

83/2004



UNIVERZITA KARLOVA V PRAZE

MATEMATICKO-FYZIKÁLNÍ FAKULTA

TROJA

BAYESOVSKÉ METODY

RNDr. MARIE HUŠKOVÁ, CSc.



SPN Praha

O B S A H

P ř e d m l u v a

5

Kapitola 1. Ú v o d

1.1 Formulace problematiky

7

1.2 Bayesova věta a její použití

10

Kapitola 2. V o l b a a p r i o r n í h o r o z d ě l e n í

2.1 Úvod

13

2.2 Konjugované systémy hustot

15

2.3 Princip neurčitosti, Jeffreysova hustota a limitní aposteriorní hustoty

26

2.4 Empirické bayesovské metody

35

Kapitola 3. S t a t i s t i c k é r o z h o d o v a c í

f u n k c e

3.1 Formulace problému

38

3.2 Bayesovské rozhodovací funkce

43

Kapitola 4. ů l o h a o d h a d u

4.1 Úvod

50

4.2 Bodový odhad; jednorozměrný případ

50

4.3 Bodový odhad; vícerozměrný případ

61

4.4 Věrohodnostní množiny

65

Kapitola 5.	T e s t o v á n í h y p o t é z	
5.1	Úvod	
5.2	Ztrátové funkce používané při testování hypotéz	70
5.3	Testy při $\lambda(\Theta_0) = 0$	70
5.4	Testy o střední hodnotě normálního rozdělení	76
		79
A p e n d i x :	Přehled použitých rozdělení	86
Literatura		93

Bayesovské metody představují jeden ze základních přístupů teoreticko-pravděpodobnostního myšlení i matematicko-statistických vyhodnocovacích metod.

Vychází se z předpokladu, že naše informace (apriorní znalost, zkušenost) o hodnotě neznámého parametru může být vyjádřena pomocí pravděpodobnostního rozdělení, tj. neznámý parametr můžeme považovat za náhodnou veličinu. K závěrům o hodnotě neznámého parametru využijeme jednak apriorní informaci o hodnotě parametru, jednak experimentální výsledky (nezávislé na této apriorní informaci). Tento přístup byl a dosud je předmětem kritiky mnoha statistiků.

Na bayesovské metody však můžeme hledět jako na metody, které nám poskytují jisté řešení statistických problémů.

Bez ohledu na výše zmíněnou kritiku mohou být bayesovské metody užitečné v řadě praktických situací, především v případech, kdy jsou dostupné výsledky obdobných experimentů z minulosti (např. při kontrole jakosti výrobků).

Účelem skriptu je vyloužit základy bayesovských metod v úlohách testování hypotéz a teorie odhadu. Skriptum je rozděleno do šesti kapitol. V první kapitole jsou vloženy základní principy bayesovského přístupu k řešení statistických problémů. V druhé se čtenář seznámí s možnostmi volby apriorního rozdělení, v další se základy teorie rozhodovacích funkcí. Čtvrtá a pátá kapitola je věnována teorii odhadu resp. testování hypotéz. Poslední kapitola obsahuje přehled použitých rozdělení.

Skriptum bylo napsáno jako pomůcka k přednášce Matematická statistika II., ale může sloužit i širšímu okruhu čtenářů, neboť pokud je mi známo, nebyla dosud v českém jazyce publikována samostatná knížka věnovaná pouze bayesovským metodám.

Předpokládá se, že čtenář je seznámen se základy vyšší matematiky a matematické statistiky na úrovni knihy J. Anděla: Matematická statistika, kap. 1-10, 13-15. Značení je převzato z této knihy.

Závěrem bych chtěla poděkovat recenzentovi prof. Ing. F. Fabianovi, OSc a dr. D. Vorlíčkové za podnětné připomínky a paní I. Marešové za pečlivé přepsání rukopisu.

1. Ú V O D

1.1 FORMULACE PROBLÉMATIKY

Nechť $\tilde{X} = (X_1, \dots, X_n)'$ je náhodný vektor s hustotou $r(\tilde{x}|\theta)$ vzhledem k σ -konečné míře ν_n , kde $\theta = (\theta_1, \dots, \theta_k)'$ je parametr s hodnotami z neprázdné borelovské množiny $\Theta \subset R_k$.

Při klasickém (nikoli bayesovském) přístupu k problému odhadu parametru θ nebo testování hypotézy o θ považujeme θ za neznámou konstantu popř. vektor neznámých konstant a k závěrům o hodnotě parametru θ použijeme pouze $\tilde{X} = (X_1, \dots, X_n)'$ a tvar rozdělení \tilde{X} .

Při bayesovském přístupu k závěrům o parametru θ použijeme kromě \tilde{X} ještě informaci (byť neúplnou) o parametru θ , kterou máme k dispozici nezávisle na realizaci \tilde{X} . Mluvíme o tzv. apriorní informaci. Tato informace může mít objektivní či subjektivní charakter, popř. může být kombinací informací obou typů. O objektivní apriorní informaci mluvíme, jestliže využijeme informaci z podobných úloh, problémů z minulosti. Subjektivní apriorní informace vyjadřuje názor či zkušenost nějakého subjektu. Apriorní informace se vyjadřuje předpokladem, že θ je náhodný vektor popř. náhodná veličina s rozdělením, které je více či méně známo v závislosti na tom, jak úplnou či neúplnou informaci o θ máme.

Vyjmecně pracujeme s náhodným θ jako metodou získání závěrů o parametru θ (náhodnost tedy nevyjadřuje žádnou apriorní informaci, ale slouží jako prostředek jak obdržet závěry o θ). V kapitole 3 uví- dáme, že znárodnění parametru θ je jedna z možností jak zavést uspořá- dání na množině rozhodovacích funkcí a definovat kritérium optimality.

Příklad 1.1. Uvažujeme problém odhadu kvocientu inteligence θ u určitěho dítěte na základě testu s výsledkem X . Dlouholeté výzkumy ukazují, že X má rozdělení $N(\theta, 100)$, kde θ je kvocient inteligence, že θ je obecně různé u různých dětí a lze ho považovat za náhodnou veličinu s rozdělením $N(100, 225)$. Poslední uvedený fakt lze považovat za objektivní apriorní informaci - závěr ze série předchozích realizací.

Příklad 1.2. Na základě krevní zkoušky se má rozhodnout, zda pacient trpí jistou chorobou. Z předchozích výzkumů je známo, že touto chorobou trpí asi 5 % populace. Toto je opět objektivní apriorní informace dostupná před provedením zkoušky. Při bayesovském přístupu použijeme pro zmíněné rozhodnutí jak výsledku krevní zkoušky, tak apriorní informace o procentu populace trpící touto chorobou.

Příklad 1.3. Úkolem fyzika je odhadnout jistou fyzikální konstantu θ . Fyzik má určitou představu o možných hodnotách θ . Připouští několik možných hodnot θ , příkládá jim obecně různé váhy (pravděpodobnosti) a tedy považuje je z tohoto hlediska za náhodné veličiny. Různí fyzikové mohou vyjádřit svou představu o hodnotě parametru θ obecně různými rozděleními. Tato apriorní informace je subjektivní. Opět k závěrům o hodnotě fyzikální konstanty použijeme jak výsledku (popř. výsledků) příslušného experimentu tak apriorní informace.

Příklad 1.4. Při předpovědi počasí se běžně používají nejen výsledky měření provedených v minulosti a současnosti, ale i subjektivní názory (informace) zkušených meteorologů. Kromě výsledků měření v současnosti tedy použijeme jak objektivních tak subjektivních apriorních informací.

Je nutné si povšimnout rozdílné interpretace rozdělení parametru θ v jednotlivých příkladech. Zatímco v 1. a 2. příkladě bylo rozdělení parametru θ získáno z řady (objektivních) měření a θ lze

skutečně pokládat za náhodnou veličinu, ve třetím příkladě rozdělenní parametru θ vyjadřuje "stupeň víry" v jednotlivé hodnoty parametru θ (popř. víry, že náleží do určité množiny). V prvních dvou příkladech se jedná o běžné rozdělení pravděpodobnosti s četnostní interpretací tak jak se s ní běžně setkáváme v klasické statistice. Ve třetím příkladě rozdělení pravděpodobnosti parametru zachycuje "stupeň víry" v určité hodnoty θ , pro různé subjekty je obecně toto rozdělení různé. Z těchto důvodů nepřipadá v úvahu četnostní interpretace. V tomto případě mluvíme obvykle o tzv. subjektivní pravděpodobnosti, o které se předpokládá, že vyhovuje Kolmogorově definici pravděpodobnosti. Rozdíl je pouze v interpretaci. Subjektivní pravděpodobnost vyjadřuje víru subjektu, že určitý jev nastane.

Určení subjektivní pravděpodobnosti je velkým problémem. Nejjednodušší způsob určení subjektivní pravděpodobnosti je porovnat relativní věrohodnosti. Např. chceme-li najít pravděpodobnost jevu E , tj. $P(E)$, srovnáme věrohodnosti E a jeho doplnkového jevu E^c . Příkladáme-li oběma jevům stejnou šanci, klademe $P(E) = P(E^c) = 1/2$. Příkladáme-li jevu E třikrát větší šanci než E^c , klademe $P(E) = 3/4$, $P(E^c) = 1/4$. Jiná možnost je porovnávat víry ve dvojice jevů na základě sázek. Za určitých předpokladů potom existuje jediná pravděpodobnost na uvažované σ -algebře jevů. Podrobný postup spolu s diskusí o dalších možnostech lze najít např. v [2], [5].

Se subjektivní pravděpodobností se setkáváme i v běžném životě. Mluvíme o naději (šanci) oblíbeného fotbalového družstva. Uvažujeme o možnosti nepříznivého počasí o nejbližším víkendu a podobně. Příkladáme vlastně vždy možným výsledkům, obvykle říkáme, že ten či onen výsledek je nejpravděpodobnější, méně pravděpodobný či málo pravděpodobný.

Nyní obrátíme pozornost na výhody a nevýhody bayesovského pří-
stupu. Jeho kladem je bezesporu využití i apriorní informace. Na

druhou stranu je bayesovský přístup předmětem kritiky mnoha statistiků, která se týká v podstatě tří bodů, a to konstrukce rozdělení parametru ϑ na základě apriorní informace, použití subjektivní apriorní informace a v některých případech připustit, že ϑ je náhodná veličina. Část kritiky je filozofického rázu. Většina statistiků nemá námitky vůči postupu v příkladě 1, kde ϑ lze skutečně považovat za náhodnou veličinu a rozdělení parametru ϑ je konstruováno na základě předchozích objektivních měření. V příkladě 3 však může být obtížné považovat fyzikální konstantu θ za náhodnou veličinu s určitým rozdělením. Subjektivní informace může výrazně ovlivnit závěry o parametru ϑ (na základě apriorní informace dvou různých subjektů můžeme dojít k diametrálně odlišným závěrům).

Je-li apriorní informace velmi neurčitá či žádná, může volba apriorního rozdělení parametru ϑ způsobit značné problémy. Je nutné si uvědomit, že různé volby apriorního rozdělení mohou vést k různým závěrům. Poznamenejme, že pro n dostatečně velké a za jistých podmínek regularity závěry o ϑ nezávisí na volbě apriorního rozdělení parametru ϑ (viz Věta 2.1).

1.2 BAYESOVA VĚTA A JEJÍ POUŽITÍ

Nechť $\vartheta = (\theta_1, \dots, \theta_k)'$ je náhodný vektor s hustotou $q(\vartheta)$ vzhledem k σ -konečné míře λ na $(\mathcal{C}, \mathcal{B}(\mathcal{C}))$, kde \mathcal{C} je neprázdná borelovská podmnožina R_k , $\mathcal{B}(\mathcal{C})$ označuje borelovské podmnožiny \mathcal{C} .
 Nechť $\bar{X} = (X_1, \dots, X_n)'$ je náhodný vektor s podmíněnou hustotou $r(\bar{x}|\vartheta)$ při daném ϑ vzhledem k σ -konečné míře ν_n na (R_n, \mathcal{B}_n) , kde \mathcal{B}_n označuje borelovské podmnožiny R_n , tj.

$$P(\vartheta \in B, \bar{X} \in C) = \int_B \left(\int_C r(\bar{x}|\vartheta) d\nu(\bar{x}) \right) q(\vartheta) d\lambda(\vartheta), \quad (1.1)$$

kde B a C jsou libovolné měřitelné množiny.

Věta 1.1 (Bayesova). Pro podmíněnou hustotu $\pi(\theta|\bar{x})$ náhodného vektoru θ při daném \bar{x} platí

$$\begin{aligned} \pi(\theta|\bar{x}) &= \frac{q(\theta)r(\bar{x}|\theta)}{\int_{\Theta} q(\theta)r(\bar{x}|\theta)d\lambda(\theta)} && \text{je-li } \int q(\theta)r(\bar{x}|\theta)d\lambda(\theta) \neq 0, \\ &= 0 && \text{jinak.} \end{aligned} \quad (1.2)$$

Důkaz provedeme stejně jako důkaz věty III.3.14 v [1] (str. 54).

Ze vztahu (1.1) je vidět, že $r(\bar{x}|\theta)q(\theta)$ je sdružená hustota vektoru (\bar{x}', θ') vzhledem k $\nu_n \times \lambda$. Proto podle věty III.3.7 v [1] o marginální hustotě je

$$\int_{\Theta} q(\theta)r(\bar{x}|\theta)d\lambda(\theta)$$

marginální hustota vektoru \bar{x} . Tvzení věty nyní plyne z věty

III.3.13 v [1]. Q.E.D.

