{rozkłady jednomodalne z modą równą } \theta_0 \}$;

$\mathcal{Q}_{su} = \{ \text{rozkłady symetryczne i jednomodalne z modą równą } \theta_0 \}$;

$\mathcal{Q}_v = \{ Q : (1 - \varepsilon)F^0 + \varepsilon Q \text{ jest jednomodalny} \}$.

Pewne inne rodziny $\mathcal{Q}$ możemy znaleźć w pracach Moreno i Cano (1991), Lavine, Wasserman i Wolpert (1991), Bose (1992), Moreno i Pericchi (1993).

Najprościej jest wyznaczyć oscylację wielkości a posteriori gdy $\mathcal{Q} = \mathcal{P}$. Korzystając z lematu 3 otrzymujemy (patrz Berger (1990)), że

$$\sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{P})} R_x(F, L) = \sup_{\theta \in \Theta} \frac{h(\theta)}{g(\theta)}$$

oraz

$$\inf_{F \in \Gamma_\varepsilon(F^0, \mathcal{P})} R_x(F, L) = \inf_{\theta \in \Theta} \frac{h(\theta)}{g(\theta)},$$

gdzie

$$h(\theta) = (1 - \varepsilon) \int_{\Theta} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) + \varepsilon L(\theta, \delta(x)) l_x(\theta)$$

oraz

$$g(\theta) = (1 - \varepsilon) m_x(F^0) + \varepsilon l_x(\theta).$$

Podstawiając $L(\theta, \delta(x)) = 1_C(\theta)$, gdzie $C \in \mathcal{F}_\Theta$, otrzymujemy, jako szczególny przypadek, następujące twierdzenie podane przez Hubera (1973).

Twierdzenie 2. Niech F^0 będzie ustalonym rozkładem a priori określonym na przestrzeni $(\Theta, \mathcal{F}_\Theta)$ oraz niech $C \in \mathcal{F}_\Theta$. Wtedy

$$\sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{P})} F_x\{C\} = \frac{(1 - \varepsilon) m_x(F^0) F_x^0\{C\} + \varepsilon \sup_{\theta \in C} l_x(\theta)}{(1 - \varepsilon) m_x(F^0) + \varepsilon \sup_{\theta \in C} l_x(\theta)}$$

oraz

$$\inf_{F \in \Gamma_\varepsilon(F^0, \mathcal{P})} F_x\{C\} = F_x^0\{C\} \left[1 + \frac{\varepsilon \sup_{\theta \notin C} l_x(\theta)}{(1 - \varepsilon) m_x(F^0)} \right]^{-1}. \blacksquare$$

Sivaganesan (1988) przedstawia oscylację wartości oczekiwanej a posteriori i oscylację wariancji rozkładu a posteriori w rodzinie $\Gamma_\varepsilon(F^0, \mathcal{P})$. Wyniki te są analogiczne do przedstawionych poniżej wyników dotyczących rodziny typu „ ε -zaburzenie” z rodziną rozkładów zaburzających równą $\mathcal{Q}_{su}$, nie będę więc ich tu prezentować.

W przypadku rodziny $\mathcal{Q}_{su}$ do wyznaczania oscylacji wielkości $R_x(\cdot, L)$ wykorzystujemy lemat 2. Przedstawimy ten wynik przy dodatkowym założeniu, że $\Theta = \mathcal{R}^1$ i że rozkłady a priori mają gęstość względem miary Lebesgue’a.

LEMAT 5 (Sivaganesan i Berger (1989)). *Rodzina $\mathcal{Q}_{su}$ jest powłoką wypukłą zbioru Φ_{su} postaci*

$$\Phi_{su} = \{Q^z = U(\theta_0 - z, \theta_0 + z), z > 0\},$$

gdzie $U(a, b)$ oznacza rozkład jednostajny na przedziale (a, b) . ■

Na mocy lematu 2 zadanie sprowadza się do badania oscylacji przy rodzinie rozkładów a priori $\Gamma_\varepsilon(F^0, \Phi_{su})$. Otrzymujemy następujący rezultat.

TWIERDZENIE 3 (Sivaganesan i Berger (1989)). *Niech $\Theta = \mathcal{R}^1$. Wtedy*

$$\begin{aligned} \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q}_{su})} R_x(F, L) &= \sup_{F \in \Gamma_\varepsilon(F^0, \Phi_{su})} R_x(F, L) = \sup_{z > 0} \frac{a_0 + H^L(z)}{a + H_0(z)}, \\ \inf_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q}_{su})} R_x(F, L) &= \inf_{F \in \Gamma_\varepsilon(F^0, \Phi_{su})} R_x(F, L) = \inf_{z > 0} \frac{a_0 + H^L(z)}{a + H_0(z)}, \end{aligned}$$

gdzie

$$\begin{aligned} a &= \frac{(1 - \varepsilon)m_x(F^0)}{\varepsilon}, \\ a_0 &= a \cdot R_x(F^0, L), \\ H_0(z) &= m_x(Q^z) \end{aligned}$$

oraz

$$H^L(z) = \begin{cases} \frac{1}{2z} \int_{\theta_0 - z}^{\theta_0 + z} L(\theta, \delta(x)) l_x(\theta) d\theta, & \text{jeżeli } z \neq 0, \\ L(\theta_0, \delta(x)) l_x(\theta_0), & \text{jeżeli } z = 0. \end{cases}$$

D o w ó d. Dowód wynika z lematu 2 oraz następującej równości:

$$\int_{\Theta} L(\theta, \delta(x)) l_x(\theta) Q(d\theta) = \int_0^\infty H^L(z) d\gamma(z),$$

gdzie γ jest pewnym rozkładem prawdopodobieństwa na rodzinie rozkładów Φ_{su} . Szczegóły dowodu można znaleźć w pracy Sivaganesan i Berger (1989). ■

Analogiczny rezultat otrzymujemy dla rodziny rozkładów $\Gamma_\varepsilon(F^0, \mathcal{Q}_u)$ wykorzystując rodzinę

$$\Phi_u = \{Q^z = U(\theta_0, \theta_0 + z), z > 0\} \cup \{Q^z = U(\theta_0 - z, \theta_0), z > 0\},$$

wstawiając w treści twierdzenia 3 w miejsce $\Gamma_\varepsilon(F^0, \Phi_{su})$ rodzinę $\Gamma_\varepsilon(F^0, \Phi_u)$ oraz definiując

$$H^L(z) = \begin{cases} \frac{1}{z} \int_{\theta_0}^{\theta_0 + z} L(\theta, \delta(x)) l_x(\theta) d\theta, & \text{jeżeli } z \neq 0, \\ L(\theta_0, \delta(x)) l_x(\theta_0), & \text{jeżeli } z = 0. \end{cases}$$

Przejdźmy teraz do badania globalnych miar odporności, to znaczy do badania odwzorowania $\overline{R}_x$, gdy rozkłady a priori przebiegają rodzinę $\Gamma_\varepsilon(F^0, \mathcal{Q})$, $\mathcal{Q} \subseteq \mathcal{P}$. Rozpatrzmy metrykę σ w przestrzeni rozkładów prawdopodobieństwa. Interesuje nas wielkość:

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) = \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q})} \sigma(F_x, F_x^0).$$

Następujące twierdzenie przedstawia górne ograniczenie dla tej wielkości, uogólniając twierdzenie z pracy Fortini i Ruggeri (1990).

TWIERDZENIE 4 (Boratyńska (1992)). *Dla każdego $x \in \mathcal{X}$ i każdego $\varepsilon \in [0, 1)$*

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) \leq \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{m_x(F^0)}{\sup_{Q \in \mathcal{Q}} m_x(Q)} \right]^{-1}.$$

Jeżeli istnieje rozkład $\tilde{Q} \in \mathcal{Q}$ taki, że rozkłady $\tilde{Q}$ i F^0 są wzajemnie osobliwe oraz dla każdego $Q \in \mathcal{Q}$

$$m_x(Q) \leq m_x(\tilde{Q}),$$

to

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) = \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{m_x(F^0)}{\sup_{Q \in \mathcal{Q}} m_x(Q)} \right]^{-1}.$$

Dowód. Przypomnijmy, że

$$\sigma(P, Q) = \sup_{A \in \mathcal{F}_\Theta} |P(A) - Q(A)| = \frac{1}{2} \int_{\Theta} |P(d\theta) - Q(d\theta)|,$$

dla dowolnych rozkładów P i Q . Podstawiając rozkłady a posteriori F_x^Q i F_x^0 odpowiednio w miejsce P i Q , otrzymujemy:

$$\begin{aligned} \rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) &= \sup_{Q \in \mathcal{Q}} \frac{1}{2} \int_{\Theta} |F_x^Q(d\theta) - F_x^0(d\theta)| \\ &= \frac{1}{2} \sup_{Q \in \mathcal{Q}} \left\{ \frac{\varepsilon m_x(Q)}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)} \cdot \int_{\Theta} |Q_x(d\theta) - F_x^0(d\theta)| \right\} \\ &\leq \sup_{Q \in \mathcal{Q}} \frac{\varepsilon m_x(Q)}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)}. \end{aligned}$$

Funkcja $\frac{\varepsilon m_x(Q)}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)}$ jest funkcją rosnącą zmiennej $m_x(Q)$, stąd otrzymujemy nierówność:

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) \leq \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{m_x(F^0)}{\sup_{Q \in \mathcal{Q}} m_x(Q)} \right]^{-1}.$$

Przy założeniach drugiej części twierdzenia mamy:

$$\int_{\Theta} |Q_x(d\theta) - F_x^0(d\theta)| \leq \int_{\Theta} |\tilde{Q}_x(d\theta) - F_x^0(d\theta)| = 2$$

oraz

$$\frac{\varepsilon m_x(Q)}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(Q)} \leq \frac{\varepsilon m_x(\tilde{Q})}{(1 - \varepsilon)m_x(F^0) + \varepsilon m_x(\tilde{Q})},$$

a stąd

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, Q), F^0) = \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{m_x(F^0)}{\sup_{Q \in \mathcal{Q}} m_x(Q)} \right]^{-1}. \blacksquare$$

WNIOSEK 1. Niech $\hat{\theta} \in \Theta$ będzie estymatorem największej wiarygodności parametru θ . Jeżeli $\delta_{\hat{\theta}} \in \mathcal{Q}$, gdzie $\delta_{\hat{\theta}}$ oznacza rozkład jednopunktowy skupiony w punkcie $\hat{\theta}$ oraz $F^0\{\{\hat{\theta}\}\} = 0$, to

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, Q), F^0) = \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{m_x(F^0)}{l_x(\hat{\theta})} \right]^{-1}.$$

W szczególności teza zachodzi gdy $\mathcal{Q} = \mathcal{P}$. $\blacksquare$

Przykład 1. Niech zmienna losowa X ma rozkład normalny $\mathcal{N}(\theta, \tau^2)$, gdzie τ^2 jest znane. Niech rozkład a priori F^0 parametru θ będzie rozkładem normalnym $\mathcal{N}(\theta_0, \tau_0^2)$, gdzie oba parametry są znane. Jeżeli $\mathcal{P}$ jest rodziną wszystkich rozkładów prawdopodobieństwa na $(\Theta, \mathcal{F}_\Theta)$, to dla każdego $x \in \mathcal{R}^1$ otrzymujemy:

$$\rho_x^\sigma(\Gamma_\varepsilon(F^0, \mathcal{P}), F^0) = \left[1 + \frac{1 - \varepsilon}{\varepsilon} \frac{\tau}{(\tau_0^2 + \tau^2)^{1/2}} \exp \left\{ \frac{-(\theta_0 - x)^2}{2(\tau^2 + \tau_0^2)} \right\} \right]^{-1}. \blacksquare$$

Oznaczmy przez $\mathcal{L}$ oraz π metryki Levy'ego i Prochorowa. Ponieważ dla dowolnych rozkładów prawdopodobieństwa P i Q zachodzą następujące nierówności (patrz Zołotariew (1986) i Rachev (1991)):

$$\mathcal{L}(P, Q) \leq \pi(P, Q) \leq \sigma(P, Q),$$

$$\mathcal{L}(P, Q) \leq \rho(P, Q) \leq \sigma(P, Q),$$

więc przedstawione w twierdzeniu 4 ograniczenie górne na $\rho_x^\sigma(\Gamma_\varepsilon(F^0, Q), F^0)$ jest jednocześnie ograniczeniem górnym dla $\rho_x^\mathcal{L}(\Gamma_\varepsilon(F^0, Q), F^0)$, $\rho_x^\pi(\Gamma_\varepsilon(F^0, Q), F^0)$ oraz $\rho_x^\rho(\Gamma_\varepsilon(F^0, Q), F^0)$.