Tato věta má v bayesovských metodách klíčové postavení. Je-li

θ parametr uvažovaný v minulém paragrafu, nazýváme $q(\theta)$ apriorní hustotou, neboť vyjadřuje informaci o θ ještě před realizací \bar{x} .

Podmíněnou hustotu $\pi(\theta|\bar{x})$ parametru θ pak nazýváme hustotou aposteriorní, neboť jde o hustotu parametru θ po realizaci \bar{x} .

K závěrům o parametru θ pak použijeme aposteriorní hustotu $\pi(\theta|\bar{x})$, která v sobě zahrnuje jak apriorní informaci o parametru θ tak informaci plynoucí z realizace \bar{x} . Poznamenejme, že při klasickém přístupu použijeme k závěrům o parametru θ pouze $r(\bar{x}|\theta)$.

str. 8

Příklad 1.1 (pokračování). Apriorní hustota parametru θ je $N(100, 225)$, podmíněné rozdělení X při daném θ je $N(\theta, 100)$. Pak aposteriorní hustota je $N\left(\frac{225\bar{x}+100 \cdot 100}{225+100}; 69,23\right)$, je-li x hodnota X . Kvocient inteligence můžeme odhadnout (viz str. 52) θ střední hodnotou příslušnou aposteriorní hustotě, tedy hodnotou $\frac{225}{325} \cdot x + \frac{100}{325} \cdot 100$. Příslušný rozptyl je 69,23. Zatímco při klasickém přístupu bychom použili jako odhad θ přímo x , jehož rozptyl je 100. Tedy použití bayesovského

přístupu vedlo k odhadu s menším rozptylem než při klasickém přístupu.

Povšimněme si nyní vzorce (1.2). Existují-li měřitelné funkce $h_1(\vartheta, \mathbf{x})$ a $h_2(\mathbf{x})$ takové, že

$$q(\vartheta) r(\mathbf{x}|\vartheta) = h_1(\vartheta, \mathbf{x}) h_2(\mathbf{x}), \quad (1.3)$$

pak aposteriorní hustotu $\pi(\vartheta|\mathbf{x})$ můžeme přepsat následovně.

$$\begin{aligned} \pi(\vartheta|\mathbf{x}) &= \frac{h_1(\vartheta, \mathbf{x})}{\int_{\Theta} h_1(\vartheta, \mathbf{x}) d\lambda(\vartheta)} && \text{je-li } \int_{\Theta} h_1(\vartheta, \mathbf{x}) d\lambda(\vartheta) \neq 0, \\ &= 0 && \text{jinak.} \end{aligned} \quad (1.4)$$

Odtud je vidět, že nahradíme-li v (1.2) $q(\vartheta)$ funkcí $cq(\vartheta)$, kde c je kladná konstanta, aposteriorní hustota $\pi(\vartheta|\mathbf{x})$ se nezmění. Někdy dokonce za hustotu $q(\vartheta)$ volíme tzv. nevlastní hustotu, která je definovaná jakonezáporná měřitelná funkce (nemusí být integrovatelná). Jak uvidíme v dalším, použití nevlastní hustoty někdy vede k výsledkům rozumným, jindy k nesmyslným. Proto je nutné nevlastní hustoty používat velmi opatrně.

Na závěr si ještě uvedeme vzorce pro (nepodmíněnou) střední hodnotu a rozptyl náhodné veličiny X_1 .

$$E X_1 = E(E(X_1 | \vartheta)) = \int \left(\int x_1 r(\mathbf{x}|\vartheta) d\mu(\mathbf{x}) \right) q(\vartheta) d\lambda(\vartheta), \quad (1.5)$$

$$\begin{aligned} \text{var } X_1 &= E(X_1 - E(X_1 | \vartheta))^2 + \text{var } \{E(X_1 | \vartheta)\} = \\ &= E \text{var} \{X_1 | \vartheta\} + \text{var} \{E(X_1 | \vartheta)\}. \end{aligned} \quad (1.6)$$

Nepodmíněná hustota náhodného vektoru $\tilde{\mathbf{x}}$ vzhledem k ν_n je

$$r(\tilde{\mathbf{x}}) = \int_{\Theta} q(\vartheta) r(\mathbf{x}|\vartheta) d\lambda(\vartheta). \quad (1.7)$$

2. VOLBA APRIORNÍHO ROZDĚLENÍ

2.1 ÚVOD

Jak je vidět z úvah v úvodní kapitole, měla by apriorní hustota $q(\theta)$ odrážet naše apriorní informace (subjektivní i objektivní) o parametru θ . Jakmile jsou naše informace malé nebo vůbec žádné a chceme-li použít bayesovský přístup, vytvářívá problém volby apriorního rozdělení. Možnosti se dají v podstatě klasifikovat do čtyř skupin:

1. tvar hustoty $q(\theta)$ (včetně hodnoty parametrů) vyplývá z apriorní informace;
2. jako hustotu použijeme histogram (při informaci objektivní)
 ⓐ rozložíme na sjednocení disjunktních měřitelných množin a zjistíme četnosti jednotlivých množin; při subjektivní informaci také rozložíme ⓑ a zjistíme subjektivní pravděpodobnosti jednotlivých podmnožin);
3. volíme hustotu (hladkou), která dobře aproximuje histogram;
4. předpokládáme, že hustota má určitý funkcionální tvar, neznáme pouze parametry.

Tyto možnosti odpovídají v podstatě volbám rozdělení náhodných veličin při klasickém přístupu. Při bayesovském přístupu (podobně jako při klasickém dáváme přednost pracovat s hustotami $q(\theta)$ určitěho funkcionálního typu (ať už s parametry známými či neznámými). Běžně se používají jednak tzv. systemy konjugovaných rozdělení, jednak typy rozdělení, které odpovídají tzv. principu neurčitosti. Konjugovaným systémem rozdělení rozumíme takový systém, že apriorní i aposteriorní hustota do něj patří. Tento systémům je věnován § 2.2. Princip neurčitosti používáme, pokud nemáme žádnou informaci o θ .

Dle něho za apriorní rozdělení bereme rovnoměrné rozdělení na Θ .
 Příslušnou hustotu budeme značit q_0 . Další výklad je v § 2.3.

Někdy máme apriorní informaci v následujícícm tvaru: Y_1, \dots, Y_N jsou nezávislé náhodné veličiny, které představují minulé výsledky, Y_1 má podmíněnou hustotu $r(y|\theta_1)$, $\theta_1 \in \Theta$, $i = 1, \dots, M$; $\theta_1, \dots, \theta_M$ jsou nezávislé náhodné vektory, θ_i má hustotu $q(\theta)$. Pomocí náhodných veličin Y_1, \dots, Y_N odhadneme nepodmíněné rozdělení

$$r(y) = \int r(y|\theta)q(\theta)d\lambda(\theta)$$

náhodných veličin Y_i , $i = 1, \dots, M$ nebo také nepodmíněnou střední hodnotu EY_i a nepodmíněný rozptyl $\text{var } Y_i$, což nám za jistých předpokladů umožní odhadnout buď $q(\theta)$ nebo aspoň některé momenty tohoto rozdělení. Takové metody jsou známy pod názvem empirické bayesovské metody. Podrobněji se s nimi seznámíme v § 2.4.

Na závěr tohoto odstavce si zformulujeme tvrzení, které říká, že pro n dost velká a za poměrně obecných dalších předpokladů aposteriorní hustota téměř nezávisí na apriorní.

Věta 2.1. Nechť $\Theta \subset R_n$ je neprázdná borelovská množina. Mějme na borelovských podmnožinách Θ definovanu nějakou σ -konečnou míru λ .

Nechť $r(x|\theta)$ je hustota náhodného vektoru $\bar{X} = (X_1, \dots, X_n)$ vzhledem k nějaké σ -konečné míře ν , při dané hodnotě $\theta \in \Theta$. Nechť $q(\theta)$ je nezáporná omezená měřitelná funkce na Θ . Předpokládejme, že pro daný $\bar{x} \in R_n$ platí

$$0 < \int_{\Theta} r(\bar{x}|\theta)d\lambda(\theta) < +\infty, \quad 0 < \int_{\Theta} r(\bar{x}|\theta)q(\theta)d\lambda(\theta) < +\infty.$$

Označme

$$\pi_0(\theta|\bar{x}) = \frac{r(\bar{x}|\theta)}{\int_{\Theta} r(\bar{x}|\theta)d\lambda(\theta)}, \quad \pi(\theta|\bar{x}) = \frac{r(\bar{x}|\theta)q(\theta)}{\int_{\Theta} r(\bar{x}|\theta)q(\theta)d\lambda(\theta)} \quad (2.1)$$

Nechť existuje taková borelovská množina $A \subset \Theta$, že pro daná čísla a, b, c ($0 \leq a < 1$, $b > 0$, $c \geq 0$) platí:

$$\int_A \pi_0(\vartheta|\bar{x}) d\lambda(\vartheta) > 1 - a, \quad (2.2)$$

$$m = \inf_{\vartheta \in A} q(\vartheta) > 0, \quad (2.3)$$

$$\sup_{\vartheta \in A} q(\vartheta) \leq (1+h)m,$$

$$\sup_{\vartheta \in \bar{A}} q(\vartheta) \leq (1+c)m.$$

Pak platí

$$\int_{\bar{\Theta}} |\pi(\vartheta|\bar{x}) - \pi_0(\vartheta|\bar{x})| d\lambda(\vartheta) \leq \max\left(\frac{a+b}{1-a}, \frac{a+b+c}{1+a+b+c}\right) + \frac{a(2-a+c)}{1-a}.$$

Důkaz lze najít např. v [1] kap. XVI.3, Věta 1 (str. 288).

Věta říká, že aposteriorní hustoty $\pi_0(\vartheta|\bar{x})$ a $\pi(\vartheta|\bar{x})$ se nebudou příliš lišit, jestliže a a b budou dostatečně malá nezáporná čísla a c nebude příliš velká. Uvědomíme-li si význam čísel a, b, c , předpoklady znamenají, že na množině A je koncentrována velká část pravděpodobnosti odpovídající $\pi_0(\vartheta|\bar{x})$, $q(\vartheta)$ musí být na A prakticky konstantní a nemílová a omezená na $\Theta - A$. Je-li X_1, \dots, X_n náhodný výběr z rozdělení $r(x|\vartheta)$, pak při dostatečně velkém n je často aposteriorní hustota $\pi_0(\vartheta|\bar{x})$ koncentrována kolem nějakého bodu. Pak lze najít množinu A (většinou k -rozměrný interval) splňující (2.2) a takovou, že $\lambda(A)$ je velmi malá a široká třída hustot $q(\vartheta)$ splňuje (2.3), pak většinou stačí, aby $q(\vartheta)$ byla hladká na A a omezená na $\Theta - A$.

2.2 KONJUGOVANÉ SYSTÉMY HUSTOT

Nechť $\bar{X} = (X_1, \dots, X_n)$ je náhodný výběr z rozdělení s hustotou

$r(x|\theta)$ vzhledem k σ -konečné míře ν , $\theta \in \Theta \neq \emptyset$, $\Theta \in \mathcal{B}_k$.

Systém Q apriorních hustot $q(\theta)$ nazveme systemem konjugovaných hustotami $\{r(x|\theta), \theta \in \Theta\}$, jestliže při dost velkém n a při libovolných hodnotách $\bar{X} = \bar{x}$, které splňují

$$0 < \int_{\Theta} \prod_{i=1}^n r(x_i|\theta) q(\theta) d\lambda(\theta) < +\infty,$$

patří aposteriorní hustota do systému Q .

Vezmeme-li za Q systém všech hustot vzhledem k σ -konečné míře λ , jde samozřejmě o systém hustot konjugovaných s $\{r(x|\theta); \theta \in \Theta\}$. Pro praktické účely se příliš nehodí. Vhodné budou systémy obsahující co nejméně hustot.

Poznamenejme, že použití konjugovaných systémů hustot vede většinou k poměrně jednoduchým výsledkům.

Nyní si uvedeme metodu konstrukce konjugovaných systémů hustot využívající postaćující statistiky. Označme $T_n(X_1, \dots, X_n)$ postaćující statistiku příslušející systému hustot $\{r(x_i|\theta); \theta \in \Theta\}$. Předpokládejme, že existuje n_0 takové, že pro vš. $n > n_0$ existují nezáporné funkce g_n a h_n takové, že

$$\prod_{i=1}^n r(x_i|\theta) = g_n(T_n(\bar{x}); \theta) h_n(\bar{x}), \quad (2.4)$$

kde $T_n(X_1, \dots, X_n)$ je r -rozměrná postaćující statistika, r nezávislá na n . Označme $S_n = \{t; t = T_n(\bar{X})\}$ množinu vš. bodů $t \in R_r$, kterých může nabývat náhodný vektor $T_n(\bar{X})$. Předpokládejme, že pro každé $t \in S_n$ platí

$$0 < \int_{\Theta} g_n(t; \theta) d\lambda(\theta) < +\infty.$$

Pak systém hustot

$$\left\{ f_{n,t}(\theta); t \in S_n, n > n_0 \right\} \quad (2.5)$$

je systém konjugovaný s $\{r(x|\theta); \theta \in \Theta\}$, kde

$$f_{n,t}(\theta) = g_n(t; \theta) \left(\int_{\Theta} g_n(t; \theta) d\lambda(\theta) \right)^{-1}. \quad (2.6)$$

Ukážeme si, že toto tvrzení skutečně platí. Nechť X_1, \dots, X_n a Y_1, \dots, Y_m jsou dva nezávislé výběry z rozdělení s hustotou $r(\mathbf{x}|\theta)$. Nechť $m \geq n_0$, $n \geq n_0$. Sdružená hustota obou výběrů má tvar

$$\prod_{i=1}^n r(x_i|\theta) \cdot \prod_{j=1}^m r(y_j|\theta) =$$

takže

$$= g_{n+m}(\pi_{n+m}(\bar{x}, \bar{y}), \theta) h_{n+m}(\bar{x}, \bar{y}) = g_n(\pi_n(\bar{x}), \theta) g_m(\pi_m(\bar{y}), \theta) \cdot h_n(\bar{x}) h_m(\bar{y}). \quad (2.7)$$

Pro jednoduchoost nyní předpokládejme, že $h_n(\bar{x}) > 0$ pro každé $n > n_0$. Nechť $t \in S_m$. Pak existuje takové $\bar{y} \in R_m$, že $t = \pi_m(\bar{y})$. Je-li apriorní hustota parametru θ rovna $f_m, t(\theta)$, pak podle Bayesovy věty je aposteriorní hustota dána vzorcem

$$\pi(\theta|\bar{x}) = k \prod_{i=1}^n r(x_i|\theta) f_m, t(\bar{y}),$$

kde k je normalizační konstanta. Odtud dále dostaneme

$$\begin{aligned} \pi(\theta|\bar{x}) &= k g_n(\pi_n(\bar{x}), \theta) h_n(\bar{x}) g_m(t, \theta) \left(\int_{\Theta} g_m(t, \theta) d\lambda(\theta) \right)^{-1} = \\ &= \frac{k g_n(\pi_n(\bar{x})) g_m(\pi_m(\bar{y})) h_n(\bar{x}) h_m(\bar{y})}{h_m(\bar{y}) \int_{\Theta} g_m(t, \theta) d\lambda(\theta)}. \end{aligned}$$

Tedy z (2.7) vyplývá

$$\pi(\theta|\bar{x}) = c g_{n+m}(y, \theta) h_{n+m}(\bar{x}, \bar{y}),$$

$$\text{kde } y = \pi_{n+m}(\bar{x}, \bar{y}), \quad c = \frac{k}{\int_{\Theta} g_m(t, \theta) d\lambda(\theta) h_m(\bar{y})}.$$

Odtud je již vidět, že

$$\pi(\theta|\bar{x}) = g_{n+m}(y, \theta) \left(\int_{\Theta} g_{n+m}(y, \theta) d\lambda(\theta) \right)^{-1}, \quad (2.8)$$

kde $y \in S_{n+m}$. Tím jsme dokázali, že π patří do systému (2.5).

Systém hustot (2.5) někdy nazýváme přirozený konjugovaný systém. Jak uvidíme z následujících příkladů často pracujeme se systémem hustot, který je o něco bohatší než přirozený konjugovaný systém; budeme ho nazývat obvyklým.

Nyní si uvedeme přehled konjugovaných systémů hustot, které přísluší nejběžněji používaným systémům rozdělení $\{r(x|\theta); \theta \in \Theta\}$. Podobně probereme případ binomického a normálního (jedno- i vícerozměrného) rozdělení.

Jednorozměrná rozdělení

Binomické rozdělení s parametry (m, θ) ; m je dáno; $\Theta = \langle 0, 1 \rangle$.

Postačující statistika je $\sum_{i=1}^n X_i$, tedy $S_n = \{0, 1, 2, \dots, nm\}$. Vzhledem k (A.1) je

$$\xi_n(y; \theta) = \theta^y (1-\theta)^{nm-y}, \quad y \in S_n$$

$$h_n(\bar{x}) = \prod_{i=1}^n \binom{m}{x_i}.$$

Odtud a z (2.6) plyne

$$f_{n,y}(\theta) = \theta^y (1-\theta)^{nm-y} (B(y+1, nm-y+1))^{-1} \quad \theta \in (0, 1)$$

$n = 1, 2, \dots$; $y = 0, 1, 2, \dots, nm$.

Pedy přirozený konjugovaný systém je systém beta rozdělení s parametry $(y+1, nm-y+1)$, kde $y = 0, 1, \dots, nm$, $n = 1, 2, 3, \dots$

Obvyklý konjugovaný systém je systém beta rozdělení s parametry (α, β) kde $\alpha > 0$, $\beta > 0$.

Je-li a priori rozdělení beta rozdělení s parametry (α, β) , je a posteriorní rozdělení beta rozdělení s parametry $(\alpha + \sum_{i=1}^n X_i, \beta + nm - \sum_{i=1}^n X_i)$. Marginální rozdělení (X_1, \dots, X_n) je podle (1.7)

$$P(X_1=x_1, \dots, X_n=x_n) = \frac{B(\alpha + \sum_{i=1}^n x_i, \beta + nm - \sum_{i=1}^n x_i)}{B(\alpha, \beta)}, \quad x_i = 0, \dots, m, \quad i = 1, \dots, n.$$

Poissonovo rozdělení s parametrem θ ; $\Theta = (0, +\infty)$.

Postačující statistika je $\sum_{i=1}^n X_i$, tedy $S_n = \{0, 1, 2, \dots\}$.

Přirozený konjugovaný systém vzhledem k Lebesguově míře je systém gama rozdělení s parametry (m, t) , $t = 1, 2, \dots; m = 1, 2, \dots$. Obvyklý konjugovaný systém tvoří systém gama rozdělení s parametry (a, p) , $a > 0, p > 0$. Je-li apriorní rozdělení gama rozdělení (m, t) , je aposteriorní rozdělení též gama s parametry $(m+n; t + \sum_{i=1}^n X_i)$.

Negativně binomické rozdělení s parametry (s, θ) ; s je známé, $\theta = \langle 0, 1 \rangle$.

Postačující statistika je $\sum_{i=1}^n X_i$; $S_n = \{0, 1, 2, \dots\}$.

Přirozený konjugovaný systém je systém beta rozdělení s parametry $(sm+1, t)$, $t = 1, 2, \dots; m = 1, 2, \dots$

Obvyklý konjugovaný systém je systém beta rozdělení s parametry

$$(\alpha, \beta), \quad \alpha > 0, \quad \beta > 0.$$

Je-li apriorní rozdělení beta rozdělení (α, β) , pak aposteriorní rozdělení je opět beta rozdělení $(\alpha + sm, \beta + \sum_{i=1}^n X_i)$.

Exponenciální rozdělení s parametrem θ ; $\theta = (0, +\infty)$.

Postačující statistika je $\sum_{i=1}^n X_i$, tedy $S_n = (0, +\infty)$.

Přirozený konjugovaný systém je systém gama rozdělení s parametry (t, m) , $t > 0, m = 1, 2, \dots$

Obvyklý konjugovaný systém je systém gama rozdělení (t, m) , $t > 0, m > 0$.

Apriorní hustotě gama rozdělení (t, m) odpovídá aposteriorní hustota gama rozdělení $(t + \sum_{i=1}^n X_i, m+n)$.

Rovnoměrné rozdělení na $(0, \theta)$; $\theta = (0, +\infty)$.

Postačující statistika je $\max_{1 \leq i \leq n} X_i$, tedy $S_n = (0, +\infty)$.

Přirozený konjugovaný systém je systém Paretových rozdělení s parametry (m, t) , $m = 1, 2, \dots; t > 0$.

Obvyklý konjugovaný systém je systém Paretových rozdělení (m, t) , $m > 0; t > 0$.

Apriorní hustotě - Paretova rozdělení (m, t) odpovídá aposteriorní hustota - Paretova rozdělení $(m+n, \max(t, X_1, \dots, X_n))$.

Rovnoměrné rozdělení na $(\theta_1, \theta_2]$: $\Theta = \{(\theta_1, \theta_2)\}; \theta_1 < \theta_2\}$.

Postačující statistiky jsou $(\min_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} X_i)$, tedy $S_n = \{(t_1, t_2)\}; t_1 < t_2\}$.

Přirozený konjugovaný systém je systém dvourozměrných Paretoových rozdělení (t_1, t_2, m) , $t_1 < t_2$, $m = 1, 2, \dots$

Obvyklý konjugovaný systém je systém dvourozměrných Paretoových rozdělení s parametry (r_1, r_2, α) , $r_1 < r_2$, $\alpha > 0$. Je-li apriorní rozdělení Paretovo rozdělení (t_1, t_2, m) , pak aposteriorní rozdělení je opět Paretovo s parametry $(\min(t_1, X_1, \dots, X_n), \max(t_2, X_1, \dots, X_n), m+n)$.

Normální rozdělení $N(\mu, \sigma_0^2)$, $\sigma_0^2 > 0$ známé, $\Theta = R_1$.

Postačující statistika je $\sum_{i=1}^n X_i$, $S_n = R_1$.

Přirozený konjugovaný systém je systém normálních rozdělení $(a, \sigma_0^2/n)$, $a \in R_1$, $n = 1, 2, \dots$

Obvyklý konjugovaný systém je systém rozdělení $N(a, b^2)$, $a \in R_1$, $b^2 > 0$. Je-li apriorní rozdělení $N(a, b^2)$, pak aposteriorní je $N(\mu_1, \sigma_1^2)$, kde

$$\mu_1 = \frac{\sum_{i=1}^n X_i b^2 + a \sigma_0^2}{n b^2 + \sigma_0^2}, \quad \sigma_1^2 = \frac{b^2 \sigma_0^2}{n b^2 + \sigma_0^2}. \quad (2.9)$$

Marginální rozdělení X_i je podle (1.7)

$$\begin{aligned} r(x) &= \int_{-\infty}^{+\infty} (2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{j=1}^n (x_j - \mu)^2\right\} (2\pi b^2)^{-1/2} \exp\left\{-\frac{1}{2b^2}(\mu - a)^2\right\} d\mu \\ &= (2\pi)^{-n/2} (n b^2 + \sigma_0^2)^{-1/2} (\sigma_0^{-2})^{(n-1)/2} \exp\left\{-\frac{1}{2} \left[\sum_{i=1}^n (x_i - \bar{x})^2 / \sigma_0^2 + (2.10) \right. \right. \\ &\quad \left. \left. + n(\bar{x} - a)^2 / (n b^2 + \sigma_0^2) \right] \right\} (x_1, \dots, x_n)' \in R_n. \end{aligned}$$

$$E X_1 = E \mu = a \quad (2.11)$$

$$\text{var } X_1 = \sigma_0^2 + b^2 \quad (2.12)$$

$$\text{cov}(X_i, X_j) = b^2, \quad i \neq j \quad (2.13)$$

Normální rozdělení $N(\mu, \sigma^2)$, μ_0 známé, $\Theta = (0, +\infty)$.

Postačující statistika je $\sum_{i=1}^n (X_i - \mu_0)^2$; $S_n = (0, +\infty)$.

Přirozený konjugovaný systém pro $1/\sigma^2$ je systém gama rozdělení s parametry $(t, m/2)$, $t > 0$, $m = 1, 2, \dots$. Obvykle bereme $t > 0$, $m > 0$. Je-li apriorní hustota gama hustota $(t, m/2)$, pak aposteriorní hustota je též gama s parametry $(t + \sum_{i=1}^n (X_i - \mu_0)^2/2, (m+n)/2)$.

Pozor! Za neznámý parametr bereme $1/\sigma^2$ (nikoli σ^2), neboť systém konjugovaných rozdělení pro $1/\sigma^2$ má jednodušší tvar. $1/\sigma^2$ se někdy nazývá parametr přemnosti.

Normální rozdělení $N(\mu, \sigma^2)$, μ i σ^2 neznámé, $\Theta = R_1 \times (0, +\infty)$.