Zauważmy, że to ograniczenie, przy ustalonej wartości x zmiennej losowej X , dąży do 0, gdy ε dąży do 0. Stąd, jako wniosek, otrzymujemy następujące twierdzenie, mówiące o infinitezymalnej odporności rozkładu a posteriori w metrykach σ , $\mathcal{L}$, π oraz ρ w rodzinie otoczeń $\Gamma_\varepsilon(F^0, Q)$, $\varepsilon > 0$.

TWIERDZENIE 5. Dla każdej rodziny rozkładów $\mathcal{Q} \subseteq \mathcal{P}$

$$\lim_{\varepsilon \rightarrow 0} \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q})} \sigma(F_x, F_x^0) = 0,$$

$$\lim_{\varepsilon \rightarrow 0} \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q})} \mathcal{L}(F_x, F_x^0) = 0,$$

$$\lim_{\varepsilon \rightarrow 0} \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q})} \pi(F_x, F_x^0) = 0,$$

$$\lim_{\varepsilon \rightarrow 0} \sup_{F \in \Gamma_\varepsilon(F^0, \mathcal{Q})} \rho(F_x, F_x^0) = 0, \blacksquare$$

Rozważania dotyczące badania oscylacji rozkładów a posteriori w innych metrykach można znaleźć między innymi w pracy Gelfand i Dey (1991).

3.2. Rodziny rozkładów sprzężonych. W wyborze rodziny rozkładów a priori Γ można czasami ograniczyć się do pewnej podrodziny rodziny rozkładów sprzężonych do rodziny $\{P_\theta : \theta \in \Theta\}$. Przypuśćmy, że ta podrodzina rozkładów sprzężonych jest parametryzowana przez $\lambda \in \Lambda$ i oznaczmy ją przez

$$\Gamma_\lambda = \{F^\lambda : \lambda \in \Lambda\}.$$

Przykład 2.

a) Jeżeli $\{P_\theta : \theta > 0\}$ jest rodziną rozkładów wektorowej zmiennej losowej $X = (X_1, X_2, \dots, X_n)$ o niezależnych współrzędnych o jednakowym rozkładzie Poissona $P(\theta)$, to za rodzinę rozkładów a priori można przyjąć na przykład rodzinę

$$\Gamma_\alpha = \{\mathcal{G}(\alpha, \beta_0) : \alpha \in [\alpha_1, \alpha_2]\},$$

gdzie $\alpha_1, \alpha_2, \beta_0$ są ustalonymi liczbami oraz $\mathcal{G}(\alpha, \beta_0)$ oznacza rozkład gamma z parametrami α i β_0 .

b) Jeżeli $\{P_\theta : \theta \in \mathcal{R}^1\}$ jest rodziną rozkładów wektorowej zmiennej losowej $X = (X_1, X_2, \dots, X_n)$ o niezależnych współrzędnych o jednakowym rozkładzie normalnym $\mathcal{N}(\theta, b^2)$, b^2 ustalone, to za rodzinę rozkładów a priori można przyjąć na przykład rodzinę

$$\Gamma_\tau = \{F^\tau : F^\tau = \mathcal{N}(\mu, \tau^2), \mu \text{ jest znane}, \tau^2 \in [\tau_1^2, \tau_2^2]\},$$

gdzie $\mathcal{N}(\mu, \tau^2)$ oznacza rozkład normalny ze średnią μ i wariancją τ^2 oraz liczby τ_1^2, τ_2^2 są ustalone. $\blacksquare$

Wadą tego typu klas rozkładów a priori może być mała różnorodność rozkładów, z drugiej jednak strony praca z rodzinami sprzężonymi jest stosunkowo prosta rachunkowo.

Przedstawię dwie konstrukcje estymatorów optymalnych przy zaburzeniach generowanych przez rodziny sprzężone.

1. Męczarski i Zieliński (1991) wyznaczają estymator optymalny (w sensie definicji 2) parametru θ przy kwadratowej funkcji straty w klasie estymatorów liniowych w modelu z przykładu 2 punktu a. Rodzinę Γ_α uważamy

za rodzinę zaburzeń pewnego ustalonego rozkładu $\mathcal{G}(\alpha_0, \beta_0)$. Okazuje się, że otrzymany estymator jest równy

$$\hat{\theta}_\infty(x) = \frac{\frac{1}{2}(\alpha_1 + \alpha_2 + 1) + S_n}{\beta_0 + n},$$

gdzie $S_n = \sum_{j=1}^n X_j$. Jeżeli $\alpha_1 = \alpha_0 - \delta$ i $\alpha_2 = \alpha_0 + \delta$, to ma on oscylację ryzyka a posteriori rzędu $O(\delta^2)$ przy δ dążącym do 0, podczas gdy estymator bayesowski $\hat{\theta}_{Bay}^{\alpha_0}(x) = E_{\mathcal{G}_x(\alpha_0, \beta_0)}(\theta)$ (również funkcja liniowa obserwacji) ma oscylację rzędu $O(\delta)$. Rząd infinitezymalnej odporności estymatora $\hat{\theta}_\infty$ na małe zaburzenia parametru kształtu rozkładu a priori $\mathcal{G}(\alpha_0, \beta_0)$ jest więc wyższy niż estymatora bayesowskiego $\hat{\theta}_{Bay}^{\alpha_0}$ i zachodzi to dla dowolnego $x \in \mathcal{X}$. Ceną za wyższy rząd zbieżności jest wyższe ryzyko a posteriori estymatora $\hat{\theta}_\infty$ przy rozkładzie a priori $\mathcal{G}(\alpha_0, \beta_0)$ niż estymatora $\hat{\theta}_{Bay}^{\alpha_0}$. Ale zauważmy, że estymator $\hat{\theta}_\infty$ jest zarazem estymatorem warunkowo Γ -minimaksowym w klasie estymatorów liniowych, dla każdego $x \in \mathcal{X}$ spełniony jest bowiem warunek:

$$\inf_{\hat{\theta} \in \mathcal{D}} \sup_{F \in \Gamma_\alpha} R_x(F, L_{\hat{\theta}(x)}) = \sup_{F \in \Gamma_\alpha} R_x(F, L_{\hat{\theta}_\infty(x)}),$$

gdzie $\mathcal{D}$ jest rodziną estymatorów liniowych parametru θ .

Podobny problem rozważa Męczarski (1993). Autor wyznacza najlepszy estymator w klasie wszystkich estymatorów dla parametru $e^{-\theta}$ w tym samym modelu. Estymator ten ma postać:

$$\hat{\theta}_\infty = \frac{1}{2} \cdot \frac{e^{\beta_2 \alpha_1} - e^{\beta_2 \alpha_2}}{e^{\beta_1 \alpha_1} - e^{\beta_1 \alpha_2}} \exp(S_n(\beta_2 - \beta_1)),$$

gdzie $\beta_i = \frac{\beta_0 + n}{\beta_0 + n + i}$, $i = 1, 2$, oraz $S_n = \sum_{j=1}^n X_j$.

2. Rozważmy teraz model z przykładu 2 punktu b. Określoną tu mamy zmienną losową $X = (X_1, X_2, \dots, X_n)$, gdzie X_i , $i = 1, 2, \dots, n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie normalnym $\mathcal{N}(\theta, b^2)$, b^2 jest znaną liczbą i przestrzeń $\Theta = \mathcal{R}^1$.

Chcemy wyznaczyć estymator optymalny $\hat{\theta}_\infty$ parametru θ w klasie wszystkich estymatorów przy kwadratowej funkcji straty, gdy rodzina rozkładów a priori ma postać

$$\Gamma_\tau = \{F^\tau : F^\tau = \mathcal{N}(\mu, \tau^2), \mu \text{ jest znane i } \tau^2 \in [\tau_1^2, \tau_2^2]\}.$$

Rodzinę Γ_τ traktujemy jako otoczenie rozkładu F^{τ_0} , gdzie $\tau_0^2 \in [\tau_1^2, \tau_2^2]$. Otrzymujemy następujące twierdzenie.

Twierdzenie 6 (Boratyńska i Męczarski (1994)). *Estymatorem optymalnym (w sensie definicji 2) parametru θ w klasie wszystkich estymatorów*

jest estymator postaci:

$$\hat{\theta}_{\infty}(x) = \begin{cases} \frac{\lambda_1 + \lambda_2}{2} v_n(x) + \frac{1}{2v_n(x)} + \mu, & \text{gdy } v_n(x) \neq 0, \\ \hat{\theta}(x), & \text{gdy } v_n = 0, \end{cases}$$

gdzie

$$\lambda_i = \left(\frac{n}{b^2} + \frac{1}{\tau_i^2} \right)^{-1}, \quad i = 1, 2,$$

oraz

$$v_n(X) = \frac{\sum_{i=1}^n X_i - n\mu}{b^2}$$

i $\hat{\theta}$ jest dowolnym estymatorem na zbiorze $\{X : v_n(X) = 0\}$. Oscylacja ryzyka R_x estymatora $\hat{\theta}_{\infty}(x)$ wynosi

$$\rho_x(\Gamma_{\tau}, R_x) = \begin{cases} [\frac{1}{2}v_n(x)(\lambda_1 - \lambda_2)]^2, & \text{gdy } v_n(x) \neq 0, \\ \lambda_2 - \lambda_1, & \text{gdy } v_n(x) = 0, \end{cases}$$

oraz

$$\rho_x(\Gamma_{\tau}, R_x) = O(\delta^2) \quad \text{przy } \delta \rightarrow 0,$$

gdzie $\tau_i^2 = \tau_0^2 + (-1)^i \delta$, $i = 1, 2$.

Dowód. Dowód jest przedstawiony w pracy Boratyńska i Męczarski (1994). ■

Zwróćmy uwagę na to, że estymator bayesowski $\hat{\theta}_{Bay}^{\tau_0}(x) = E_{F^{\tau_0}}(\theta)$ ze względu na rozkład a priori F^{τ_0} ma oscylację ryzyka a posteriori w klasie Γ_{τ} rzędu $O(\delta)$, gdy $\tau_i^2 = \tau_0^2 + (-1)^i \delta$, $i = 1, 2$ i $\delta \rightarrow 0$.

Estymator $\hat{\theta}_{\infty}$ jest estymatorem o najmniejszej oscylacji ryzyka a posteriori, ale jednocześnie jego ryzyko a posteriori $R_x(F^{\tau}, L)$ spełnia:

$$\lim_{v_n(x) \rightarrow 0} R_x(F^{\tau}, L_{\hat{\theta}_{\infty}(x)}) = +\infty.$$

Aby uniknąć zbyt dużego wzrostu ryzyka a posteriori proponujemy dwa rodzaje ograniczeń na klasę rozważanych estymatorów. Pierwszym jest ograniczenie się do klasy estymatorów postaci:

$$\mathcal{D}_{\Gamma} = \{\hat{\theta} : \forall x \in \mathcal{X} \quad \exists \tau \in [\tau_1, \tau_2] \quad \hat{\theta}(x) = \hat{\theta}_{Bay}^{\tau}(x)\},$$

gdzie $\hat{\theta}_{Bay}^{\tau}(x)$ jest estymatorem bayesowskim przy rozkładzie a priori F^{τ} . Są to więc takie estymatory, które dla odpowiedniego $\tau \in [\tau_1, \tau_2]$ są „zwykłymi” estymatorami bayesowskimi. Otrzymany estymator optymalny w tej klasie jest zarazem estymatorem warunkowo Γ -minimaksowym w klasie wszystkich estymatorów parametru θ . Drugim rodzajem ograniczenia jest rozważanie tylko estymatorów ε -bayesowskich względem rozkładu a priori F^{τ_0} . Znalezione w każdej z tych klas estymatory są infinitesimalnie odporne z oscylacją

ryzyka a posteriori nie większą niż oscylacja ryzyka estymatora $\hat{\theta}_{Bay}^{\tau_0}$ jednostajnie względem x (dokładne wyniki patrz Boratyńska i Męczarski (1994)).

Wielowymiarowy model normalny ze sprzężonymi rodzinami rozkładów a priori badają Leamer (1982), DasGupta i Studden (1988, 1989), Polasek i Potzelberger (1991).