Postačující statistika je $(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2)$, tedy $S_n = R_1 \times (0, +\infty)$ a

$$g_n(t_1, t_2; \mu, \sigma^2) = \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}(t_2 - 2t_1\mu + n\mu^2)\right\} \quad t_1 \in R_1, \quad t_2 > t_1^2 n^{-1}$$

Odtud a z (2.6) plyne

$$f_{n, t_1, t_2}(\mu, \sigma^2) = \left(\frac{n}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{n}{2\sigma^2}(\mu - t, n^{-1})^2\right\} \quad (2.14)$$

$$\frac{1}{\sigma^{-n+1}} \exp\left\{-\frac{1}{2\sigma^2}(t_2 - n t_1^2)\right\} \left(\Gamma\left(\frac{n+1}{2}\right)\right)^{-1} \left(\frac{t_2 - t_1^2 n}{2}\right)^{\frac{n+1}{2}} \quad \mu \in R_1, \quad \sigma^{-2} > 0$$

Přirozený konjugovaný systém pro parametry $(\mu, 1/\sigma^2)$ je systém rozdělení vymezený následovně: podmíněné rozdělení μ při daném $1/\sigma^2$ je $N(a, \sigma^2 r^{-1})$ a marginální rozdělení $1/\sigma^2$ je gama rozdělení (c, d) , kde $a \in R_1$, $r = 1, 2, \dots$; $2d = 1, 2, \dots$, $c > 0$. Obvykle bereme $a \in R_1$, $r > 0$, $c > 0$, $d > 0$. Takovito rozdělení budeme nazývat normální - gama s parametry (a, r, c, d) . Apriornímu rozdělení popsaného typu odpovídá aposteriorní rozdělení normální - gama s parametry $(\mu^*, r+n, c^*, d^*)$, kde

$$\mu^* = \frac{r\bar{a} + n\bar{X}}{n+r} \quad (2.15)$$

$$c^* = c + \frac{1}{2} \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{rn(\bar{X} - a)^2}{2(r+n)} \quad (2.16)$$

$$d^* = d + n/2$$

(2.17)

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Tedy a priori sdrúžená hustota $(\mu, 1/\sigma^2)$ (t.j. hustota normálního - gamma rozdělení s parametry (a, r, c, d)) je

$$\left(\frac{r}{2\pi\sigma^2}\right)^{1/2} \exp\left\{-\frac{r}{2\sigma^2}(\mu-a)^2\right\} \frac{\sigma^d}{\Gamma(d)} \left(\frac{1}{\sigma^2}\right)^{d-1} \exp\left\{-\frac{c}{\sigma^2}\right\}, \quad (2.18)$$

$$\mu \in R_1, \quad \sigma^{-2} > 0,$$

což implikuje, že marginální hustota μ je až na násobící konstantu rovna

$$\left(1 + \frac{1}{2d}\right) \cdot \frac{dr(\mu-a)^2}{\sigma^2}^{-(2d+1)/2} \quad \mu \in R_1. \quad (2.19)$$

Jinými slovy $(\mu-a)(dr/c)^{1/2}$ má t-rozdělení o $2d$ stupních volnosti (pro $2d$ přirozené). Marginální hustota (x_1, \dots, x_n) je

$$\begin{aligned} r(x_1, \dots, x_n) &= \int_0^{+\infty} \int_{-\infty}^{+\infty} (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\} (2\pi\sigma^2/r)^{-1/2} \\ &\quad \exp\left\{-\frac{r}{2\sigma^2}(\mu-a)^2\right\} \frac{\sigma^d}{\Gamma(d)} \left(\frac{1}{\sigma^2}\right)^{d-1} \exp\left\{-\frac{c}{\sigma^2}\right\} d\sigma^{-2} d\mu = \\ &= \frac{\sigma^d}{\Gamma(d)} (2r)^{-n/2} \left(\frac{r}{n+r}\right)^{1/2} \int_0^{+\infty} (\sigma^{-2})^{\frac{n}{2}+d-1} \exp\left\{-\frac{\sigma^*}{\sigma^2}\right\} d\sigma^{-2} = \\ &= \frac{\Gamma\left(\frac{n}{2}+d\right)}{\Gamma(d)} \left(\frac{r}{n+r}\right)^{1/2} (2r)^{-n/2} \frac{\sigma^d}{\sigma^{d+n/2}}, \quad (x_1, \dots, x_n) \in R_n. \end{aligned} \quad (2.20)$$

Všerobná rozdělení

Multinomiální rozdělení s parametry (s, p) , $p = (p_1, \dots, p_k)$; s je známé,

$$\textcircled{a} = \{p = (p_1, \dots, p_k); 0 < p_i \leq 1, i = 1, \dots, k, \sum_{i=1}^k p_i = 1\}.$$

Postačující statistika je $\left\{ \sum_{j=1}^n X_{1j}, i = 1, \dots, k \right\}$, tedy $S_n =$
 $= \left\{ t_1, \dots, t_k \right\}; t_1 = 0, 1, \dots, sn; i = 1, \dots, k; \sum_{i=1}^k t_i = sn \}$.

Přirozený konjugovaný systém je systém Dirichletových rozdělení s parametry (t_1, \dots, t_k) , $t_1 = 1, \dots, s$, $i = 1, \dots, k$, $s = 1, 2, \dots$. Obvykle používáme systém s $t_1 > 0$, $i = 1, \dots, k$. Je-li apriorní rozdělení Dirichletovo s parametry (t_1, \dots, t_k) , pak aposteriorní rozdělení je také Dirichletovo s parametry $(t_1 + \sum_{j=1}^n X_{1j}, \dots, t_k + \sum_{j=1}^n X_{kj})$.

Vícerozměrné normální rozdělení $N_k(\mu, \Sigma_0)$, kde Σ_0 je známá symetrická pozitivně definitní matice typu $k \times k$, $\otimes = R_k$.

Postačující statistika je $\bar{X} = (\bar{X}_1, \dots, \bar{X}_k)'$, $\bar{X}_1 = \frac{1}{n} \sum_{j=1}^n X_{1j}$. Podle (2.6)

a (A.29)

$$g_n(\bar{x}; \mu) = \exp \left\{ -\frac{n}{2} (\mu - \bar{x})' \Sigma_0^{-1} (\mu - \bar{x}) \right\}$$

$$h_n(\bar{x}) = (2\pi)^{\frac{nk}{2}} (\det \Sigma_0)^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (\bar{X}_i - \bar{X})' \Sigma_0^{-1} (\bar{X}_i - \bar{X}) \right\}$$

Tedy přirozený konjugovaný systém je systém normálních rozdělení $N_k(\hat{a}, \hat{\Sigma}/n)$, $\hat{a} \in R_k$, $n = 1, 2, \dots$. Obvykle bereme systém $N_k(\hat{a}, b \hat{\Sigma})$, $\hat{a} \in R_k$, $b > 0$.

Apriorní hustotě $N_k(\hat{a}, \hat{B})$, B pozitivně definitní symetrická typu $k \times k$, odpovídá aposteriorní hustota $N_k(\hat{a}^*, \hat{B}^*)$, kde

$$\hat{a}^* = (n \hat{\Sigma}_0^{-1} + \hat{B}^{-1})^{-1} (n \hat{\Sigma}_0^{-1} \hat{a} + \hat{B}^{-1} \hat{a}) \quad (2.21)$$

$$\hat{B}^* = (n \hat{\Sigma}_0^{-1} + \hat{B}^{-1})^{-1}. \quad (2.22)$$

Vícerozměrné normální rozdělení $N_k(\mu_0, \Sigma)$, μ_0 je známé, Σ pozitivně definitní symetrická neznámá matice $k \times k$, $\otimes = \{B; \text{symetrická pozitivně definitní matice } k \times k\}$.

Postačující statistika je $\bar{Y} = (Y_{1j})_{1,j=1,\dots,k}$,

$$Y_{1j} = \sum_{q=1}^n (X_{1q} - \mu_{10}) (X_{jq} - \mu_{j0}),$$

$\mu_0 = (\mu_{10}, \dots, \mu_{k0})'$; $S_n = \{B \text{ symetrická pozitivně definitní matice } k \times k\}$. Podle (2.6) a (A.29) a (A.31)

$$g_n(Y; \Sigma^{-1}) = (\det \Sigma^{-1})^{n/2} \exp\left\{-\frac{1}{2} \text{tr}\{\Sigma^{-1} Y\}\right\}$$

$$\int_{\mathcal{C}} g_n(Y; \Sigma^{-1}) d\Sigma^{-1} = (\det Y)^{-(n+k+1)/2} c_{k, n+k+1}^{-1}$$

kde $c_{k, n}$ je dáno (A.32), $\text{tr}\{\cdot\}$ označuje stopu matice. Přírozensý konjugovaný systém pro Σ^{-1} je systém (centrálních) Wishartových rozdělení s r stupni volnosti a parametrickou maticí B , kde $r = k+1, \dots, B \in \mathcal{C}$. Je-li a priori rozdělení Wishartovo s a stupni volnosti a parametrickou maticí R , pak aposteriorní rozdělení je opět Wishartovo s $a+n$ stupni volnosti a parametrickou maticí R' splňující

$$R'^{-1} = R^{-1} + Y.$$

Vícerozměrné normální rozdělení $N_k(\mu, \Sigma)$, $\mu \in \mathbb{R}^k$, $\Sigma > 0$ neznámé, $\mathcal{C} = \{g, B > 0; g \in \mathbb{R}^k, B \text{ symetrická pozitivně definitní matice typu } k \times k\}$.

Postačující statistika je (\bar{X}, S) , kde

$$S = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})',$$

a $S_n = R_k \times \{B; B - \text{symetrická pozitivně definitní matice typu } k \times k\}$. Dle (2.6) a (A.29) platí

$$g_n(\bar{X}, S; \mu, \Sigma) = (\det \Sigma)^{-n/2} \exp\left\{-\frac{n}{2}(\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu)\right\} \cdot \exp\left\{-\text{tr} \Sigma^{-1} S\right\}. \quad (2.24)$$

Odtud plyne, že přírozensý konjugovaný systém pro (μ, Σ^{-1}) můžeme popsat následovně: podmíněné rozdělení μ při daném Σ^{-1} je $N_k(g, r^{-1} \Sigma^{-1})$, marginální rozdělení Σ^{-1} je k -rozměrné Wishartovo rozdělení s q stupni volnosti a parametrickou maticí Y , přičemž $g \in \mathbb{R}^k$, $r=1, 2, \dots$;

\tilde{Y} je symetrická pozitivně definitní matice typu $k \times k$, $q > k-1$. Obvykle pracujeme s $r > 0$ nikoli jen s r přirozenými. Sdružené rozdělení μ a $\tilde{\Sigma}$ má tvar

$$(2\pi)^{-k/2} (\det \tilde{\Sigma})^{-1/2} \exp\left\{-\frac{r}{2} \text{tr}(\tilde{\Sigma}^{-1}(\mu-\tilde{\theta})(\mu-\tilde{\theta})')\right\} r^{k/2}. \quad (2.25)$$

$$\cdot c_{kq} (\det \tilde{Y})^{-q/2} (\det \tilde{\Sigma})^{-(q-k-1)/2} \exp\left\{-\frac{1}{2} \text{tr}(\tilde{Y}^{-1}\tilde{\Sigma}^{-1})\right\} \cdot$$

Odtud lze integrací získat marginální rozdělení μ . Stejně vlastně vypočítat integrál

$$\int (\det \tilde{\Sigma})^{-(q-k)/2} \exp\left\{-\frac{1}{2} \text{tr}(\tilde{\Sigma}^{-1}(r(\mu-\tilde{\theta})(\mu-\tilde{\theta})' + \tilde{Y}^{-1}))\right\} d\tilde{\Sigma}^{-1}$$

a uvědomit si, že funkce pod integrálem je až na násobící konstantu rovna hustotě k -rozměrného Wishartova rozdělení s $(q+1)$ stupni volnosti a perametrickou maticí $(r(\mu-\tilde{\theta})(\mu-\tilde{\theta})' + \tilde{Y}^{-1})^{-1}$. Odtud dostaneme, že marginální hustota μ je až na násobící konstantu rovna

$$(\det(\tilde{Y}^{-1} + r(\mu-\tilde{\theta})(\mu-\tilde{\theta})'))^{-(q+1)/2} =$$

$$=(\det \tilde{Y}^{-1} \cdot (1 + r(\mu-\tilde{\theta})(\mu-\tilde{\theta})'))^{-\frac{q+1}{2}}.$$

Porovnáním s (A.33) je vidět, že μ má k -rozměrné t -rozdělení s $(q-k+1)$ stupni volnosti a parametry $\tilde{\theta}$ a $\tilde{Y}^{-1}r^{-1}(q-k+1)^{-1}$. Apriornímu rozdělení (2.25) odpovídá aposteriorní rozdělení, které lze popsat následovně: podmíněné rozdělení μ při daném $\tilde{\Sigma}$ je $N_k(\mu^*, (r+n)^{-1}\tilde{\Sigma})$, rozdělení $\tilde{\Sigma}^{-1}$ je k -rozměrné Wishartovo s $(q+n)$ -stupni volnosti a perametrickou maticí \tilde{Y}^* , kde

$$\mu^* = \frac{r\bar{\mu} + n\bar{X}}{r + n}, \quad (2.26)$$

$$\tilde{Y}^* = \tilde{Y}^{-1} + \tilde{S} + \frac{r\tilde{M}}{r + n} (\bar{g}-\bar{X})(\bar{g}-\bar{X})'. \quad (2.27)$$

K odvození tohoto výsledku lze použít fakt, že aposteriorní hustota je až na násobící konstantu rovna součinu $g_n(\tilde{Y}, \tilde{\Sigma}, \mu, \tilde{\Sigma})$ dané (2.24)

a pravé strany (A.29).

Je-li $\tilde{I}_k = \sigma^2 I_k$, kde I_k je jednotková matice typu $k \times k$, je obvyklý konjugovaný systém pro (μ, σ^{-2}) tvořen hustotami

$$\exp\left\{-\frac{r}{2\sigma^2}(\mu-\hat{g})'(\mu-\hat{g})\right\} (r(2\pi\sigma^2)^{-1})^{1/2} \cdot (\sigma^{-2})^{d-1} c^d \Gamma(d)^{-1} \exp\{-c\sigma^{-2}\} \quad \mu \in R_k, \sigma^{-2} > 0, \quad (2.28)$$

kde parametry (\hat{g}, r, c, d) probíhají množinu $R_k \times (0, +\infty) \times (0, +\infty) \times (0, +\infty)$, tj. podmíněné rozdělení μ při daném σ^{-2} je $N_k(\hat{g}, \sigma^2 r^{-1} I_k)$ a marginální rozdělení σ^{-2} je gama rozdělení s parametry (c, d) . Marginální rozdělení μ je až na násobící konstantu rovno

$$\left(1 + \frac{1}{2d} \frac{dr(\mu-\hat{g})'(\mu-\hat{g})}{c} - (2d+1)/2\right)^{-1}, \quad \mu \in R_k.$$

Apriornímu rozdělení (2.28) odpovídá aposteriorní rozdělení

$$\exp\left\{-\frac{r+n}{2\sigma^2}(\mu-\mu^*)'(\mu-\mu^*)\right\} ((r+n)(2\pi\sigma^2)^{-1})^{1/2} \cdot (\sigma^{-2})^{d+n/2-1} c^{r+n/2} (\Gamma(d+\frac{n}{2}))^{-1} \exp\{c^r \sigma^{-2}\}, \quad \mu \in R_k, \sigma^{-2} > 0, \quad (2.30)$$

kde μ^* je dáno (2.26)

$$c^{r*} = c + \frac{1}{2} \sum_1 (\bar{X}_1 - \bar{X})'(\bar{X}_1 - \bar{X}) + \frac{rn}{2(r+n)} (\bar{X} - \hat{g})'(\bar{X} - \hat{g}) \quad (2.31)$$

2.3 PRINCIP NEURČITOSTI, JEFFREYSOVA HUSTOTA, LIMITNÍ APOSTERIORNÍ HUSTOTY

Necht $r(\bar{X}|\hat{g})$ je podmíněná hustota náhodného vektoru $\bar{X}=(X_1, \dots, X_n)'$ při dané hodnotě parametru $\hat{g} \in \mathcal{C} \in \mathcal{R}_k$, $\mathcal{C} \neq \emptyset$. Pokládáme-li \hat{g} za náhodný vektor, o němž víme jen to, že $\hat{g} \in \mathcal{C}$, vzniká problém jak volit apriorní rozdělení. Některým řešením tohoto problému se budeme věnovat

v tomto paragrafu.

Podle principu neurčitosti bereme za apriorní rozdělení g rovnoměrné rozdělení na Θ . Příslušnou hustotu budeme značit $q_0(\theta)$ a příslušnou aposteriorní hustotu $\pi_0(\theta|\bar{x})$. Pro Θ nanejvýš spočetnou půjde o hustotu vzhledem k čítecí míře. Je-li Lebesgueova míra Θ kladná, bude q_0 hustota vzhledem k Lebesgueově míře. V obou případech je q_0 rovna identicky kladné konstantě, pro účely výpočtu aposteriorního rozdělení ji vzhledem k (1.4) můžeme klást rovnu 1. Je-li Θ nekonečná spočetná nebo je-li Lebesgueova míra Θ nekonečná, je hustota $q_0(\theta)$ nevlastní.

Je-li např. (X_1, \dots, X_n) náhodný výběr z alternativního rozdělení s parametrem $\theta \in (0, 1)$ a nemáme-li o parametru θ žádná informace, volíme apriorní hustotu (vzhledem k Lebesgueově míře) rovnu

$$q_0(\theta) = 1 \quad \text{pro } \theta \in (0, 1).$$

Příslušná aposteriorní hustota je

$$\pi_0(\theta|\bar{x}) = (B(\sum_1^n x_i + 1, n - \sum_1^n x_i + 1))^{-1} \theta^{\sum_1^n x_i} (1-\theta)^{n - \sum_1^n x_i} \quad \theta \in (0, 1),$$

tj. beta rozdělení s parametry $(\sum_1^n x_i + 1, n - \sum_1^n x_i + 1)$.

Je-li X_1, \dots, X_n náhodný výběr z $N(\mu, \sigma^2)$, kde $\sigma^2 > 0$ je známé a μ je parametr, o němž pouze víme, že $\mu \in R_1$. Pak podle principu neurčitosti vezmeme za apriorní hustotu (vzhledem k Lebesgueově míře)

$$q_0(\mu) = 1 \quad \mu \in R_1$$

a odpovídající aposteriorní hustota je opět $N(\bar{x}, n^{-1})$.

Definujeme si nyní náhodnou veličinu

$$Z = \bar{X} - \mu.$$

Z výše řečeného plyne, že podmíněné rozdělení Z při daném \bar{X} je $N(0, \sigma^2/n)$ a že rovněž podmíněné rozdělení Z při daném μ je

$N(0, \sigma^2/n)$. Tudiž Z a \bar{X} jsou nezávislé náhodné veličiny a podobně Z a μ jsou nezávislé náhodné veličiny. Var Z , var \bar{X} a var μ jsou konečné a tedy

$$\begin{aligned} \text{var } \bar{X} &= \text{var } \mu + \text{var } Z \\ \text{var } \mu &= \text{var } \bar{X} + \text{var } Z, \end{aligned}$$

což implikuje, že var $Z = 0$ a tedy Z je skoro jistě konstanta. Toto je spor s tím, že Z má rozdělení $N(0, \sigma^2/n)$, $\sigma^2 > 0$. V tomto případě tedy není použití principu neurčitosti vhodné, neboť vede k nesmyslným závěrům.

Je vidět, že při použití principu neurčitosti musíme postupovat velmi opatrně. Princip neurčitosti má ještě další nevýhodu, změníme-li totiž parametrizaci modelu (tj. místo parametru θ vezmeme $\eta = \theta^2$), apriorní hustota nového parametru už není rovnoměrná.

Nechť je X_1, \dots, X_n výběr z Poissonova rozdělení s parametrem $\theta \in (0, +\infty)$. Apriorní hustotě

$$\begin{aligned} q_0(\theta) &= 1 & \theta \in (0, +\infty) \\ &= 0 & \theta \notin (0, +\infty) \end{aligned} \quad (2.32)$$

přísluší aposteriorní hustota gama rozdělení s parametry $(n, \sum_{i=1}^n X_i + 1)$.

Definujeme nový parametr $\lambda = \theta^{1/2}$, pak apriorní hustota parametru (odpovídající (2.14)) je

$$\begin{aligned} q(\lambda) &= \lambda & \lambda > 0 \\ &= 0 & \lambda \leq 0. \end{aligned} \quad (2.33)$$

Nedy není již konstantní. Je vidět, že nevíme-li nic o hodnotách θ , máme již jakousi informaci o λ , což je paradoxní závěr. Nevíc aposteriorní hustota $\pi(\lambda | \bar{x})$ příslušná apriorní hustotě (2.33) je

$$\begin{aligned} \pi(\lambda | \bar{x}) &= \frac{\prod_{i=1}^n X_i + 1}{\Gamma(\sum_{i=1}^n X_i + 1)} e^{-\lambda^2 n} \lambda^{2 \sum_{i=1}^n X_i + 1} & \lambda > 0 \\ &= 0 & \lambda \leq 0 \end{aligned}$$

Zatímco apriorní hustotě

$$q^*(\lambda) = 1 \quad \lambda > 0$$

$$= 0 \quad \lambda \leq 0$$

přísluší aposteriorní hustota

$$\pi^*(\lambda | \bar{x}) = \frac{\sum_{i=1}^n x_i + \frac{1}{2}}{\Gamma(\frac{1}{2}x_i + \frac{1}{2})} e^{-\lambda^2 n} \lambda^{\sum_{i=1}^n x_i} \quad \lambda > 0$$

$$= 0 \quad \lambda \leq 0.$$

Pedy aposteriorní hustoty π a π^* jsou různé, ačkoli oba postupy, kterými jsme k nim dospěli, jsou z hlediska logického rovnocenné.

Tyto úvahy vedly k závěru, že místo podle principu neurčitosti bychom měli volit apriorní rozdělení takové, aby nezáviselo na počáteční parametrizaci modelu. Následující věta nám dává řešení pro případ, že λ je Lebesgueova míra a $\lambda(\Theta) > 0$.

Nejprve si však připomeneme pojmy regulárního systému hustot a Fisherovy informační matice.

Řekneme, že systém hustot $\{r(x|\theta), \theta \in \Theta\}$ je regulární, jsou-li splněny tyto podmínky:

- Θ je neprázdná otevřená množina intervalů v R_k .
- Množina $M = \{\bar{x}; r(\bar{x}|\theta) > 0\}$ nezávisí na θ .
- Pro skoro všechna $\bar{x} \in M$ (vzhledem k σ -konečné míře ν_n) existuje konečná parciální derivace $r_{i1}'(\bar{x}|\theta) = \frac{\partial r(\bar{x}|\theta)}{\partial \theta_i}$ $i=1, \dots, k$.
- Pro každé i a pro vš. $\theta \in \Theta$ platí $\int_M r_{i1}'(\bar{x}|\theta) d\nu_n(\bar{x}) = 0$.
- Pro každou dvojici (i, j) existuje konečný integrál

$$J_{ij}(\theta) = \int_M \frac{r_{i1}'(\bar{x}|\theta) r_{j1}'(\bar{x}|\theta)}{r^2(\bar{x}|\theta)} r(\bar{x}|\theta) d\nu_n(\bar{x}).$$

$f)$ Matice $J(\theta) = (J_{ij}(\theta))_{i,j=1, \dots, k}$ je pozitivně definitní pro každé $\theta \in \Theta$.

Matici $J(\theta)$ nazýváme Fisherovou informační maticí.

Věta 2.3 (Jeffreysova). Necht náhodný vektor $\tilde{X} = (X_1, \dots, X_n)'$ má při daném parametru $\theta \in \Theta \in \mathcal{B}_k$ hustotu $r(\tilde{x}|\theta)$ vzhledem k nějaké σ -konečné míře ν_n . Předpokládejme, že systém hustot $\{r(\tilde{x}|\theta); \theta \in \Theta\}$ je regulární a má Fisherovu informační matici $J(\theta)$. Položme

$$c = \int_{\Theta} r(\tilde{x}|\theta) (\det J(\theta))^{1/2} d\theta \quad (2.34)$$

a předpokládejme, že $0 < c < +\infty$.

Budiž H regulární prosté zobrazení množiny Θ na $\Theta^* \in \mathcal{B}_k$. Označme $\eta = H(\theta)$ a $r^*(\tilde{x}|\eta) = r(\tilde{x}|H^{-1}(\eta))$. Pak $\{r^*(\tilde{x}|\eta); \eta \in \Theta^*\}$ je regulární systém hustot. Označíme-li $J^*(\eta)$ Fisherovu informační matici, pak pro libovolnou množinu B splňující podmínky $B \in \Theta$, $B \in \mathcal{B}_k$ platí

$$\begin{aligned} \int_B c r(\tilde{x}|\theta) (\det J(\theta))^{1/2} d\theta &= \\ &= \int_{H(B)} c r^*(\tilde{x}|\eta) (\det J^*(\eta))^{1/2} d\eta. \end{aligned} \quad (2.35)$$

Důkaz. Regularita systému $\{r^*(\tilde{x}|\eta), \eta \in \Theta^*\}$ je zřejmá až na to, že se musí dokázat pozitivní definitnost matice $J^*(\eta)$. Necht $\theta = (\theta_1, \dots, \theta_k)'$, $\eta = (\eta_1, \dots, \eta_k)'$. Ze vztahu

$$\frac{\partial \ln r(\tilde{x}|\theta)}{\partial \theta_i} = \frac{\partial \ln r^*(\tilde{x}|\eta)}{\partial \theta_i} = \sum_{j=1}^k \frac{\partial \ln r^*(\tilde{x}|\eta)}{\partial \eta_j} \cdot \frac{\partial \eta_j}{\partial \theta_i}$$

dostaneme pro prvky $J_{ij}(\theta)$ a $J_{ij}^*(\eta)$ informačních matic $J(\theta)$ a $J^*(\eta)$ rovnost

$$J_{ij}(\theta) = E \left(\frac{\partial \ln r(\tilde{x}|\theta)}{\partial \theta_i} \cdot \frac{\partial \ln r(\tilde{x}|\theta)}{\partial \theta_j} \Big| \theta \right) = \sum_{u=1}^k \sum_{v=1}^k \frac{\partial \eta_u}{\partial \theta_i} J_{uv}^*(\eta) \frac{\partial \eta_v}{\partial \theta_j}. \quad (2.36)$$

Označíme-li

$$D = \left(\frac{\partial \eta_u}{\partial \theta_i} \right)_{i,u=1,\dots,k}$$

můžeme (2.36) napsat pomocí matic jako

$$J(\mathfrak{g}) = \mathbb{D}' J^*(\eta) \mathbb{D}.$$

Přitom \mathbb{D} je regulární matice, neboť \mathbb{D} je jakobián zobrazení H a toto zobrazení je podle předpokladu regulární. Z pozitivní definitnosti matice $J(\mathfrak{g})$ plyne i pozitivní definitnost matice $J^*(\eta)$ pro $\forall \eta \in \Theta^*$.

Podle věty o substituci v mnohonásobných integrálech platí

$$\begin{aligned} & \int_B \circ r(\bar{x}|\mathfrak{g}) (\det J(\mathfrak{g}))^{1/2} d\mathfrak{g} = \\ & = \int_{H(B)} \circ r(\bar{x}|H^{-1}(\eta)) (\det(\mathbb{D}' J^*(\eta) \mathbb{D}))^{1/2} (\det \mathbb{D})^{-1} d\eta. \end{aligned}$$

Tvrzení věty nyní plyne, použijeme-li vztah

$$\det(\mathbb{D}' J^*(\eta) \mathbb{D})^{1/2} (\det \mathbb{D})^{-1} = (\det(J^*(\eta)))^{1/2}.$$

Q.E.D.

Je vidět, že za předpokladů věty je apriorní hustota parametru \mathfrak{g} rovna funkci $(\det J(\mathfrak{g}))^{1/2}$ (nebo jakémukoliv kladnému násobku této funkce) a aposteriorní hustota parametru \mathfrak{g} je rovna $c r(\bar{x}|\mathfrak{g})$.