3.3. Rodziny rozkładów z ograniczonymi gęstościami. Niech U_1, U_2 będą miarami na $(\Theta, \mathcal{F}_\Theta)$ spełniającymi warunek

$$U_1\{\Theta\} \leq 1 \leq U_2\{\Theta\}$$

oraz niech u_1, u_2 będą ich gęstościami względem miary μ dominującej rodzinę rozkładów a priori.

DEFINICJA 6 (Lavine (1991a)). Klasą rozkładów z ograniczonymi gęstościami nazywamy klasę rozkładów prawdopodobieństwa postaci:

$$\Gamma_{DB}^{u_1, u_2} = \{F : \forall \theta \in \Theta \quad u_1(\theta) \leq f(\theta) \leq u_2(\theta) \text{ i } F\{\Theta\} = 1\},$$

gdzie f oznacza gęstość rozkładu F względem miary μ .

Zgodnie z lematem o linearyzacji, aby wyznaczyć oscylację wielkości a posteriori $R_x(F, L)$, wystarczy wyznaczyć funkcje:

$$\begin{aligned} \bar{g}(q) &= \sup_{F \in \Gamma_{DB}^{u_1, u_2}} \int_{\Theta} [L(\theta, \delta(x)) - q] l_x(\theta) F(d\theta), \\ \underline{g}(q) &= \inf_{F \in \Gamma_{DB}^{u_1, u_2}} \int_{\Theta} [L(\theta, \delta(x)) - q] l_x(\theta) F(d\theta). \end{aligned}$$

Funkcje $\bar{g}$ oraz $\underline{g}$ wyznacza się podobnie, więc w dalszej części będziemy zajmować się tylko funkcją $\bar{g}$. Do jej wyznaczenia wykorzystamy następujące twierdzenie.

Niech $Z : \Theta \longrightarrow \mathcal{R}^1$ będzie zmienną losową określoną na przestrzeni $(\Theta, \mathcal{F}_\Theta)$ i niech $\tilde{F}$ będzie jej rozkładem prawdopodobieństwa, gdy rozkład prawdopodobieństwa na Θ jest równy F . Wartość oczekiwaną tej zmiennej losowej oznaczmy przez $E_F Z$, czyli

$$E_F Z = \int_{\mathcal{R}^1} z \tilde{F}(dz).$$

Dla każdego $t \in \mathcal{R}^1$ niech

$$A_t = \{\theta : Z(\theta) \geq t\}$$

oraz

$$B_t = \{\theta : Z(\theta) < t\}.$$

Niech Π_b oznacza miarę na przestrzeni $(\Theta, \mathcal{F}_\Theta)$ taką, że dla dowolnego $C \in \mathcal{F}_\Theta$

$$\Pi_b\{C\} = U_2\{A_b \cap C\} + U_1\{B_b \cap C\},$$

gdzie $b = \inf\{t : U_2\{A_t\} + U_1\{B_t\} \leq 1\}$. Przez π_b oznaczamy gęstość miary Π_b względem miary μ .

Twierdzenie 7 (Lavine (1991a)). *Jeżeli dla każdego $t \in \mathcal{R}^1$ spełniony jest warunek $U_2\{\{\theta : Z(\theta) = t\}\} = 0$, to*

$$\sup_{F \in \Gamma_{DB}^{u_1, u_2}} E_F(Z) = \int_{\Theta} Z(\theta) \Pi_b(d\theta).$$

Dowód. Niech $F \in \Gamma_{DB}^{u_1, u_2}$ i niech

$$A = \{\theta : \pi_b(\theta) \geq f(\theta)\}$$

oraz

$$B = \{\theta : \pi_b(\theta) < f(\theta)\}.$$

Wtedy $A \subseteq A_b$ i $B \subseteq B_b$, więc dla każdego rozkładu $F \in \Gamma_{DB}^{u_1, u_2}$ mamy

$$\begin{aligned} & \int_{\Theta} Z(\theta) \Pi_b(d\theta) - \int_{\Theta} Z(\theta) F(d\theta) \\ &= \int_A Z(\theta) (\Pi_b(d\theta) - F(d\theta)) - \int_B Z(\theta) (F(d\theta) - \Pi_b(d\theta)) \\ &\geq \int_A b(\Pi_b(d\theta) - F(d\theta)) - \int_B b(F(d\theta) - \Pi_b(d\theta)) = 0 \end{aligned}$$

Teraz teza twierdzenia wynika stąd, że $\Pi_b \in \Gamma_{DB}^{u_1, u_2}$. ■

Wykorzystamy teraz to twierdzenie do konstrukcji funkcji $\bar{g}(q)$ oraz $\underline{g}(q)$ dla pewnej szczególnej klasy rozkładów z ograniczonymi gęstościami, a mianowicie dla klasy

$$\Gamma_{DB}^k(F^0) = \left\{ F : \forall \theta \in \Theta \quad \frac{1}{k} f^0(\theta) \leq f(\theta) \leq k f^0(\theta) \right\},$$

gdzie k jest pewną stałą większą od 1, $\Theta \subseteq \mathcal{R}^m$, a rozkład F^0 jest ustalony i ciągły względem miary Lebesgue'a z gęstością równą f^0 . Rodzina $\Gamma_{DB}^k(F^0)$ jest szczególnie użyteczna przy badaniu infinitesimalnej odporności, co sprowadza się do badania interesujących nas wielkości przy $k \rightarrow 1$.

Załóżmy, że rozkład $\tilde{F}^0$ zmiennej losowej Z jest rozkładem ciągłym. Wprowadźmy oznaczenia:

$$\bar{g}_k(Z) = \sup_{F \in \Gamma_{DB}^k(F^0)} \int_{\Theta} Z(\theta) F(d\theta).$$

$$\underline{g}_k(Z) = \inf_{F \in \Gamma_{DB}^k(F^0)} \int_{\Theta} Z(\theta) F(d\theta).$$

Na mocy Twierdzenia 7 otrzymujemy (por. Wasserman i Kadane (1992)), że

$$\bar{g}_k(Z) = \frac{1}{k} \int_{\inf_{\theta \in \Theta} Z(\theta)}^{b_k} z \tilde{F}^0(dz) + k \int_{b_k}^{\sup_{\theta \in \Theta} Z(\theta)} z \tilde{F}^0(dz),$$

gdzie $b_k = (\tilde{F}^0)^{-1}(\frac{k}{k+1})$ i analogicznie

$$\underline{g}_k(Z) = k \int_{\inf_{\theta \in \Theta} Z(\theta)}^{c_k} z \tilde{F}^0(dz) + \frac{1}{k} \int_{c_k}^{\sup_{\theta \in \Theta} Z(\theta)} z \tilde{F}^0(dz),$$

gdzie $c_k = (\tilde{F}^0)^{-1}(\frac{1}{k+1})$.

Dla poszukiwanych funkcji $\bar{g}(q)$ oraz $\underline{g}(q)$ mamy teraz

$$\bar{g}(q) = \bar{g}_k(Z_q) \quad \text{ i } \quad \underline{g}(q) = \underline{g}_k(Z_q),$$

gdzie

$$Z_q(\theta) = [L(\theta, \delta(x)) - q] l_x(\theta).$$

Otrzymujemy wtedy:

$$\begin{aligned} \bar{g}(q) &= \frac{1}{k} \int_{\bar{B}_q} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) + k \int_{\bar{A}_q} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) \\ &\quad - q \left[\frac{1}{k} \int_{\bar{B}_q} l_x(\theta) F^0(d\theta) + k \int_{\bar{A}_q} l_x(\theta) F^0(d\theta) \right] \end{aligned}$$

oraz

$$\begin{aligned} \underline{g}(q) &= k \int_{\underline{B}_q} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) + \frac{1}{k} \int_{\underline{A}_q} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) \\ &\quad - q \left[k \int_{\underline{B}_q} l_x(\theta) F^0(d\theta) + \frac{1}{k} \int_{\underline{A}_q} l_x(\theta) F^0(d\theta) \right], \end{aligned}$$

gdzie

$$\begin{aligned} \bar{A}_q &= \{\theta : [L(\theta, \delta(x)) - q] l_x(\theta) \geq b_k\}, & \bar{B}_q &= \Theta - \bar{A}_q, \\ \underline{A}_q &= \{\theta : [L(\theta, \delta(x)) - q] l_x(\theta) \geq c_k\}, & \underline{B}_q &= \Theta - \underline{A}_q. \end{aligned}$$

Korzystając z lematu o linearyzacji (por. Lemat 4) mamy

$$\sup_{F \in \Gamma_{DB}^k(F^0)} R_x(F, L) = \inf\{q : \bar{g}(q) \leq 0\},$$

oraz

$$\inf_{F \in \Gamma_{DB}^k(F^0)} R_x(F, L) = \sup\{q : \underline{g}(q) \geq 0\}.$$

Wartości q , dla których spełnione są równania $\bar{g}(q) = 0$ oraz $\underline{g}(q) = 0$ mogą być teraz wyznaczone na drodze numerycznej; wyznaczają one, odpowiednio, prawe strony dwóch ostatnich wzorów.

W celu zbadania oscylacji infinityzmalnej zastanówmy się teraz, co dzieje się z wielkością $\bar{g}_k(Z)$, gdy k dąży do 1. Tę lokalną zmienność można wyrazić przez pochodną $\frac{\partial \bar{g}_k(Z)}{\partial k}$ w punkcie $k = 1$: będziemy ją oznaczać przez $DZ(1)$. Zakładając, jak poprzednio, że dystrybuanta $\tilde{F}^0$ jest ciągła, po prostych rachunkach otrzymujemy, że

$$DZ(1) = \int_{m_z}^{\infty} t \tilde{f}^0(t) dt - \int_{-\infty}^{m_z} t \tilde{f}^0(t) dt,$$

gdzie m_z jest medianą zmiennej losowej Z .

Pochodna $DZ(1)$ przedstawia lokalną zmienność wartości oczekiwanej a priori zmiennej losowej Z w rodzinie rozkładów $\Gamma_{DB}^k(F^0)$ przy k dążącym do 1. Powyższe wyniki można wykorzystać do badania oscylacji interesujących nas wielkości a posteriori, jeżeli tylko rodzina rozkładów a posteriori, generowana przez daną rodzinę rozkładów a priori, jest rodziną rozkładów z ograniczonymi gęstościami. Pewien warunek na to jest podany w następującym twierdzeniu Ruggeri i Wasserman (1991).

Niech $\Gamma_{DB}^{k,x}(F^0)$ oznacza rodzinę rozkładów a posteriori odpowiadających rozkładowi a priori z rodziny $\Gamma_{DB}^k(F^0)$. Wtedy

$$\Gamma_{DB}^{\sqrt{k}}(F_x^0) \subseteq \Gamma_{DB}^{k,x}(F^0) \subseteq \Gamma_{DB}^{k^2}(F_x^0).$$

Wprowadźmy oznaczenie:

$$D_x L = \left. \frac{\partial \sup_{F \in \Gamma_{DB}^k(F^0)} R_x(F, L)}{\partial k} \right|_{k=1}.$$

Niech $\tilde{F}_x^0$ będzie rozkładem zmiennej losowej $L(\theta, \delta(x))$ z gęstością $\tilde{f}_x^0$, gdy rozkład prawdopodobieństwa na przestrzeni Θ jest równy F_x^0 z gęstością f_x^0 względem miary Lebesgue'a, to znaczy niech $\tilde{F}_x^0$ będzie takim rozkładem, że

$$\forall A \in \mathcal{B}(\mathcal{R}^1) \quad \tilde{F}_x^0\{A\} = F_x^0\{L^{-1}(A, \delta(x))\}.$$

Niech m oznacza medianę rozkładu $\tilde{F}_x^0$. Zachodzi następujące twierdzenie.