$\cdot (\det J(\mathfrak{g}))^{1/2}$ a že je to pravděpodobnostní hustota. Funkci $k \cdot (\det J(\mathfrak{g}))^{1/2}$ kde k je libovolné kladné číslo, budeme říkat Jeffreysova apriorní hustota. Z tvrzení věty plyne, že při Jeffreysové volbě apriorní hustoty parametrů \mathfrak{g} a η jsou obě aposteriorní pravděpodobnosti stejné a nemůže dojít k paradoxnímu výsledku jako u principu neurčitosti. Někdy je Jeffreysova hustota nevlastní.

Uvedeme si Jeffreysovy hustoty pro některé případy.

Pro binomické rozdělení s parametry (m, p) , m je známé, je Jeffreysova hustota beta hustota $(1/2, 1/2)$. Aposteriorní hustota je beta $(\frac{1}{2} + \sum_{i=1}^n x_i, \frac{1}{2} + m - \sum_{i=1}^n x_i)$.

Pro Poissonovo rozdělení s parametrem λ má Jeffreysova hustota tvar

$$q(\lambda) = \lambda^{-1/2} \quad \lambda > 0, \\ = 0 \quad \lambda \leq 0.$$

Je tedy nevládní a aposteriorní rozdělení je gama s parametry $(n, \sum_{i=1}^n x_i + 1/2)$.

Pro negativně binomické rozdělení s parametry (a, p) , a - známé, je Jeffreysova hustota

$$q(p) = p^{-1} (1-p)^{-1/2} \quad p \in (0, 1) \\ = 0 \quad p \notin (0, 1).$$

Jde o nevládní hustotu a aposteriorní hustota je beta s parametry $(an, \sum_{i=1}^n x_i + 1/2)$.

Pro normální rozdělení $N(\mu, \sigma_0^2)$, $\sigma_0^2 > 0$, je Jeffreysova hustota konstantní (vady této hustoty jsou v první části tohoto paragrafu).

Pro normální rozdělení $N(\mu_0, \sigma^2)$, μ_0 známé, je Jeffreysova hustota rovna σ^{-2} pro $\sigma^{-2} > 0$. Jde o nevládní hustotu. Aposteriorní hustota σ^{-2} je gama rozdělení $(\sum_{i=1}^n (x_i - \mu_0)^2 / 2, n/2)$.

Pro normální rozdělení $N(\mu, \sigma^2)$, μ i $\sigma^2 > 0$ neznámé je Jeffreysova hustota dána vzorcem

$$q(\mu, \sigma^{-2}) = \sigma^{-2}, \quad \mu \in R_1, \quad \sigma^{-2} > 0.$$

Opět jde o nevládní hustotu a aposteriorní hustota (μ, σ^{-2}) se dá popsat následovně: podmíněné rozdělení μ při daném $1/\sigma^2$ je $N(\bar{x}, \sigma^2/n)$ a marginální rozdělení $1/\sigma^2$ je gama s parametry $(\frac{1}{2} \sum_{i=1}^n (x_i - \bar{x})^2, n/2)$.

Další možná volba apriorní hustoty parametru θ je vyjít z konjugovaného systému hustot (většinou tzv. obvyklého) $\{q(\theta; \lambda); \lambda \in \Lambda\}$ příslušného $\{r(\underline{x}|\theta); \theta \in \Theta\}$ (Λ je většinou otevřená borelovská množina a volíme přímo aposteriorní hustotu $r^*(\theta|\underline{x})$, kterou dostaneme jako limitu aposteriorních hustot z $\{q(\theta; \lambda); \lambda \in \Lambda\}$ pro λ konvergující k nějakému bodu na hranici množiny Λ . Formálně odpovídá tento způsob tomu, že za apriorní hustotu $q^*(\theta)$ vezmeme limitu apriorních hustot z $\{q(\theta; \lambda), \lambda \in \Lambda\}$ pro λ konvergující k nějakému bodu na hranici množiny Λ . Hustota $q^*(\theta)$ je obvykle nevlastní. V řadě případů jsou potom bayesovské odhady a bayesovské testy shodné s klasickými. Aposteriorní hustoty získané právě popsaným způsobem budeme nazývat limitní aposteriorní hustoty.

Uvedeme si několik příkladů:

Pro binomické rozdělení (n, p) je konjugovaný systém systémem beta rozdělení (α, β) , $\alpha > 0$, $\beta > 0$. Limita aposteriorního rozdělení pro $\alpha \rightarrow 0$ a $\beta \rightarrow 0$ je beta rozdělení $(\sum_{i=1}^n x_i, m - \sum_{i=1}^n x_i)$, což odpovídá apriornímu rozdělení

$$q^*(p) = (p(1-p))^{-1} \quad p \in (0, 1)$$

$$= 0 \quad p \notin (0, 1).$$

Pro Poissonovo rozdělení s parametrem θ je systém konjugovaných hustot tvořen systémem gama rozdělení (a, t) , $a > 0$, $t > 0$. Limita aposteriorního rozdělení pro $a \rightarrow 0$, $t \rightarrow 0$ je gama rozdělení $(n, \sum_{i=1}^n x_i)$. Proto též aposteriorní rozdělení dostaneme, jestliže za apriorní rozdělení zvolíme

$$q^*(\theta) = \theta^{-1} \quad \theta > 0$$

$$= 0 \quad \theta \leq 0.$$

Stejnou úvahou dospějeme pro negativně binomické rozdělení (s, p) k tomu, že za aposteriorní rozdělení $r^*(p|\underline{x})$ vezmeme beta rozdělení $(sn, \sum_{i=1}^n x_i)$, které odpovídá apriorní hustotě

$$q^*(p) = (p(1-p))^{-1} \quad p \in (0,1) \\ = 0 \quad p \notin (0,1).$$

Pro exponenciální rozdělení s parametrem θ vezmeme za aposteriorní hustotu $\pi^*(\theta|\bar{x})$ gama hustotu $(\sum_1^n x_i, n)$, které odpovídá apriorní hustota

$$q^*(\theta) = \theta^{-1} \quad \theta > 0 \\ = 0 \quad \theta \leq 0.$$

Pro normální rozdělení $N(\mu, \sigma_0^2)$, $\sigma_0^2 > 0$ známé, vezmeme za aposteriorní hustotu $N(\bar{x}, \sigma_0^2/n)$, které odpovídá apriorní hustotě rovnoměrné na R_1 . (U apriorního rozdělení $N(a, b^2)$ klademe $b^2 = \infty$.)

Pro normální rozdělení $N(\mu_0, \sigma^2)$, μ_0 známé, vezmeme za aposteriorní hustotu gama s parametry $(\sum_1^n (x_i - \mu_0)^2/2, n/2)$, které odpovídá apriorní hustotě

$$q^*(\sigma^{-2}) = \sigma^{2^2} \quad \sigma^{-2} > 0 \\ = 0 \quad \sigma^{-2} \leq 0.$$

Pro normální rozdělení $N(\mu, \sigma^2)$, μ i $\sigma^2 > 0$ neznámé, bereme v aposteriorní hustotě $r \rightarrow 0$, $\sigma \rightarrow 0$, $d \rightarrow -\frac{1}{2}$. Pak podmíněné aposteriorní rozdělení μ při daném $1/\sigma^2$ je $N(\bar{x}, \sigma^2/n)$ a marginální aposteriorní rozdělení $1/\sigma^2$ je gama s parametry $(\sum_1^n (x_i - \bar{x})^2/2, (n-1)/2)$. Toto aposteriorní rozdělení odpovídá apriornímu

$$q^*(\mu, 1/\sigma^2) = (\sigma^2)^{-3/2} \quad \mu \in R_1, \quad 1/\sigma^2 > 0. \quad (2.37)$$

Marginální aposteriorní rozdělení (tj. při daném x) náhodné veličiny

$$\frac{\mu - \bar{x}}{\sqrt{S_n}} \sqrt{n},$$

kde

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (2.38)$$

je t -rozdělení s $(n-1)$ stupni volnosti.

Obdobné výsledky platí i pro vícerozměrné normální rozdělení.

2.4 EMPIRICKÉ BAYESOVSKÉ METODY

Jedná se o metody volby apriorního rozdělení $q(\theta)$, jestliže máme k dispozici výsledky z minulosti v následujícím tvaru. Y_1, \dots, Y_N jsou nezávislé náhodné veličiny, Y_1 má podmíněnou hustotu $r(Y|\theta_1)$ (vzhledem k σ -konečné míře ν), $\theta_1 \in \Theta$, $l = 1, \dots, N$, $\theta_1, \dots, \theta_N$ jsou nezávislé náhodné vektory, θ_1 má hustotu $q(\theta)$ (vzhledem k σ -konečné míře λ). Ředy nepodmíněné rozdělení Y_1 je

$$r_q(Y) = \int_{\Theta} r(Y|\theta) q(\theta) d\lambda(\theta) \quad (2.39)$$

a odtud plyne pro nepodmíněnou střední hodnotu a rozptyl

$$EY_1 = E(E(Y_1|\theta)) \quad (2.40)$$

$$\text{var } Y_1 = E(\text{var}(Y_1|\theta)) + E(E(Y_1|\theta) - EY_1)^2. \quad (2.41)$$

za předpokladu konečnosti EY_1 resp. $\text{var } Y_1$.

Na základě Y_1, \dots, Y_N můžeme odhadnout hustotu $r_q(Y)$ popř. příslušnou distribuční funkci některou běžnou metodou. Označíme-li $\hat{r}(Y)$ odhad hustoty $r(Y)$ a dosadíme-li do (2.39) dostáváme funkcionální rovnice, které lze obecně jen velmi těžko řešit. Místo řešení funkcionální rovnice někdy minimalizujeme vzdálenost \hat{r} a r_q vzhledem k q a za odhad \hat{q} vezmeme hustotu q , pro kterou je dosaženo minimum.

Za vzdálenost nejčastěji volíme

$$\int \hat{r}(Y) \log \left(\frac{\hat{r}_q(Y)}{\hat{r}(Y)} \right) d\nu(Y).$$

V obecném případě je řešení složité. Řeší se jen ve speciálních případech. Další informace o tomto postupu lze nalézt v [2] a [7].

Spíše než najít odhad popsanou obecnou metodou se nejdříve podíváme, zda variabilita náhodné veličiny Y_1 (popsané $r(y)$), která vznikne složením variability θ (popsané $q(\theta)$) a variability Y_1 při pevném θ_1 (popsané $r(y|\theta)$) je způsobená hlavně variabilitou θ_1 . V kladném případě můžeme za odhad $q(\theta)$ v řadě případů vzít odhad $\hat{r}(y)$. Variabilitou často míníme rozptyl. Je-li v tomto případě $\text{var}(Y_1|\theta)$ malá ve srovnání s $\text{var } \theta$, pak bereme za odhad $q(\theta)$ funkci $r(y)$ (pro θ jednorozměrné).

Oba vyložené postupy nemají příliš velkou naději na použití v praxi, neboť v prvním případě je obtížné najít řešení a druhá metoda je příliš intuitivní. Používají se zatím jen v určitých velmi speciálních případech viz např. [7].

Mnohem větší naději na úspěch má metoda, kterou si nyní vyložíme. Předpokládáme, že $q(\theta)$ má určitý funkcionální tvar, u kterého známe jen parametry, tj. předpokládáme, že a priori hustota je $q(\theta; \underline{\alpha})$, kde $\underline{\alpha} = (\alpha_1, \dots, \alpha_s)'$ je vektor konstant, který neznáme, víme jen, že $\underline{\alpha} \in \Lambda$, kde $\emptyset \neq \Lambda \in \mathcal{B}_s$. Tedy také marginální hustotu Y_1 známe až na vektor $\underline{\alpha}$, který můžeme odhadnout pomocí Y_1, \dots, Y_n některou klasickou metodou odhadu např. metodou momentů (viz [1]). Při ní klademe

$$EY_1^j = \frac{1}{N} \sum_{j=1}^N Y_1^j \quad j = 1, 2, \dots, s$$

(za předpokladu konečnosti příslušných momentů) a řešíme vzhledem k $\underline{\alpha}$. Řešení označíme $\hat{\underline{\alpha}}$. Za a priori hustotu vezmeme $q(\theta; \hat{\underline{\alpha}})$.

Tato metoda má široké použití. Dá se vhodně kombinovat s konjugovanými rozděleními, které poskytují jen funkcionální tvar rozdělení. Za funkcionální tvar vezmeme rozdělení konjugované s $r(y|\theta)$ a parametry odhadneme podle výše popsané metody. Ve většině rozdělení uvažovaných v 2.2 vede tato kombinace metod k rozumným výsledkům.

Empirické bayesovské metody se používají spíše v úlohách odhadu než při testování hypotéz. V kapitole 4 jsou uvedeny 2 příklady odhadu parametru, jestliže apriorní hustota byla získána empirickou bayesovskou metodou.

Nechť Y_1, \dots, Y_N jsou nezávislé náhodné veličiny, Y_1 s rozdělením $N(\theta, \sigma_0^2)$ a představují výsledky z minulosti. $\theta_1, \dots, \theta_N$ jsou nezávislé náhodné veličiny. θ_1 má rozdělení $N(\mu_q, \sigma_q^2)$, kde μ_q a $\sigma_q^2 > 0$ neznáme. Nepodmíněné rozdělení Y_1 je $N(\mu_q, \sigma_q^2)$, kde $\sigma_q^2 = \sigma_q^2 + \sigma_0^2$. Metodou momentů získáme odhady $\hat{\mu}_q$ a $\hat{\sigma}_q^2$ pro μ_q a σ_q^2 , a to

$$\hat{\mu}_q = \frac{1}{N} \sum_{i=1}^N Y_i = \bar{Y} \quad (2.39)$$

$$\hat{\sigma}_q^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2. \quad (2.40)$$

Tedy za apriorní rozdělení vezmeme $N(\hat{\mu}_q, \hat{\sigma}_q^2)$, kde

$$\hat{\sigma}_q^2 = \hat{\sigma}_q^2 - \sigma_0^2 \quad \text{je-li } \hat{\sigma}_q^2 > \sigma_0^2 \quad (2.41)$$

$$= 0 \quad \text{je-li } \hat{\sigma}_q^2 \leq \sigma_0^2$$

(odhad $\hat{\sigma}_q^2$ musí být nezáporný).

3. STATISTICKÉ ROZHODOVACÍ FUNKCE

3.1 FORMULACE PROBLÉMU

V této kapitole se budeme zabývat tzv. statistickými rozhodovacími úlohami, které zahrnují jako speciální případy úlohu odhadu a testování hypotéz. Nejprve si uvedeme nezbytné značení a definice.

Nechť $\mathbf{X} = (X_1, \dots, X_n)'$ je náhodný vektor s hustotou $r(\mathbf{x}|\theta)$ vzhledem k σ -konečné míře γ_n na (R_n, \mathcal{A}_n) , $\theta \in \Theta$ je parametr, Θ je neprázd-
ná borelovská podmnožina R_k . Označme \mathcal{D} množinu možných rozhodnutí
(závěrů) o parametru θ a d prvek množiny \mathcal{D} .

Dále označme $L(\theta, d)$ ztrátovou funkci, která nám udává číselně jakou ztrátu utrpíme, jestliže skutečná hodnota parametru je θ a přijmeme rozhodnutí d . Ztrátová funkce je tedy zobrazení z $\Theta \times \mathcal{D}$ do R_1 . Budeme předpokládat, že existuje k konečné takové, že $L(\theta, d) \geq k$ pro vš. $\theta \in \Theta$ a vš. $d \in \mathcal{D}$. Pokud budeme uvažovat σ -algebry podmnožin \mathcal{A} a \mathcal{D} , budeme předpokládat, že L je měřitelná funkce vzhledem k součinnové σ -algebře, přičemž na R_1 bereme σ -algebru borelovských množin \mathcal{B}_1 .

Dále budeme pracovat s rozhodovacími funkcemi d , kterými budeme rozumět zobrazení z R_n do \mathcal{D} . Jinak řečeno hodnota $d(\mathbf{x})$, $\mathbf{x} \in R_n$ představuje rozhodnutí (závěr) o parametru θ , je-li $\mathbf{x} = \mathbf{X}$.

Riziko $R(\theta, d)$ příslušné rozhodovací funkci d , je-li skutečná hodnota parametru θ , definujeme následovně

$$R(\theta, d) = R(L(\theta, d(\mathbf{X}))) \mathbb{1}(\theta) = \int L(\theta, d(\mathbf{x})) r(\mathbf{x}|\theta) d\mu(\mathbf{x}). \quad (3.1)$$

Značení $R(\cdot|\theta)$ jsme použili, přestože v právě uvažované situaci je θ vektor konstant. Vzorec však budeme používat i při bayesovském pří-
stupu, kde θ považujeme za náhodný vektor a použítí nepodmíněné střed-
ní hodnoty by mohlo vést k nejasnostem. Riziko $R(\theta, d)$ představuje
střední ztrátu způsobenou volbou rozhodovací funkce d , je-li skuteč-

ná hodnota parametru θ .

V dalším se omezíme jen na rozhodovací funkce d , pro něž je $R(\theta, d) < +\infty$ pro vš. $\theta \in \Theta$. Množinu takovýchto rozhodovacích funkcí označme Δ .

Punkci $R: \Theta \times \Delta \rightarrow R_1$ nazýváme rizikovou funkcí.

Statistickým rozhodovacím problémem budeme rozumět trojici

(Θ, Δ, R) .

Nechť $\bar{X} = (X_1, \dots, X_n)'$ je náhodný výběr z alternativního rozdělení s parametrem θ . Uvažujme úlohu odhadu parametru θ . Zřejmě $\Theta = \langle 0, 1 \rangle$, množina \mathcal{D} možných rozhodnutí je množina hodnot ^{odhadů} parametru θ a Δ je množina odhadů parametru θ . Měříme-li ztrátu jako čtverec rozdílu parametru θ a jeho odhadu $d(\bar{X})$, můžeme psát

$$L_1(\theta, d(\bar{X})) = (\theta - d(\bar{X}))^2$$

a tedy pro rizikovou funkci máme

$$R_1(\theta, d) = E((\theta - d(\bar{X}))^2 | \theta).$$

Pro úlohu testu hypotézy $H_0: \theta \in \Theta_0$ proti alternativě $H_1: \theta \in \Theta_1$, $\Theta_0 \cap \Theta_1 = \emptyset$, $\Theta_0 \cup \Theta_1 = \langle 0, 1 \rangle$ je $\mathcal{D} = \{d_0, d_1\}$, kde $d_1 = \{H_1 \text{ platí}\}$.

Obvykle předpokládáme, že při správném rozhodnutí je ztráta nulová a při nesprávném rozhodnutí je ztráta rovna nějaké kladné konstantě a . Pak můžeme pro ztrátovou funkci psát:

$$\begin{aligned} L_2(\theta, d_1) &= 0 & \theta \in \Theta_1 & \quad i = 0, 1 \\ &= a & \theta \notin \Theta_1 \end{aligned} \quad (3.2)$$

a pro rizikovou funkci máme:

$$\begin{aligned} R_2(\theta, d) &= a E(I\{d(\bar{X}) = d_1\}; \theta) = a \cdot P(d(\bar{X}) = d_1; \theta) & \theta \in \Theta_0 \\ &= a E(I\{d(\bar{X}) = d_0\}; \theta) = a \cdot P(d(\bar{X}) = d_0; \theta) & \theta \in \Theta_1, \end{aligned} \quad (3.3)$$

kde $I\{A\}$ je indikátor množiny A .

Odtud je vidět, že riziková funkce je a násobek pravděpodobnosti chybného rozhodnutí a množina $\{\bar{x}; \delta(\bar{x}) = d_1\}$ je oblast přijetí H_1 příslušná rozhodovací funkci δ .

Na množině Δ rozhodovacích funkcí zavedeme uspořádání a ekvivalenci. Rozhodovací funkce δ_1 je R-lepší (popř. lepší vzhledem k rizikové funkci R) než rozhodovací funkce δ_2 , jestliže

$$R(\theta, \delta_1) \leq R(\theta, \delta_2) \quad \text{pro vš. } \theta \in \Theta.$$

a jestliže ostrá nerovnost platí pro aspoň jedno θ .

Rozhodovací funkce δ_1 je R-ekvivalentní rozhodovací funkci δ_2 , jestliže

$$R(\theta, \delta_1) = R(\theta, \delta_2) \quad \text{pro vš. } \theta \in \Theta.$$

Přirozenou optimální rozhodovací funkcí by byla funkce $\delta^* \in \Delta$, taková, že

$$\min_{\delta \in \Delta} R(\theta, \delta) = R(\theta, \delta^*) \quad \text{pro vš. } \theta \in \Theta.$$

Taková rozhodovací funkce existuje jen výjimečně. Proto definujeme optimalitu v nějakém slabším smyslu.

Při bayesovském přístupu předpokládáme, že parametr θ je náhodný vektor s hustotou $q(\theta)$ vzhledem k σ -konečné míře λ na $(\mathcal{Q}, \mathcal{B}(\mathcal{Q}))$ a definujeme bayesovskou rizikovou funkci následovně

$$\begin{aligned} \rho(q, \delta) &= E R(\theta, \delta) = \\ &= \int_{\Theta} \left(\int_{\mathcal{X}} L(\theta, \delta(\bar{x})) r(\bar{x} | \theta) d\nu(\bar{x}) \right) q(\theta) d\lambda(\theta). \end{aligned} \quad (3.4)$$

$\delta^* \in \Delta$. Optimální (bayesovská) rozhodovací funkce $\delta^*(q)$ je definována následovně

$$\min_{\delta \in \Delta} \rho(q, \delta) = \rho(q, \delta^*). \quad (3.5)$$

Číslo $\rho^*(q) = \rho(q, \delta^*)$ nazýváme bayesovské riziko.

Další řešení může poskytnout tzv. invarianční princip.

Nechť je systém hustot $\{r(\underline{x}|\vartheta); \vartheta \in \Theta\}$ invariantní vůči grupě transformací g , tj. pro každé $g \in \mathcal{G}$ a $\vartheta \in \Theta$ existuje jediné řešení $\vartheta^0 \in \Theta$ takové, že $\underline{y} = g(\underline{x})$ má hustotu $r(\underline{y}|\vartheta^0)$ (\underline{x} má hustotu $r(\underline{x}|\vartheta)$). Označme

$$\vartheta^0 = g(\vartheta).$$

Pak je rozumné uvažovat ztrátové funkce invariantní vůči g , tj. takové ztrátové funkce $L(g, d)$, že pro každé $g \in \mathcal{G}$ a $d \in \mathcal{D}$ existuje $d^0 \in \mathcal{D}$ takové, že

$$L(g, d) = L(g(\vartheta), d^0) \quad \text{pro vš. } \vartheta \in \Theta.$$

Označme

$$d^0 = g(d).$$

Rozhodovací funkci $d(\underline{x})$ nazveme invariantní vůči grupě g , jestliže pro vš. $\underline{x} \in R_{\mathcal{U}}$ a $g \in \mathcal{G}$ platí

$$d(g(\underline{x})) = g(d(\underline{x})).$$

Označme Δ_I množinu takovýchto rozhodovacích funkcí.

Pro rizikovou funkci můžeme dokázat následující tvrzení:

Věta 3.1. Pro rizikovou funkci $R(g, d')$ odpovídající invariantní ztrátové funkci L (vůči grupě g) platí

$$R(g, d') = R(g(\vartheta), d') \quad \text{pro vš. } \vartheta \in \Theta, \text{ vš. } g \in \mathcal{G} \\ \text{a vš. } d' \in \Delta_I.$$

Důkaz. Použijeme-li postupně invariance ztrátové funkce, invariance rozhodovací funkce a invariance rozdělení $\{r(\underline{x}|\vartheta); \vartheta \in \Theta\}$, obdržíme

$$R(g, d') = E(L(g, d'(\underline{x}))/\vartheta) = \\ = E(L(g(\vartheta), g(d'(\underline{x}))/\vartheta)) = \\ = E(L(g(\vartheta), d'(g(\underline{x}))/\vartheta)) = \\ = E(L(g(\vartheta), d'(\underline{x}))/g(\vartheta)) = \\ = R(g(\vartheta), d'(\underline{x})).$$

Q.E.D.

Statistický rozhodovací problém (Θ, Δ, R) nazveme invariantní vůči skupě \mathcal{G} , jestliže $\Delta = \Delta_I$ a R odpovídá invariantní ztrátové funkci.

Optimální bayesovskou rozhodovací funkcí invariantní vůči \mathcal{G} budeme rozumět rozhodovací funkci d^* takovou, že

$$\min_{d \in \Delta_I} \varrho(q, d^*) = \varrho(q, d^*), \quad (3.6)$$

kde ϱ je definováno (3.4) a odpovídá invariantní ztrátové funkci.

V řadě případů vede tento postup k tomu, že riziková funkce $R(\theta, d^*)$ nezávisí na θ a optimální bayesovská rozhodovací funkce d^* invariantní vůči \mathcal{G} splňuje

$$\min_{d \in \Delta_I} R(\theta, d^*) = R(\theta, d^*),$$

a tedy nezávisí na q .

Ukážeme si to na příkladě. Systém hustot

$$\left\{ (2\pi)^{-n/2} \exp\left\{-\frac{1}{2} \sum_{i=1}^n (x_i - \theta)^2\right\}; \theta \in R_1 \right\}$$

odpovídající výběru o rozsahu n a z rozdělení $N(\theta, 1)$ je invariantní vůči posunutí, tj. \mathcal{G} je tvořena transformacemi

$$g(x) = (x_1 - c, \dots, x_n - c)', \quad c \in R_1, \quad \mathbf{x} = (x_1, \dots, x_n)' \in R_n.$$

Zřejmě $\bar{g}(\theta) = \theta - c$, $\theta \in R_1$, $c \in R_1$ a budeme uvažovat ztrátové funkce

L s vlastností: pro každé $d \in \mathcal{D}$ existuje $d_0 \in \mathcal{D}$ takové, že

$$L(\theta, d) = L(\theta - c, d_0) \quad \text{pro vš. } c \in R_1, \theta \in \Theta.$$

Tohoto je například splněno pro $\mathcal{D} \subset R_1$ a

$$L(\theta, d) = |\theta - d|^2. \quad (3.7)$$

Pak $\bar{g}(d) = d - c$ a množina Δ_I rozhodovacích funkcí invariantních vůči

\mathcal{G} je tvořena funkcemi $d^*(\mathbf{x})$ splňujícími

$$d^*(x_1 - c, \dots, x_n - c) = d^*(x_1, \dots, x_n) - c \quad \text{pro vš. } c \in R_1 \text{ a } \mathbf{x} \in R_n.$$

Nedy věty 3.1 implikuje, že riziková funkce nezávisí na θ a je rovna

$$R(\theta, \delta^0) = R(0, \delta^0) = (2\pi)^{-n/2} \int_{R^n} \dots \int (\delta^0(\underline{x}))^2 \exp\left\{-\frac{1}{2} \sum_{i=1}^n x_i^2\right\} dx_1 \dots dx_n.$$

V dalším se budeme zabývat pouze bayesovským přístupem.