Twierdzenie 8 (Ruggeri i Wasserman (1991)). *Jeżeli dystrybuanta $\tilde{F}_x^0$ jest funkcją ciągłą, to*

$$\begin{aligned} \frac{1}{2} \left\{ \int_m^{\infty} t \tilde{f}_x^0(t) dt - \int_{-\infty}^m t \tilde{f}_x^0(t) dt \right\} &\leq D_x L \\ &\leq 2 \left\{ \int_m^{\infty} t \tilde{f}_x^0(t) dt - \int_{-\infty}^m t \tilde{f}_x^0(t) dt \right\}. \quad \blacksquare \end{aligned}$$

Podobnie możemy znaleźć ograniczenie na pochodną

$$\left. \frac{\partial \inf_{F \in \Gamma_{DB}^k(F^0)} R_x(F, L)}{\partial k} \right|_{k=1}.$$

Wyznaczenie tych pochodnych daje nam wiedzę o szybkości zmiany wielkości a posteriori przy k bliskim 1. Pochodna $D_x L$ może też być wykorzystywana do badania szybkości zmiany supremum ryzyka a posteriori dla wybranej reguły decyzyjnej, na przykład reguły warunkowo Γ -minimaksowej (por. definicja 3), gdy rozpatrujemy coraz mniejsze otoczenia $\Gamma_{DB}^k(F^0)$ rozkładu a priori F^0 . Może ono także służyć do porównywania reguł ze względu na szybkość tej zmiany.

3.4. Rodziny rozkładów oparte na ilorazie gęstości. Rodziny rozkładów oparte na ilorazie gęstości zostały wprowadzone w pracy DeRobertis i Hartigan (1981). Niech l i u będą funkcjami ograniczonymi, nieujemnymi, mierzalnymi, określonymi na przestrzeni Θ i takimi, że $l(\theta) \leq u(\theta)$.

DEFINICJA 7. Rodziną rozkładów opartą na ilorazie gęstości nazywamy rodzinę rozkładów postaci:

$$\Gamma_{DR}^{l,u} = \left\{ F : \forall \theta, \theta' \in \Theta \quad \frac{f(\theta)}{f(\theta')} \leq \frac{u(\theta)}{l(\theta')} \right\}.$$

Zauważmy, że dla ustalonych $\theta_0 \in \Theta$ oraz $F \in \Gamma_{DR}^{l,u}$ mamy

$$\forall \theta \in \Theta \quad \frac{f(\theta_0)}{u(\theta_0)} l(\theta) \leq f(\theta) \leq \frac{f(\theta_0)}{l(\theta_0)} u(\theta).$$

Oznacza to, że F należy do pewnej rodziny rozkładów z ograniczonymi gęstościami. Rzeczywiście obie rodziny są podobne. Związek pomiędzy obiema przedstawia następujący lemat.

LEMAT 6 (Wasserman i Kadane (1992)). *Dla dowolnych, nieujemnych, ograniczonych i mierzalnych funkcji u_1, u_2 określonych na przestrzeni Θ , takich że $u_1(\theta) \leq u_2(\theta)$ i $\int_{\Theta} u_2(\theta) \mu(d\theta) < \infty$, mamy:*

$$\bigcup_{a \leq \alpha \leq b} \Gamma_{DB}^{\alpha u_1, \alpha u_2} = \Gamma_{DR}^{u_1, u_2} \subset \Gamma_{DB}^{\hat{u}_1, \hat{u}_2},$$

gdzie

$$a = \frac{1}{U_2\{\Theta\}}, \quad b = \frac{1}{U_1\{\Theta\}},$$

oraz

$$\hat{u}_1 = \frac{u_1(\theta)}{U_2\{\Theta\}}, \quad \hat{u}_2 = \frac{u_2(\theta)}{U_1\{\Theta\}}.$$

Jeżeli $U_2\{\Theta\} > U_1\{\Theta\}$ i jeżeli u_1, u_2 są funkcjami ciągłymi, to inkluzja jest właściwa.

Dowód. podany jest w pracy Wasserman i Kadane (1992). ■

Ogólne wyniki dotyczące klasy $\Gamma_{DR}^{l,u}$, podobne do wyników przedstawionych dla rodziny $\Gamma_{DB}^{u_1,u_2}$ można znaleźć w pracach Lavine (1991a) i DeRobertis i Hartigan (1981). Nasze rozważania ograniczymy tylko do szczególnej klasy postaci:

$$\Gamma_{DR}^k(F^0) = \left\{ F : \forall \theta, \theta' \in \Theta \quad \frac{f(\theta)}{f(\theta')} \leq k \frac{f^0(\theta)}{f^0(\theta')} \right\}.$$

Zachodzi następujący związek (DeRobertis i Hartigan (1981), Ruggeri i Wasserman (1991)):

$$\sup_{F \in \Gamma_{DR}^k(F^0)} R_x(F, L) = \sup_{F \in \Gamma_{DR}^k(F_x^0)} E_F L$$

oraz

$$\inf_{F \in \Gamma_{DR}^k(F^0)} R_x(F, L) = \inf_{F \in \Gamma_{DR}^k(F_x^0)} E_F L.$$

Stąd wyliczenie oscylacji wielkości $R_x(F, L)$ w rodzinie $\Gamma_{DR}^k(F^0)$ sprowadza się do wyznaczenia oscylacji wartości oczekiwanej $E_F L$ w rodzinie $\Gamma_{DR}^k(F_x^0)$. Została ona wyznaczona w pracy Ruggeri i Wasserman (1991). Tam też przedstawione są pewne własności infinitezymalne, podobne do zaprezentowanych w rozdziale 3.3, i przykłady zastosowań.

4. Analiza odporności rozkładu a posteriori w rodzinie $\Gamma_\varepsilon(F^0)$ rozkładów a priori. Wróćmy teraz do zdefiniowanej w rozdziale drugim rodziny rozkładów będącej ε -otoczeniem w metryce Kolmogorowa rozkładu F^0 , to znaczy do rodziny

$$\Gamma_\varepsilon(F^0) = \{ F : \rho(F, F^0) \leq \varepsilon \}.$$

Pokażemy, że w pewnych modelach statystycznych (reprezentowanych przez funkcję wiarygodności) funkcjonal $R_x(\cdot, L)$ oraz odwzorowanie $\bar{R}_x$ są infinitezymalnie odporne. Przedstawimy też górne ograniczenie na oscylację ρ_x^0 i ρ_x^ρ przy małych ε , gdzie (przypominamy)

$$\begin{aligned} \rho_x^0(\Gamma_\varepsilon(F^0), F^0, R_x) &= \sup_{F \in \Gamma_\varepsilon(F^0)} |R_x(F, L) - R_x(F^0, L)|, \\ \rho_x^\rho(\Gamma_\varepsilon(F^0), F^0) &= \sup_{F \in \Gamma_\varepsilon(F^0)} \rho(F_x, F_x^0) \end{aligned}$$

(por. str. 5). Będziemy zakładać, że przestrzeń Θ jest przedziałem w $\mathcal{R}^1$, a $\mathcal{F}_\Theta$ jest σ -ciałem zbiorów borelowskich. Zakładamy też, że rozkłady a priori F mają gęstość f względem miary Lebesgue'a na przestrzeni $\mathcal{R}^1$, czyli rozważaną rodziną $\mathcal{P}$ będzie teraz rodzina rozkładów prawdopodobieństwa na przedziale Θ , absolutnie ciągłych względem miary Lebesgue'a.

Koncentrujemy się na badaniu oscylacji interesujących nas wielkości na rodzinach $\Gamma_\varepsilon(F^0)$ rozkładów a priori. Otrzymane wyniki dla tych rodzin można też wykorzystać przy badaniu odporności w innych rodzinach rozkładów a priori. Zachodzą następujące lematy pokazujące, że pewne rodziny rozkładów a priori stanowią podrodziny rodziny $\Gamma_\varepsilon(F^0)$.

LEMAT 7. Niech $\Gamma_\varepsilon(F^0, \mathcal{Q}) = \{F^Q = (1 - \varepsilon)F^0 + \varepsilon Q : Q \in \mathcal{Q}\}$ i $\mathcal{Q} \subset \mathcal{P}$. Dla każdego $\varepsilon \in [0, 1]$ zachodzi:

$$\Gamma_\varepsilon(F^0, \mathcal{Q}) \subset \Gamma_\varepsilon(F^0).$$

Dowód. Niech $F^Q = (1 - \varepsilon)F^0 + \varepsilon Q$ i $Q \in \mathcal{Q}$, wtedy

$$\begin{aligned} \rho(F^Q, F^0) &= \sup_{\theta \in \Theta} |(1 - \varepsilon)F^0(\theta) + \varepsilon Q(\theta) - F^0(\theta)| \\ &= \sup_{\theta \in \Theta} \varepsilon |F^0(\theta) - Q(\theta)| \leq \varepsilon. \blacksquare \end{aligned}$$

LEMAT 8. Jeżeli istnieją funkcje h_1, h_2 , będące gęstościami rozkładów prawdopodobieństwa na Θ , takie, że

$$\begin{aligned} u_1(\theta) &\geq f^0(\theta) - \varepsilon h_1(\theta), \\ u_2(\theta) &\leq f^0(\theta) + \varepsilon h_2(\theta), \end{aligned}$$

to

$$\Gamma_{DB}^{u_1, u_2} \subset \Gamma_\varepsilon(F^0).$$

Dowód. Niech $F \in \Gamma_{DB}^{u_1, u_2}$. Wtedy dla dowolnego zbioru $C \in \mathcal{F}_\Theta$ mamy

$$-\varepsilon \int_C h_1(\theta) \mu(d\theta) \leq \int_C (f(\theta) - f^0(\theta)) \mu(d\theta) \leq \varepsilon \int_C h_2(\theta) \mu(d\theta),$$

czyli

$$-\varepsilon \leq \int_C (f(\theta) - f^0(\theta)) \mu(d\theta) \leq \varepsilon,$$

stąd $\rho(F, F^0) \leq \varepsilon$. $\blacksquare$

Zauważmy, że szczególnym przypadkiem rodziny rozkładów z ograniczonymi gęstościami, spełniającej założenia lematu 8, jest rodzina $\Gamma_{DB}^k(F^0)$ przy $k \leq 1 + \varepsilon$.

Korzystając z lematów 6 i 8 otrzymujemy:

WNIOSEK 2. Jeżeli $k \leq 1 + \varepsilon$, to

$$\Gamma_{DR}^k(F^0) \subset \Gamma_\varepsilon(F^0). \blacksquare$$

Podobny związek zachodzi dla pewnych rodzin rozkładów sprzężonych. Przyjmując $\Gamma_\delta(F^{\lambda_0}) = \{F^\lambda : |\lambda - \lambda_0| < \delta\}$ jako otoczenie rozkładu a priori

F^{λ_0} , gdzie F^λ są rozkładami sprzężonymi do rodziny $\{P_\theta : \theta \in \Theta\}$, indeksowanymi parametrem λ , otrzymujemy:

$$\forall \varepsilon > 0 \exists \delta \forall \lambda |\lambda - \lambda_0| < \delta \Rightarrow \rho(F^\lambda, F^{\lambda_0}) < \varepsilon,$$

czyli $\Gamma_\delta(F^{\lambda_0}) \subseteq \Gamma_\varepsilon(F^{\lambda_0})$.

Przykład 3. Niech $F^{\tau_0} = \mathcal{N}(\mu, \tau_0^2)$, μ, τ_0 ustalone oraz

$$\Gamma_\delta(F^{\tau_0}) = \{F^\tau : F^\tau = \mathcal{N}(\mu, \tau^2) \text{ i } |\tau^2 - \tau_0^2| < \delta\},$$

wtedy $\Gamma_\delta(F^{\tau_0}) \subseteq \Gamma_\varepsilon(F^{\tau_0})$ dla odpowiedniego ε oraz $\delta \rightarrow 0 \Leftrightarrow \varepsilon \rightarrow 0$. ■

Pewne wnioski odnośnie do infinitesimalnej odporności przy innych „standardowych” metrykach sformułowaliśmy już w rozdziale 3.1.

4.1. Ciągłość odwzorowań $R_x(\cdot, L)$ i $\overline{R}_x$. Na początek podamy dwa lematy wykorzystywane w dowodach twierdzeń dotyczących ciągłości odwzorowań $R_x(\cdot, L)$ i $\overline{R}_x$.

LEMAT 9. Niech $g : \Theta \rightarrow \mathcal{R}^1$ będzie funkcją ciągłą i ograniczoną. Wtedy dla każdego rozkładu prawdopodobieństwa F^0 i dla każdego $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że dla każdego rozkładu prawdopodobieństwa F na Θ

$$\rho(F, F^0) < \delta \Rightarrow \left| \int_{\Theta} g(\theta) F(d\theta) - \int_{\Theta} g(\theta) F^0(d\theta) \right| < \varepsilon.$$

Dowód. Dla dowolnych rozkładów prawdopodobieństwa F i G zachodzi:

$$\mathcal{L}(F, G) \leq \rho(F, G),$$

gdzie $\mathcal{L}$ jest metryką Levy’ego. Metryka Levy’ego $\mathcal{L}$ metryzuje słabą topologię na przestrzeni rozkładów prawdopodobieństwa na prostej rzeczywistej, co daje tezę. ■

LEMAT 10. Niech $g : \Theta \rightarrow \mathcal{R}^1$ będzie funkcją o wahanii skończonym. Wtedy dla każdego rozkładu prawdopodobieństwa F^0 , absolutnie ciągłego względem miary Lebesgue’a, i dla każdego $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że dla każdego rozkładu F absolutnie ciągłego względem miary Lebesgue’a

$$\rho(F, F^0) < \delta \Rightarrow \left| \int_{\Theta} g(\theta) F(d\theta) - \int_{\Theta} g(\theta) F^0(d\theta) \right| < \varepsilon.$$

Dowód. Całkując przez części otrzymujemy:

$$\left| \int_{\Theta} g(\theta) (F(d\theta) - F^0(d\theta)) \right| \leq \rho(F, F^0) \cdot V(g),$$

gdzie $V(g)$ oznacza całkowite wahanie funkcji g . Stąd już natychmiast otrzymujemy tezę. ■

Twierdzenie 9 (infinitesimalna odporność funkcjonału $R_x(\cdot, L)$). Niech $x \in \mathcal{X}$ będzie ustalonym punktem przestrzeni prób. Niech funkcje $l_x(\cdot)$ oraz

$L(\cdot, \delta(x)) \cdot l_x(\cdot)$ będą funkcjami ciągłymi i ograniczonymi lub funkcjami o wahanii skończonym.

Wtedy dla każdego rozkładu a priori $F^0 \in \mathcal{P}$ i dla każdego $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że dla każdego rozkładu a priori $F \in \mathcal{P}$

$$\rho(F, F^0) < \delta \Rightarrow |R_x(F, L) - R_x(F^0, L)| < \varepsilon,$$

o ile $R_x(F^0, L) < \infty$.

D o w ó d. Mamy:

$$\begin{aligned} & |R_x(F, L) - R_x(F^0, L)| \\ &= \left| \frac{\int_{\Theta} L(\theta, \delta(x)) l_x(\theta) F(d\theta)}{m_x(F)} - \frac{\int_{\Theta} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta)}{m_x(F^0)} \right| \\ &= \frac{1}{m_x(F) \cdot m_x(F^0)} \left| \int_{\Theta} L(\theta, \delta(x)) l_x(\theta) (F(d\theta) - F^0(d\theta)) m_x(F^0) \right. \\ &\quad \left. + \int_{\Theta} L(\theta, \delta(x)) l_x(\theta) F^0(d\theta) [m_x(F^0) - m_x(F)] \right| \\ &\leq \frac{(R_x(F^0, L) |m_x(F^0) - m_x(F)| + |\int_{\Theta} L(\theta, \delta(x)) l_x(\theta) (F(d\theta) - F^0(d\theta))|)}{m_x(F)}. \end{aligned}$$

Korzystając z lematu 9 i 10 przy funkcji g równej $l_x(\cdot)$ lub równej $L(\cdot, \delta(x)) \cdot l_x(\cdot)$ oraz z implikacji

$$|m_x(F) - m_x(F^0)| < \varepsilon_1 \Rightarrow m_x(F) > m_x(F^0) - \varepsilon_1$$

otrzymujemy tezę. ■

Przejdźmy teraz do badania ciągłości odwzorowania $\overline{R}_x: \mathcal{P} \rightarrow \mathcal{P}$, określonego wzorem $\overline{R}_x(F) = F_x$ (por. str.5). Podobnie jak twierdzenie 9 dowodzimy następujące twierdzenie.

Twierdzenie 10 (infinitezymalna odporność rozkładu a posteriori w metryce Kołmogorowa). Niech $x \in \mathcal{X}$ będzie ustalonym punktem przestrzeni prób. Jeżeli funkcja $l_x(\cdot)$ jest funkcją ciągłą i ograniczoną lub funkcją o wahanii skończonym, to dla każdego rozkładu a priori $F^0 \in \mathcal{P}$ i dla każdego $\varepsilon > 0$ istnieje liczba $\delta > 0$ taka, że dla każdego rozkładu a priori $F \in \mathcal{P}$

$$\rho(F, F^0) < \delta \Rightarrow \rho(F_x, F_x^0) < \varepsilon. \quad \blacksquare$$

Bezpośrednio z powyższych twierdzeń oraz z lematów 7, 8 i wniosku 2 otrzymujemy następujący wniosek mówiący o infinitezymalnej odporności funkcjonału $R_x(\cdot, L)$ i rozkładu a posteriori w metryce Kołmogorowa w innych rodzinach rozkładów a priori.

WNIOSEK 3. Niech funkcje $l_x(\cdot)$ i $L(\cdot, \delta(x))l_x(\cdot)$ spełniają założenia twierdzenia 9. Wtedy zachodzą następujące równości:

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \rho_x(\Gamma_\varepsilon(F^0, \mathcal{Q}), R_x) &= 0, & \lim_{\varepsilon \rightarrow 0} \rho_x^\rho(\Gamma_\varepsilon(F^0, \mathcal{Q}), F^0) &= 0, \\ \lim_{k \rightarrow 1} \rho_x(\Gamma_{DB}^k(F^0), R_x) &= 0, & \lim_{k \rightarrow 1} \rho_x^\rho(\Gamma_{DB}^k(F^0), F^0) &= 0, \\ \lim_{k \rightarrow 1} \rho_x(\Gamma_{DR}^k(F^0), R_x) &= 0, & \lim_{k \rightarrow 1} \rho_x^\rho(\Gamma_{DR}^k(F^0), F^0) &= 0. \blacksquare \end{aligned}$$

Przykład 4. Niech zmienna losowa X ma rozkład normalny $\mathcal{N}(\theta, b^2)$, b^2 jest ustaloną liczbą, $\theta \in \Theta = \mathcal{R}^1$. Niech rodzina rozkładów a priori ma postać

$$\Gamma_\varepsilon(F^0, \mathcal{Q}) = \{(1 - \varepsilon)F^0 + \varepsilon Q : Q \in \mathcal{Q}\},$$

gdzie F^0 jest rozkładem normalnym z ustalonymi parametrami, a $\mathcal{Q}$ jest pewną podrodziną rodziny rozkładów ciągłych na Θ . Niech $L(\theta, \delta(x)) = w(\theta, \delta(x))$, gdzie w jest wielomianem zmiennej θ . Wtedy dla każdego $x \in \mathcal{X}$

$$\lim_{\varepsilon \rightarrow 0} \rho_x(\Gamma_\varepsilon(F^0, \mathcal{Q}), R_x) = 0. \blacksquare$$

Przykład 5. Załóżmy, że zmienna losowa X ma rozkład prawdopodobieństwa jak w przykładzie 4 i rodzina rozkładów a priori jest postaci:

$$I_\tau^\delta = \{F^\tau : F^\tau = \mathcal{N}(\mu, \tau^2), \mu \text{ jest ustalone}, \tau \in [\tau_0 - \delta, \tau_0 + \delta]\},$$

gdzie τ_0 jest ustalone. Niech $L(\theta, \delta(x)) = w(|\theta - \delta(x)|)$, gdzie w jest wielomianem. Wówczas przy funkcji straty $L(\theta, \delta(x))$ każdy estymator jest infinitesimalnie odporny ze względu na ryzyko a posteriori. $\blacksquare$

W szczególności w przykładach 4 i 5 możemy podstawić $L(\theta, \delta(x)) = |\theta - \delta(x)|^2$ i otrzymujemy, że w przypadku kwadratowej funkcji straty każdy estymator jest infinitesimalnie odporny. Odporny jest więc na przykład estymator bayesowski dla rozkładu a priori F^0 i estymator ε -bayesowski.

4.2. Różniczkowalność odwzorowań $R_x(\cdot, L)$ i $\overline{R}_x$. Rozpatrując odporność infinitesimalną procedury statystycznej naturalne wydaje się rozważenie pochodnej jako miary wrażliwości procedury na małe zaburzenia badanego modelu. Pochodną Frecheta operatora przyporządkowującego rozkładowi a priori jego rozkład a posteriori, gdy w przestrzeni miar $\mathcal{P}$ określona jest metryka σ , wyznaczają Diaconis i Freedman (1986).

Wyznamy pochodne odwzorowań $R_x(\cdot, L) : \mathcal{P} \rightarrow \mathcal{R}^1$ i $\overline{R}_x : \mathcal{P} \rightarrow \mathcal{P}$ przy odległości w przestrzeni $\mathcal{P}$ zdefiniowanej przez metrykę Kołmogorowa i przy ustalonych funkcji wiarygodności $l_x(\cdot)$ oraz funkcji $L(\cdot, \delta(x))$. Obliczenie pochodnej funkcjonału $R_x(\cdot, L_{\delta(x)})$ dla każdego $x \in \mathcal{X}$ umożliwia wyodrębnienie zbioru tych wartości zmiennej losowej X , dla których małe zmiany w rozkładzie a priori powodują duże zmiany badanych wielkości a

posteriori. Obliczenie pochodnych funkcjonału $R_x(\cdot, L_{\delta(x)})$ dla różnych procedur statystycznych δ umożliwia ich porównywanie ze względu na odporność infitezymalną.

Niech $\mathcal{G}$ oznacza przestrzeń skończonych przeliczalnie addytywnych funkcji Δ , określonych na σ -ciele $\mathcal{F}_\Theta$, o wartościach w $\mathcal{R}^1$, absolutnie ciągłych względem miary Lebesgue'a i takich, że $\Delta\{\Theta\} = 0$. W przestrzeni $\mathcal{G}$ definiujemy normę w następujący sposób:

$$\|\Delta\|_\infty = \sup_{\theta \in \Theta} |\Delta(\theta)|,$$

gdzie $\Delta(\theta) = \Delta\{(-\infty, \theta)\}$. Przestrzeń $\mathcal{G}$ jest unormowaną przestrzenią liniową nad $\mathcal{R}^1$.

Zauważmy, że jeżeli $F_1, F_2 \in \mathcal{P}$, to $F_1 - F_2 \in \mathcal{G}$ i

$$\|F_1 - F_2\|_\infty = \rho(F_1, F_2).$$

Będziemy korzystali z następującej definicji pochodnej (patrz na przykład Clarke (1983), Serfling (1991)) odwzorowania o wartościach w przestrzeni liniowej unormowanej.

DEFINICJA 8. Mówimy, że operator $T : \mathcal{P} \rightarrow \mathcal{Y}$ jest różniczkowalny w punkcie $F^0 \in \mathcal{P}$, w odniesieniu do metryki ρ w przestrzeni $\mathcal{P}$ wtedy i tylko wtedy, gdy istnieje ciągle operator liniowy $T'(F^0) : \mathcal{G} \rightarrow \mathcal{Y}$ taki, że

$$\frac{\|T(F) - T(F^0) - T'(F^0)(F - F^0)\|}{\rho(F, F^0)} \rightarrow 0,$$

gdy $\rho(F, F^0)$ dąży do 0.

Niech $\mathcal{V}$ oznacza klasę funkcji rzeczywistych na $\mathcal{R}^1$ o wahanu skończonym i dla $h \in \mathcal{V}$ oznaczmy przez $V(h)$ wahanie całkowite funkcji h . Niech $h(\cdot) = h^+(\cdot) - h^-(\cdot)$ będzie rozkładem Jordana funkcji h i niech $h^*(\cdot) = h^+(\cdot) + h^-(\cdot)$. Wtedy

$$V(h) = \int_{\mathcal{R}^1} dh^*(x).$$

Zauważmy, że w przypadku, gdy h jest funkcją różniczkowalną i

$$\int_{\mathcal{R}^1} \left| \frac{dh(x)}{dx} \right| dx < \infty$$

mamy

$$V(h) = \int_{\mathcal{R}^1} \left| \frac{dh(x)}{dx} \right| dx.$$

Wszystkie całki poniżej będą całkami Lebesgue'a-Stieltjesa na prostej. Dla ustalonych x, δ zdefiniujemy odwzorowanie $z : \Theta \rightarrow \mathcal{R}^1$ wzorem

$$z(\theta) = [R_x(F^0, L) - L(\theta, \delta(x))]l_x(\theta).$$

Dla ustalonych $l_x(\cdot) \in \mathcal{V}$ oraz $L(\cdot, \delta(x)) \in \mathcal{V}$ rozważmy funkcjonal $R_x(\cdot, L)$ jako odwzorowanie z przestrzeni $\mathcal{P}$ w $\mathcal{R}^1$ postaci $F \mapsto R_x(F, L)$. Zachodzi następujące twierdzenie.

Twierdzenie 11. *Jeżeli funkcje $l_x(\cdot)$ oraz $l_x(\cdot)L(\cdot, \delta(x))$ są funkcjami o wahanu skończonym, to dla każdego rozkładu $F^0 \in \mathcal{P}$ funkcjonal $R_x(\cdot, L) : \mathcal{P} \rightarrow \mathcal{R}^1$ jest różniczkowalny. Pochodna tego funkcjonału wyraża się wzorem*

$$R'_x(F^0, L)\Delta = \frac{\int_{\Theta} L(\theta, \delta(x))l_x(\theta)\Delta(d\theta) - R_x(F^0, L)m_x(\Delta)}{m_x(F^0)}, \quad \Delta \in \mathcal{G}.$$

Dowód. Mamy

$$\begin{aligned} R_x(F, L) - R_x(F^0, L) &= \frac{\int_{\Theta} L(\theta, \delta(x))l_x(\theta)F(d\theta)}{m_x(F)} - \frac{\int_{\Theta} L(\theta, \delta(x))l_x(\theta)F^0(d\theta)}{m_x(F^0)} \\ &= \frac{\int_{\Theta} L(\theta, \delta(x))l_x(\theta)(F - F^0)(d\theta) - R_x(F^0, L)m_x(F - F^0)}{m_x(F^0)} \\ &\quad + r_x(F^0, F - F^0, L), \end{aligned}$$

gdzie

$$\begin{aligned} r_x(F^0, F - F^0, L) &= \frac{m_x(F - F^0)}{m_x(F^0) + m_x(F - F^0)} \\ &\quad \cdot \frac{R_x(F^0, L)m_x(F - F^0) - \int_{\Theta} L(\theta, \delta(x))l_x(\theta)(F - F^0)(d\theta)}{m_x(F^0)}. \end{aligned}$$

Oznaczmy $F - F^0 = \bar{\Delta} \in \mathcal{G}$. Pokażemy, że

$$|r_x(F^0, \bar{\Delta}, L)| = o(\|\bar{\Delta}\|_{\infty}).$$

Całkując przez części otrzymujemy:

$$|m_x(\bar{\Delta})| \leq \|\bar{\Delta}\|_{\infty} \cdot V(l_x)$$

i stąd

$$(1) \quad m_x(\bar{\Delta}) \rightarrow 0 \quad \text{przy} \quad \|\bar{\Delta}\|_{\infty} \rightarrow 0,$$

oraz

$$(2) \quad \left| \int_{\Theta} [R_x(F^0, L)l_x(\theta) - L(\theta, \delta(x))l_x(\theta)]\bar{\Delta}(d\theta) \right| \leq \|\bar{\Delta}\|_{\infty} \cdot V(z),$$

gdzie $z(\theta) = [R_x(F^0, L) - L(\theta, \delta(x))]l_x(\theta)$ jest funkcją o wahanu skończonym na mocy założeń.

Korzystając z (2) otrzymujemy

$$|r_x(F^0, \bar{\Delta}, L)| \leq \left| \frac{m_x(\bar{\Delta})}{m_x(F^0) + m_x(\bar{\Delta})} \right| \cdot \|\bar{\Delta}\|_{\infty} \cdot \frac{V(z)}{m_x(F^0)},$$

a z (1) mamy

$$\left| \frac{m_x(\bar{\Delta})}{m_x(F^0) + m_x(\bar{\Delta})} \right| \rightarrow 0 \quad \text{przy} \quad \|\bar{\Delta}\|_\infty \rightarrow 0.$$

Należy jeszcze pokazać, że $R'_x(F^0, L)$ jest ciągłym funkcjonałem liniowym określonym na przestrzeni $\mathcal{G}$. Policzmy $|R'_x(F^0, L)\Delta|$. Całkując przez części otrzymujemy:

$$\begin{aligned} |R'_x(F^0, L)\Delta| &= \frac{1}{m_x(F^0)} \left| \int_{\Theta} (L(\theta, \delta(x))l_x(\theta) - R_x(F^0, L)l_x(\theta))\Delta(d\theta) \right| \\ &\leq \|\Delta\|_\infty \cdot \frac{V(z)}{m_x(F^0)}, \end{aligned}$$

co dowodzi ograniczoności (a więc i ciągłości) funkcjonału $R'_x(F^0, L)$. ■

Ostatnia nierówność w dowodzie twierdzenia 11 przedstawia górne ograniczenie na normę pochodnej $R'_x(F^0, L)$. Okazuje się, że prawdziwe jest następujące twierdzenie.

TWIERDZENIE 12. *Przy założeniach twierdzenia 11 norma pochodnej $R'_x(F^0, L)$ wynosi*

$$\|R'_x(F^0, L)\| = \frac{V_{\text{int}\Theta}(z)}{m_x(F^0)},$$

gdzie $V_{\text{int}\Theta}(z)$ jest wahaniem funkcji z po wnętrzu zbioru Θ .

D o w ó d. Oznaczmy końce przedziału Θ literami a i b (a, b mogą przyjmować wartości $-\infty, +\infty$). Niech $\{a_n\}_{n=1}^\infty$ i $\{b_n\}_{n=1}^\infty$ będą ciągami takimi, że

- 1) $\forall n \in \mathcal{N} \quad a_n, b_n \in \Theta, \quad \lim_{n \rightarrow \infty} a_n = a, \quad \lim_{n \rightarrow \infty} b_n = b,$
- 2) funkcja z jest ciągła w punktach a_n, b_n

oraz

- 3) $V_{\{a_n, b_n\}}(z) \rightarrow V_{(a, b)}(z),$

gdzie $V_A(z)$ oznacza całkowite wahanie funkcji z na zbiorze A .

Wtedy

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathcal{N} \quad \forall n > n_0 \quad V_{(a, b)}(z) < V_{[a_n, b_n]}(z) + \varepsilon.$$

Pokażemy teraz, że $\|R'_x(F^0, L)\| \leq \frac{V_{(a, b)}(z)}{m_x(F^0)}$. Całkując przez części otrzymujemy

$$|R'_x(F^0, L)\Delta| = \left| \frac{\int_{\Theta} \Delta(\theta) dz(\theta)}{m_x(F^0)} \right| \leq \frac{V_{(a, b)}(z)}{m_x(F^0)} \cdot \|\Delta\|_\infty$$

(funkcja Δ spełnia warunek $\Delta(a) = \Delta(b) = 0$).

Aby pokazać równość rozpatrzmy funkcjonal $H_n : \mathcal{C}[a_n, b_n] \rightarrow \mathcal{R}^1$ postaci

$$H_n(\Delta) = \frac{\int_{a_n}^{b_n} \Delta(\theta) dz(\theta)}{m_x(F^0)},$$

gdzie $\mathcal{C}[a_n, b_n]$ oznacza przestrzeń funkcji ciągłych na zbiorze $[a_n, b_n]$. Z twierdzenia Riesz o reprezentacji funkcjonałów liniowych mamy

$$\|H_n\| = \frac{V_{[a_n, b_n]}(z)}{m_x(F^0)},$$

stąd dla każdego $\varepsilon > 0$ istnieje funkcja $\Delta_n \in \mathcal{C}[a_n, b_n]$ taka, że

$$\|\Delta_n\|_\infty = \sup_{\theta \in [a_n, b_n]} |\Delta_n| = 1$$

oraz

$$|H_n(\Delta_n)| > \frac{V_{[a_n, b_n]}(z)}{m_x(F^0)} - \varepsilon.$$

Funkcję Δ_n możemy przedłużyć do funkcji ciągłej $\tilde{\Delta}_n$ określonej na zbiorze $\mathcal{O}$ takiej, że $\tilde{\Delta}_n(a) = \tilde{\Delta}_n(b) = 0$ i $\|\tilde{\Delta}_n\|_\infty = 1$ ($\tilde{\Delta}_n$ jako funkcja zbiorów należy do $\mathcal{G}$). Wtedy

$$\begin{aligned} |R'_x(F^0, L)\tilde{\Delta}_n| &= \frac{|H_n(\Delta_n) + \int_a^{a_n} \tilde{\Delta}_n(\theta) dz(\theta) + \int_{b_n}^b \tilde{\Delta}_n(\theta) dz(\theta)|}{m_x(F^0)} \\ &\geq \frac{V_{[a_n, b_n]}(z)}{m_x(F^0)} - \varepsilon - \tilde{\Delta}_n(V_{(a, a_n)}(z) + V_{(b_n, b)}(z)) \geq \frac{V_{(a, b)}(z)}{m_x(F^0)} - 3\varepsilon, \end{aligned}$$

co daje tezę. ■

WNIOSEK 4. Jeżeli $l_x(\cdot)$ oraz $L(\cdot, \delta(x))$ są różniczkowalne, to

$$\|R'_x(F^0, L)\| = \frac{\int_{\mathcal{O}} \left| \frac{\partial}{\partial \theta} [(R_x(F^0, L) - L(\theta, \delta(x)))l_x(\theta)] \right| d\theta}{m_x(F^0)}.$$

Dowód. Weźmy ciąg przeliczalnie addytywnych funkcji zbiorów $\{\Delta_n\}_{n=1}^\infty$ takich, że $\|\Delta\|_\infty \leq 1$, zbieżny do funkcji Δ spełniającej warunki:

$$\Delta(\theta) = \begin{cases} -1, & \text{gdy } \theta \in \left\{ \theta : \frac{\partial}{\partial \theta} [(R_x(F^0, L) - L(\theta, \delta(x)))l_x(\theta)] < 0 \right\}, \\ 1, & \text{gdy } \theta \in \left\{ \theta : \frac{\partial}{\partial \theta} [(R_x(F^0, L) - L(\theta, \delta(x)))l_x(\theta)] \geq 0 \right\}. \end{cases}$$

Wtedy

$$|R'_x(F^0, L)\Delta_n| \rightarrow \frac{\int_{\mathcal{O}} \left| \frac{\partial}{\partial \theta} [(R_x(F^0, L) - L(\theta, \delta(x)))l_x(\theta)] \right| d\theta}{m_x(F^0)}$$

przy $n \rightarrow \infty$ ■

Podobnie wyznaczamy pochodną operatora $\overline{R}_x : \mathcal{P} \rightarrow \mathcal{P}$, określonego wzorem $\overline{R}_x(F) = F_x$. Otrzymujemy następujące twierdzenie.

Twierdzenie 13. *Jeżeli funkcja $l_x(\cdot)$ jest funkcją o wahanii skończonym, to operator $\overline{R}_x$ jest różniczkowalny w każdym punkcie $F^0 \in \mathcal{P}$ i pochodna $\overline{R}_x'(F^0) : \mathcal{G} \rightarrow \mathcal{G}$ ma postać*

$$\overline{R}_x'(F^0)\Delta(d\theta) = \frac{1}{m_x(F^0)}[l_x(\theta)\Delta(d\theta) - m_x(\Delta)F_x^0(d\theta)].$$

Dowód. Oznaczmy $\overline{\Delta} = F - F^0$ i

$$\overline{r}_x(F^0, \overline{\Delta}) = \overline{R}_x(F) - \overline{R}_x(F^0) - \overline{R}_x'(F^0)\overline{\Delta},$$

wtedy

$$\overline{r}_x(F^0, \overline{\Delta})(d\theta) = \frac{m_x(\overline{\Delta})}{m_x(F^0) + m_x(\overline{\Delta})} \cdot \frac{m_x(\overline{\Delta})F_x^0(d\theta) - l_x(\theta)\overline{\Delta}(d\theta)}{m_x(F^0)}.$$

Rozumując jak w dowodzie twierdzenia 11 otrzymujemy

$$\begin{aligned} \|\overline{r}_x(F^0, \overline{\Delta})\|_\infty &= \sup_{\theta \in \Theta} |\overline{r}_x(F^0, \overline{\Delta})(\theta)| \\ &\leq \left| \frac{m_x(\overline{\Delta})}{m_x(F^0) + m_x(\overline{\Delta})} \right| \cdot \frac{\|\overline{\Delta}\|_\infty}{m_x(F^0)} \cdot \left\{ \sup_{\theta \in \Theta} |l_x(\theta)| + V(l_x) \right\}, \end{aligned}$$

stąd

$$\|\overline{r}_x(F^0, \overline{\Delta})\|_\infty = o(\|\overline{\Delta}\|_\infty),$$

oraz

$$\|\overline{R}_x'(F^0)\Delta\|_\infty \leq \frac{\|\Delta\|_\infty}{m_x(F^0)} \cdot \left\{ \sup_{\theta \in \Theta} |l_x(\theta)| + V(l_x) \right\},$$

więc $\overline{R}_x'(F^0)$ jest operatorem ciągłym. ■

Wynika stąd, że jeżeli pochodne operatorów $\overline{R}_x$ i $R_x(\cdot, L)$ nie znikają, to przy ε dążącym do zera mamy

$$\rho_x^0(\Gamma_\varepsilon(F^0), F^0, R_x) = O(\varepsilon)$$

oraz

$$\rho_x^\rho(\Gamma_\varepsilon(F^0), F^0) = O(\varepsilon),$$

co wyznacza rząd infinitezymalnej odporności równy $O(\varepsilon)$.

4.3. Górne ograniczenia wielkości ρ_x^0 i ρ_x^ρ . Poniższe twierdzenia szacują różnicę między badanymi wielkościami a posteriori dla rozkładów F_x i F_x^0 w terminach odległości Kołmogorowa między rozkładami F i F^0 . Niech x będzie danym punktem przestrzeni prób.

Twierdzenie 14. Jeżeli funkcje $l_x(\cdot)$ oraz $l_x(\cdot)L(\cdot, \delta(x))$ są funkcjami o wahanii skończonym oraz $\rho(F, F^0) < \frac{m_x(F^0)}{V(l_x)}$, to

$$|R_x(F^0, L) - R_x(F, L)| \leq \frac{\rho(F, F^0)}{m_x(F^0) - \rho(F, F^0) \cdot V(l_x)} \cdot V(z),$$

gdzie

$$z(\theta) = [R_x(F^0, L) - L(\theta, \delta(x))]l_x(\theta).$$

Dowód. Mamy:

$$\begin{aligned} |R_x(F^0, L) - R_x(F, L)| \\ = \frac{1}{m_x(F)} \left| \int_{\Theta} [R_x(F^0, L) - L(\theta, \delta(x))]l_x(\theta)(F^0 - F)(d\theta) \right|. \end{aligned}$$

Tezę otrzymujemy całkując przez części wyrażenie pod modulem i wykorzystując nierówności:

$$\left| \int_{\Theta} [F(\theta) - F^0(\theta)]dz(\theta) \right| \leq \rho(F, F^0) \cdot V(z)$$

oraz

$$|m_x(F) - m_x(F^0)| \leq \rho(F, F^0) \cdot V(l_x),$$

a stąd

$$m_x(F) \geq m_x(F^0) - \rho(F, F^0) \cdot V(l_x). \blacksquare$$

Twierdzenie 15. Jeżeli funkcja $l_x(\cdot)$ jest funkcją o wahanii skończonym, to

$$\rho(F_x, F_x^0) \leq \frac{\rho(F, F^0)}{\max\{m_x(F), m_x(F^0)\}} \cdot \left\{ \sup_{\theta \in \Theta} |l_x(\theta)| + V(l_x) \right\}.$$

Istnieje funkcja wiarygodności przy której nierówność jest dokładna.

Dowód. Dowód jest podobny do dowodu twierdzenia 14. Dokładnie przedstawiony jest w pracy Boratyńska i Zieliński (1993). $\blacksquare$

Przykład 6. Jako przykład ilustrujący dokładność oszacowań za pomocą nierówności z twierdzenia 14 rozpatrzmy następujący model. Niech $\mathcal{X} = \mathcal{R}^1$, $l_x(\theta) = \theta x \cdot 1_{(0, \sqrt{2/\theta})}(x)$ i $\Theta = (0, 1)$. Rozpatrzmy następujące rozkłady a priori:

$$F^0(\theta) = \begin{cases} \frac{b}{a} \theta & \text{dla } 0 < \theta < a, \\ \frac{1-b}{1-a} \theta + \frac{b-a}{1-a} & \text{dla } a < \theta < 1, \end{cases}$$

oraz

$$F(\theta) = \begin{cases} \frac{1-b}{1-a}\theta & \text{dla } 0 < \theta < 1-a, \\ \frac{b}{a}\theta + 1 - \frac{b}{a} & \text{dla } 1-a < \theta < 1, \end{cases}$$

gdzie a, b należą do przedziału $(0, 1)$ i $b > a$. Odległość

$$\rho(F^0, F) = \frac{b-a}{1-a}$$

dąży do 0 przy $b \rightarrow a$. Niech $L(\theta, \delta(x)) = \frac{1}{\theta}$. Wtedy lewa strona nierówności (oznaczymy ją przez $LHS(a, b)$) wynosi:

$$LHS(a, b) = \frac{4(b-a)}{(1+a-b)(1+b-a)},$$

podczas gdy prawa strona (oznaczymy ją przez $RHS(a, b)$) jest równa

$$RHS(a, b) = \frac{4(b-a)}{(1+a-b)(1-3b+2a-a^2+ab)}.$$

Dla wybranych wartości a i $b = 2a$ wyniki przedstawione są w poniższej tabeli.

a	LHS(a,2a)	RHS(a,2a)
0.001	0.004	0.00402
0.002	0.008	0.008081
0.003	0.012	0.012182
0.004	0.016	0.016325
0.005	0.020001	0.02051
0.006	0.024001	0.024738
0.007	0.028001	0.029008
0.008	0.032002	0.033322
0.009	0.036003	0.03768
0.01	0.040004	0.042083

■

Z twierdzeń 14 i 15 uzyskujemy natychmiast wnioski podające górne ograniczenia dla oscylacji pewnych wielkości a posteriori i rozkładu a posteriori w metryce Kolmogorowa przy małych zaburzeniach rozkładu a priori.

WNIOSEK 5. Niech $\Gamma_\varepsilon(F^0) = \{F : \rho(F, F^0) < \varepsilon\}$ będzie otoczeniem rozkładu a priori F^0 i niech x będzie ustalonym punktem przestrzeni prób.

A) Jeżeli funkcja $l_x(\cdot)$ jest funkcją o wahaniu skończonym, to

$$\rho_x^\rho(\Gamma_\varepsilon(F^0), F^0) \leq \frac{\varepsilon \cdot \{\sup_{\theta \in \Theta} |l_x(\theta)| + V(l_x)\}}{\max\{\inf_{F \in \Gamma_\varepsilon(F^0)} m_x(F), m_x(F^0)\}}.$$

B) Jeżeli dodatkowo funkcja $L(\cdot, \delta(x))l_x(\cdot)$ jest funkcją o wahanii skończonym oraz $\varepsilon < \frac{m_x(F^0)}{V(l_x)}$, to

$$\rho_x^0(\Gamma_\varepsilon(F^0), F^0, R_x) \leq \frac{\varepsilon \cdot V(z)}{m_x(F^0) - \varepsilon V(l_x)}. \blacksquare$$

Korzystając z lematów 7 i 8 oraz wniosku 2 otrzymujemy następujący wniosek.

WNIOSEK 6. Niech otoczenie $\Gamma(F^0)$ rozkładu a priori F^0 w rodzinie $\mathcal{P}$ będzie równe $\Gamma_\varepsilon(F^0, \mathcal{Q})$, $\Gamma_{DB}^{1+\varepsilon}(F^0)$ lub $\Gamma_{DR}^{1+\varepsilon}(F^0)$.

A) Jeżeli funkcja $l_x(\cdot)$ jest funkcją o wahanii skończonym, to

$$\rho_x^0(\Gamma(F^0), F^0) \leq \frac{\varepsilon \cdot \{\sup_{\theta \in \Theta} |l_x(\theta)| + V(l_x)\}}{\max\{\inf_{F \in \Gamma(F^0)} m_x(F), m_x(F^0)\}}.$$

B) Jeżeli dodatkowo funkcja $L(\cdot, \delta(x))l_x(\cdot)$ jest funkcją o wahanii skończonym oraz $\varepsilon < \frac{m_x(F^0)}{V(l_x)}$, to

$$\rho_x^0(\Gamma(F^0), F^0, R_x) \leq \frac{\varepsilon \cdot V(z)}{m_x(F^0) - \varepsilon V(l_x)}. \blacksquare$$

Zauważmy, że rząd infinitesimalnej odporności we wszystkich rozpatrywanych wyżej przypadkach jest $O(\varepsilon)$, przy czym prawe strony powyższych nierówności podają ograniczenia na stałe tej zbieżności.

4.4. Uogólnienia.

A) **Metryka Levy'ego.** Jeżeli rozkład F^0 ma ograniczoną funkcję gęstości f^0 , to twierdzenia o ciągłości i o pochodnej, to znaczy twierdzenia 9, 10, 11 zachodzą przy metryce Levy'ego $\mathcal{L}(F, F^0)$. Wynika to z następujących nierówności (patrz Zołotariew (1986))

$$\mathcal{L}(F, F^0) \leq \rho(F, F^0)$$

oraz

$$\rho(F, F^0) \leq (1 + \sup_{\theta \in \Theta} |f^0(\theta)|) \mathcal{L}(F, F^0).$$

B) **Metryka całkowitego wahanía.** Rozpatrzmy w przestrzeni $\mathcal{P}$ rozkładów a priori metrykę σ i generowane przez nią otoczenie rozkładu F^0

$$\Gamma_\varepsilon^\sigma(F^0) = \{F : \sigma(F, F^0) \leq \varepsilon\}.$$

Funkcjonał $R_x(\cdot, L)$ jest ciągły i różniczkowalny także przy tej metryce. Ciągłość wynika z twierdzenia 9. Rodzina $\Gamma_\varepsilon^\sigma(F^0)$ jest podrodziną rodziny $\Gamma_\varepsilon(F^0)$. W celu wyznaczenia pochodnej rozpatrzmy w przestrzeni $\mathcal{G}$ (zdefiniowanej w rozdziale 4.2) normę

$$\|\delta\|_\sigma = \sigma(\Delta, 0) = \sup_{A \in \mathcal{F}_\Theta} |\Delta\{A\}|.$$

Odpowiednie twierdzenie przyjmuje następującą postać:

Twierdzenie 16. *Jeżeli przy ustalonym $x \in \mathcal{X}$ funkcje $l_x(\cdot)$ oraz $L(\cdot, \delta(x))l_x(\cdot)$ są ograniczone, to*

$$\dot{R}_x(F^0, L)\Delta = \frac{\int_{\Theta} L(\theta, \delta(x))l_x(\theta)\Delta(d\theta) - R_x(F^0, L)m_x(\Delta)}{m_x(F^0)}$$

gdzie $\dot{R}_x(F^0, L)$ oznacza pochodną w sensie definicji 8 w odniesieniu do metryki σ . ■

Ruggeri i Wasserman (1993) wyznaczają normę pochodnej $\|\dot{R}_x(F^0, L)\|$. Jest ona równa

$$\|\dot{R}_x(F^0, L)\| = \frac{1}{m_x(F^0)}(\bar{h} - \underline{h}),$$

gdzie

$$\bar{h} = \sup_{\theta \in \Theta} \{l_x(\theta)[L(\theta, \delta(x)) - R_x(F^0, L)]\}$$

oraz

$$\underline{h} = \inf_{\theta \in \Theta} \{l_x(\theta)[L(\theta, \delta(x)) - R_x(F^0, L)]\}.$$

C) Stabilność gęstości a posteriori. Cuevas i Sanz (1988) wyznaczają z kolei pochodną dla operatora postaci

$$B_x : \mathcal{P} \rightarrow \mathcal{L}^1 \quad B_x(F) = f_x,$$

gdzie $\mathcal{L}^1$ jest przestrzenią funkcji rzeczywistych na Θ , całkowalnych względem miary Lebesgue'a, z normą

$$\|f\|_1 = \int_{\Theta} |f(\theta)|d\theta.$$

Otrzymujemy twierdzenie.

Twierdzenie 17 (Cuevas i Sanz (1988)). *Niech $\Theta = \mathcal{R}^1$ i niech $l_x(\cdot)$ będzie ciągłą funkcją o wahanii skończonym. Niech x będzie ustalonym punktem przestrzeni prób. Wtedy operator B_x jest różniczkowalny (w sensie definicji 8) w F dla wszystkich $F \in \mathcal{P}$ w odniesieniu do metryki ρ . ■*

5. Uwagi końcowe. W pracy przedstawiłam niektóre z problemów dotyczących odporności w modelach bayesowskich. Główną uwagę poświęciłam problemom odporności infinitezymalnej. Omówiłam też istotny problem wyznaczania procedur optymalnych. Oto kilka dalszych uwag.

1. W rozważaniach dotyczących badania wielkości a posteriori wykorzystywałam funkcję wiarygodności $l_x(\cdot)$. Zauważmy, że jeżeli $T : \mathcal{X} \rightarrow \mathcal{T}$ jest statystyką dostateczną, to, jeżeli zachodzi twierdzenie o faktoryzacji, wszystkie rozważania możemy prowadzić dla funkcji wiarygodności generowanej przez statystykę dostateczną. Wtedy bowiem istnieją funkcje $g : \mathcal{T} \times \Theta \rightarrow \mathcal{R}^1$

i $h : \mathcal{X} \rightarrow \mathcal{R}^1$ takie, że $l(x, \theta) = g(T(x), \theta)h(x)$, stąd rozkład a posteriori przyjmuje postać:

$$F_x(\theta) = \frac{\int_{\{z \leq \theta\}} g(T(x), z) F(dz)}{\int_{\Theta} g(T(x), z) F(dz)}.$$

Założenia dotyczące funkcji $l_x(\cdot)$ możemy przenieść na funkcję $g(T(x), \cdot)$ nie zmniejszając ogólności twierdzeń dotyczących wielkości a posteriori. W analizie funkcjonału $R_x(F^0, L)$ oraz operatora $\bar{R}_x$ oczywiście możemy tak postąpić, gdy $\delta(x)$ jest funkcją statystyki dostatecznej T .

2. W pracy zajmowałam się wielkościami a posteriori. Istnieją też prace poświęcone badaniu stabilności ryzyka bayesowskiego, to znaczy wielkości zdefiniowanej następująco.

$$r(F, \delta) = \int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x)) P_{\theta}(dx) F(d\theta),$$

gdzie δ jest regułą decyzyjną. Przypuśćmy, że δ_0 jest regułą bayesowską przy rozkładzie a priori F^0 , to znaczy, że

$$r(F^0, \delta_0) = \inf_{\delta \in \mathcal{D}} r(F^0, \delta),$$

gdzie $\mathcal{D}$ jest klasą rozważanych reguł decyzyjnych.

Stabilność reguły δ_0 możemy mierzyć wielkością:

$$\sup_{F \in \Gamma(F^0)} [r(F, \delta_0) - \inf_{\delta \in \mathcal{D}} r(F, \delta)].$$

Infinitymalną stabilność uzyskamy jeżeli

$$r(F^n, \delta_0) - \inf_{\delta \in \mathcal{D}} r(F^n, \delta) \rightarrow 0$$

przy $F^n \rightarrow F^0$ w pewnym sensie (na przykład może to być słaba zbieżność lub zbieżność w pewnej metryce). Problem ten poruszają między innymi Kadane i Chuang (1978) oraz Salinetti (1992). Zauważmy, że jeżeli dla reguły decyzyjnej δ zachodzi

$$\int_{\Theta} \int_{\mathcal{X}} L(\theta, \delta(x)) P_{\theta}(dx) F(d\theta) = \int_{\mathcal{X}} \int_{\Theta} L(\theta, \delta(x)) F_x(d\theta) P(dx),$$

gdzie $P(dx)$ jest rozkładem brzegowym zmiennej losowej X , to

$$r(F, \delta) = \int_{\mathcal{X}} R_x(F, L_{\delta(x)}) P(dx).$$

W badaniu infinitymalnej stabilności ryzyka bayesowskiego można wykorzystać więc wyniki dotyczące ciągłości i różniczkowalności funkcjonału $R_x(F, L_{\delta(x)})$, które przedstawione zostały w powyższej w pracy.

Bibliografia

- Basu, S. i DasGupta, A. (1990): *Bayesian analysis under distribution bands*. Technical Report # 90-48, Dept. of Stat., Purdue University.
- Bednarski, T. (1985): *On minimum bias and variance estimation for parametric models with shrinking contamination*. Probab. Math. Statist. 6, 121–129.
- Berger, J.O. (1980): *Statistical decision theory*. Foundations, Concepts, and Methods. Springer-Verlag.
- Berger, J.O. (1984): *The robust Bayesian viewpoint (with discussion)*. *Robustness of Bayesian Analysis*. J. Kadane, Ed. Amsterdam: North Holland, 63–124.
- Berger, J.O. (1990): *Robust Bayesian analysis: sensitivity to the prior*. J. Statist. Plann. Inference 25, 303–328.
- Berger, J.O. i Berliner, L.M. (1986): *Robust Bayes and empirical Bayes analysis with ϵ -contaminated priors*. Ann. Statist. 14, 461–486.
- Betro, B., Męczarski, M. i Ruggeri, F. (1992): *Robust Bayesian analysis under generalized moments conditions*. To appear in J. Statist. Plann. Inference.
- Betro, B. i Ruggeri, F. (1992): *Conditional Γ -minimax actions under convex losses*. Commun. Statist. — Theory Meth. 21(4), 1051–1066.
- Bickel, P.J. (1976): *Another look at robustness: a review of Reviews and some new developments*. Scand. J. Statist. 3, 145–168.
- Boratyńska, A. (1992): *Bayesian robustness in the ϵ -contaminated class of priors*. Komunikat na konferencji „International Workshop on Bayesian Robustness”, Milano, May 1992, złożone do druku.
- Boratyńska, A. i Męczarski, M. (1994): *Robust Bayesian estimation in the one-dimensional normal model*, to appear in Statistics & Decisions 1994.
- Boratyńska, A. i Zieliński, R. (1993): *Bayesian robustness via Kolmogorov metric*. Applicationes Math. 22, 139–143.
- Bose, S. (1992): *Bayesian robustness with local and nonlocal contaminations*. An invited paper at „International Workshop on Bayesian Robustness”, Milano, May 1992.
- Box, G.E.P. (1953): *Non-normality and tests on variances*. Biometrika 40, 318–335.
- Box, G.E.P. i Andersen, S.L. (1955): *Permutation theory in the derivation of robust criteria and the study of departures from assumptions*. J. Roy. Statist. Soc. Ser. B 17, 1–34.
- Clarke, B.R. (1983): *Uniqueness and Frechet differentiability of functional solutions to maximum likelihood type equations*. Ann. Statist. 11, 1196–1205.
- Cuevas, A. i Sanz, P. (1988): *On differentiability properties of Bayes operators*. Bayesian Statistics 3. J. M. Bernardo, M. H. DeGroot, D. V. Lindley i A. F. M. Smith (eds), Oxford University Press, 569–577.
- DasGupta, A. i Studden, W.J. (1988): *Robust Bayesian analysis and optimal experimental designs in normal linear models with many parameters*. Technical Report # 88-14, Dept. of Stat. Purdue University.
- DasGupta, A. i Studden, W.J. (1989): *Frequentist behavior of robust Bayes estimates of normal means*. Statistics & Decisions 7, 333–361.
- DeRobertis, L. i Hartigan, J.A. (1981): *Bayesian inference using intervals of measures*. Ann. Statist. 9, 235–244.
- Diaconis, P. i Freedman, D. (1986): *On the consistency of Bayes estimates*. Ann. Statist. 14, 1–67.
- Fortini, S. i Ruggeri, F. (1990): *Concentration function in a robust Bayesian framework*. Quaderno IAMI 90.6, Milano: CNR-IAMI.

- Gelfand, A.E. i Dey, D.K. (1991): *On Bayesian robustness of contaminated class of priors*. Statistics & Decisions 9, 63–80.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, R.J., Stahel, W.A. (1986): *Robust statistics; the approach based on influence functions*. Wiley, New York.
- Huber, P.J. (1973): *The use of Choquet capacities in statistics*. Bull. Inst. Internat. Statist. 45, 181–191.
- Huber, P.J. (1981): *Robust statistics*. Wiley, New York.
- Kadane, J.B. i Chuang, D.T. (1978): *Stable decision problems*. Ann. Statist. 6, 1095–1110.
- Lavine, M. (1991a): *An approach to robust Bayesian analysis for multidimensional parameter spaces*. JASA 86, 400–403.
- Lavine, M. (1991b): *Sensitivity in Bayesian statistics: the prior and the likelihood*. JASA 86, 396–399.
- Lavine, M., Wasserman, L., Wolpert, R.L. (1991): *Bayesian inference with specified priors marginals*. JASA 86, 964–971.
- Leamer, E.E. (1982): *Sets of posterior means with bounded variance priors*. Econometrica 50, 725–736.
- Męczarski, M. (1993): *Stable Bayesian estimation in the Poisson model*. Zeszyty Nauk. PL - Sci. Bull. Łódź Technical Univ. No. 687-Matematyka, z.25, 83–91.
- Męczarski, M. (1993): *Stability and conditional Γ -minimaxity in Bayesian inference*. Applicationes Math. 22, 117–122.
- Męczarski, M. i Zieliński, R. (1991): *Stability of the Bayesian estimator of the Poisson mean under the inexactly specified gamma priors*. Statistics & Probability Letters 12, 329–333.
- Moreno, E. i Cano, J.A. (1991): *Robust Bayesian analysis with ε -contaminations partially known*. J. R. Statist. Soc. B 53, 143–155.
- Moreno, E. i Pericchi, L.R. (1993): *On ε -contaminated priors with quantile and piecewise unimodality constraints*. Commun. Statist. Theory Meth. 22(7), 1963–1978.
- O'Hagan, A. i Berger, J.O. (1988): *Ranges of posterior probabilities for quasi-unimodal priors with specified quantiles*. JASA 83, 503–508.
- Poetzelberger, K. i Polasek, W. (1991): *Robust HPD-regions in Bayesian regression models*. Econometrica 59, 1581–1589.
- Rachev, S.T. (1991): *Probability metrics and the stability of stochastic models*. Wiley, Chichester.
- Ruggeri, F. (1990): *Posterior ranges of functions of parameters under priors with specified quantiles*. Commun. Statist.—Theory Meth. 19, 127–144.
- Ruggeri, F. (1991): *Robust Bayesian analysis given a lower bound on the probability of a set*. Commun. Statist.—Theory Meth. 20, 1881–1891.
- Ruggeri, F. i Wasserman, L. (1991): *Density based classes of priors: infinitesimal properties and approximations*. Quaderno IAMI 91.4, CNR-IAMI, Milano.
- Ruggeri, F. i Wasserman, L. (1993): *Infinitesimal sensitivity of posterior distributions*. Canadian J. Statist. 21(2), 195–203.
- Salinetti, G. (1992): *Stability of Bayesian decisions*. An invited paper at „International Workshop on Bayesian Robustness”, Milano, May 1992.
- Serfling, R.J. (1991): *Twierdzenia graniczne statystyki matematycznej*. PWN, Warszawa.
- Sivaganesan, S. (1988): *Range of posterior measures for priors with arbitrary contaminations*. Commun. Statist.—Theory Meth. 17(5), 1591–1612.
- Sivaganesan, S. and Berger, J.O. (1989): *Ranges of posterior measures for priors with unimodal contaminations*. Ann. Statist. 17, 868–889.

- Wasserman, L. i Kadane, J. (1992): *Computing bounds on expectations*. JASA 87, 516–522.
- Zieliński, R. (1983): *Robust statistical procedures: a general approach*. Lecture Notes in Mathematics 982, Springer-Verlag, 283–295.
- Zolotarev, V.M. (1986): *Contemporary Theory of Summation of Independent Random Variables*. Nauka, Moscow. (In Russian).

Abstract

The problem of measuring the Bayesian robustness is considered. An upper bound for the oscillation of a posterior functional in terms of the Kolmogorov distance between the prior distributions is given. The norm of the Frechet derivative as a measure of local sensitivity is presented. The problem of finding optimal statistical procedures is presented.

INSTYTUT MATEMATYKI STOSOWANEJ I MECHANIKI
UNIwersytet Warszawski
BANACHA 2
02-097 WARSZAWA