Jiné řešení rozhodovacího problému poskytuje např. Princip minimu maxu, při kterém za optimální rozhodovací funkci budeme považovat rozhodovací funkci δ^0 s vlastností

$$\min_{\delta \in \mathcal{D}} \max_{g \in \mathcal{G}} R(g, \delta) = \max_{g \in \mathcal{G}} R(g, \delta^0).$$

Podrobnější informace o tomto přístupu lze nalézt např. v [3], [2].

3.2 BAYESOVSKÉ ROZHODOVACÍ FUNKCE

Bayesovská rozhodovací funkce δ^* definovaná (3.5) závisí na volbě ztrátové funkce a volbě rozdělení parametru \underline{g} . Volbě rozdělení parametru \underline{g} byla věnována druhá kapitola. Pokud se týče ztrátové funkce, budeme vycházet z předpokladu, že je dána. Nejběžnější ztrátové funkce používané v úlohách odhadu a testování hypotéz jsou uvedeny v následující částech dvou kapitolách. Existuje obecná metoda volby ztrátové funkce na základě preferencí (podrobněji viz např. [2], [5], [9]).

Statistický rozhodovací problém formulovaný v předchozím paragrafu lze modifikovat tak, aby byly vzaty v úvahu náklady na realizaci X . Místo s bayesovskou rizikovou funkcí $\varphi(q, \delta)$ (při bayesovském přístupu) pak pracujeme s tzv. totálním rizikem

$$\varphi(q, \delta, o) = \varphi(q, \delta) + E c(g, \underline{X}),$$

kde $c(g, \underline{X})$ jsou náklady spojené s realizací \underline{X} při hodnotě parametru rovné g . Další informace o tomto problému lze nalézt např. v [5].

Mějme dvě ztrátové funkce L_1 a L_2 mezi nimiž platí vztah

$$L_1(\vartheta, d) = a L_2(\vartheta, d) + b$$

pro vš. $(\vartheta, d) \in \mathfrak{G} \times \mathfrak{D}$ a nějaké $a > 0$, $b \in R_1$. Pak bayesovské rozhodovací funkce odpovídající L_1 a L_2 jsou shodné, což vzhledem k předpokladu omezenosti ztrátové funkce zdoła, vede k tomu, že bez újmy na obecnosti můžeme předpokládat nezápornost ztrátové funkce.

Nyní si zformulujeme a dokážeme tvrzení o konkrétně bayesovské-ho rizika $\varphi^*(q)$ jako funkce rozdělení q parametru ϑ .

Věta 3.2. Pro libovolné hustoty q_1, q_2 parametru ϑ vzhledem k ϵ -konečné míře λ a pro libovolné $\alpha \in (0, 1)$ platí

$$\varphi^*(\alpha q_1 + (1-\alpha)q_2) \geq \alpha \varphi^*(q_1) + (1-\alpha) \varphi^*(q_2),$$

za předpokladu, že příslušná bayesovská rizika existují.

Důkaz. Z definice bayesovské rizikové funkce máme:

$$\varphi(\alpha q_1 + (1-\alpha)q_2, d') = \alpha \varphi(q_1, d') + (1-\alpha) \varphi(q_2, d')$$

pro vš. $d' \in \Delta$, neboť $\alpha q_1 + (1-\alpha)q_2$ je opět hustota vzhledem k míře λ a φ je lineární funkcí vzhledem ke q . Z vlastností infima pak plyne

$$\inf_{d' \in \Delta} \varphi(\alpha q_1 + (1-\alpha)q_2, d') \geq \alpha \inf_{d' \in \Delta} \varphi(q_1, d') + (1-\alpha) \inf_{d' \in \Delta} \varphi(q_2, d').$$

Q.E.D.

Nyní si všimneme bayesovské rizikové funkce $\varphi(q, d')$ definované vztahem (3.4). Můžeme si ji totiž vyjádřit následovně:

$$\begin{aligned} \varphi(q, d') &= \int \left\{ \int_{\mathfrak{G}} q(\vartheta) r(\bar{x}|\vartheta) d\lambda(\vartheta) \right\} \cdot \\ &\cdot \left(\int_{\mathfrak{D}} L(\vartheta, d'(\bar{x})) r(\vartheta|\bar{x}) d\lambda(\vartheta) \right) d\nu_{\bar{x}}(\bar{x}), \end{aligned} \quad (3.8)$$

kde $r(\vartheta|\bar{x})$ je aposteriorní hustota parametru ϑ .

Tedy pro bayesovskou rozhodovací funkci δ^* odpovídající $q(q, \delta^*)$ platí (za předpokladu $\int_{\Theta} q(\theta) r(\bar{x}|\theta) d\lambda(\theta) \neq 0$)

$$\min_{\delta \in \Delta} R_{\bar{x}} \int_{\Theta} \left\{ \int_{\mathcal{X}} L(\theta, \delta(\bar{x})) r(\theta|\bar{x}) d\lambda(\theta) \right\} d\nu(\bar{x}) = \int_{\mathcal{X}} \left\{ \int_{\Theta} L(\theta, \delta^*(\bar{x})) r(\theta|\bar{x}) d\lambda(\theta) \right\} d\nu(\bar{x}). \quad (2.9)$$

Najdeme-li při pevném \bar{x} hodnotu $\delta^0(\bar{x}) \in \mathcal{D}$ takovou, že

$$\min_{\delta(\bar{x}) \in \mathcal{D}} \int_{\Theta} L(\theta, \delta(\bar{x})) r(\theta|\bar{x}) d\lambda(\theta) = \int_{\Theta} L(\theta, \delta^0(\bar{x})) r(\theta|\bar{x}) d\lambda(\theta), \quad (3.10)$$

pak je odpovídající funkce δ^0 bayesovskou rozhodovací funkcí. Hodnotu $\delta^*(\bar{x})$ tedy můžeme najít minimalizací podmíněné střední hodnoty ztrátové funkce vzhledem k aposteriorní hustotě $r(\theta|\bar{x})$ (pokud existuje bod minimalizující tuto podmíněnou střední hodnotu), tj. za podmínky $\bar{x} = \bar{x}$. Tento fakt je velice užitečný v konkrétních situacích, neboť obvykle stačí znát hodnotu δ^* pouze v bodě $\bar{x} = \bar{x}$, \bar{x} je realizace \bar{X} .

Nahradíme-li nyní v (3.10) aposteriorní hustotu $r(\theta|\bar{x})$ hustotou apriorní $q(\theta)$, dostaneme tzv. bayesovské rozhodnutí bez provedení pokusů, tj. k rozhodnutí použijeme pouze informace plynoucí z apriorního rozdělení.

Povšimněme si, že bayesovská rozhodovací funkce δ^* je založena na informaci plynoucí jak z \bar{X} , tak z apriorního rozdělení parametru θ .

Příklad 1.1 (pokračování ze str. 8). Uvažujme problém odhadu kvocientu inteligence θ . Ztrátové funkci

$$L(\theta, \delta(X)) = (\theta - \delta(X))^2$$

odpovídá bayesovská riziková funkce (viz (3.8))

$$\varphi(q, \delta^*) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} (\theta - \delta^*(x))^2 \pi(\theta|x) d\theta \right) \int_{-\infty}^{+\infty} q(\theta) \pi(x|\theta) d\theta dx,$$

kde aposteriorní hustota $\pi(\theta|x)$ je $N(\frac{225}{325}x + \frac{100}{325}, 100; 69,23)$ a apriorní hustota $q(\theta)$ je $N(100, 225)$. Najít bayesovskou rozhodovací funkci znamená najít $\delta^*(x)$, které minimalizuje

$$\int_{-\infty}^{+\infty} (\theta - \delta^*(x))^2 \pi(\theta|x) d\theta.$$

Jak vyplývá z úvah v následujícím paragrafu

$$\delta^{*}(x) = \int_{-\infty}^{+\infty} \theta \pi(\theta|x) d\theta = \frac{225}{325}x + \frac{100}{325} \cdot 100. \quad (3.11)$$

Přímým výpočtem pak dostaneme

$$\begin{aligned} \varphi(q, \delta^{*}(x)) &= 69,23 \int_{-\infty}^{+\infty} (2\pi \cdot 325)^{-1/2} \exp\left\{-\frac{1}{2} \cdot \frac{(x-100)^2}{325}\right\} dx = \\ &= 69,23. \end{aligned}$$

Zatímco bayesovské rozhodnutí δ^+ bez provedení pokusu je

$$\delta^+ = 100$$

a

$$\varphi(q, \delta^+) = \int (\theta - \delta^+)^2 q(\theta) d\theta = 225. \quad (3.12)$$

Příklad 3.1. Mějme náhodnou veličinu $X = \theta + Y$, kde Y je náhodná veličina s hustotou

$$\begin{aligned} f(y) &= e^{-y} & y \geq 0, \\ &= 0 & y < 0, \end{aligned}$$

a θ je neznámý parametr, o němž můžeme předpokládat, že má rozdělení stejné jako Y .

Uvažujme rozhodovací problém (Θ, Δ, R) , kde $\Theta = \langle 0, +\infty \rangle$, Δ je množina funkcí z $\langle 0, +\infty \rangle$, do

$$\mathcal{D} = \{d_1, d_2, d_3, d_4\},$$

kde $d_i = \langle i-1, i \rangle$, $i = 1, 2, 3$, $d_4 = \langle 3, +\infty \rangle$ a riziková funkce odpovídá ztrátové funkci dané následující tabulkou.

Klasifikace dle výsledku měření

	d_1	d_2	d_3	d_4
d_1	0	1	1	2
d_2	1	0	2	2
d_3	1	2	0	2
d_4	3	3	3	0

skutečnost

Aposteriorní rozdělení θ je rovnoměrné na $(0, X)$. Dále platí, že

$$\int_{-\infty}^{+\infty} L(\theta, d_1) \pi(\theta|X) d\theta = \int \{I\{3 < X\}\} (3X-7)/X + I\{1 < X < 3\}\} (X-1)\} / X + I\{X < 1\}\} \cdot 0$$

$$\int_{-\infty}^{+\infty} L(\theta, d_2) \pi(\theta|X) d\theta = \int \{I\{X < 1\}\} X + I\{1 < X < 2\}\} + I\{2 < X < 3\}\} (2X-3) + I\{X > 3\}\} (3X-6)\} / X$$

$$\int_{-\infty}^{+\infty} L(\theta, d_3) \pi(\theta|X) d\theta = \int \{I\{X < 1\}\} \cdot X + I\{1 < X < 2\}\} (2X-1) + I\{2 < X < 3\}\} 3 + I\{X > 3\}\} (3X-6)\} / X$$

$$\int_{-\infty}^{+\infty} L(\theta, d_4) \pi(\theta|X) d\theta = 2X^{-1} \min(3, X).$$

Nedy bayesovské rozhodnutí je následující:

- je-li $X = 2$, pak je rozhodnutí d_1 nebo d_2 ,
- je-li $X = 13/3$ " " d_1 nebo d_4 ,
- je-li $X < 13/3$ " " d_1 ,
- je-li $X > 13/3$ " " d_4 .

Na závěr tohoto paragrafu si zformulujeme větu pro případ, že Θ i δ obsahují právě 2 body. Toto je případ, se kterým se nejčastěji setkáváme při testování jednoduché nulové hypotézy proti jednoduché alternativní hypotéze. Předpokládáme, že ztrátová funkce je definována následovně:

$$\begin{aligned} L(\theta_1, d_1) &= 0, & i &= 1, 2 \\ L(\theta_1, d_2) &= a_1, \\ L(\theta_2, d_1) &= a_2, \end{aligned}$$

kde $a_1 > 0$, $\Theta = \{d_1, d_2\}$, $\Theta = \{\theta_1, \theta_2\}$.
Dále předpokládáme, že

$$P(\theta = \theta_1) = \xi, \quad P(\theta = \theta_2) = 1 - \xi, \quad \xi \in (0, 1) \text{ dáno.}$$

Pak podle (3.4) pro bayesovskou rizikovou funkci platí

$$\varphi(\xi, d^*) = a_1 \xi P(d^* | \theta = \theta_1) + a_2 (1 - \xi) P(d^* | \theta = \theta_2). \quad (3.13)$$

Pravděpodobnosti $P(d^* | \theta = \theta_1)$ a $P(d^* | \theta = \theta_2)$ jsou podmíněné pravděpodobnosti chybných rozhodnutí. Najít bayesovskou rozhodovací funkci d^* vlastně znamená najít takovou rozhodovací funkci d^* , která minimalizuje jistou lineární kombinaci podmíněných pravděpodobností chybných rozhodnutí. Řešení problému je obsaženo v následující větě.

Věta 3.3. Pro libovolné $a > 0$ a $b > 0$ definujeme rozhodovací funkci předpisem

$$\begin{aligned} d^*(x) &= d_1 & \text{je-li} & \quad a r(x | \theta_1) > b r(x | \theta_2) \\ &= d_2 & \text{je-li} & \quad a r(x | \theta_1) < b r(x | \theta_2) \\ &= \text{libovolně je-li} & \quad a r(x | \theta) = b r(x | \theta_2). \end{aligned}$$

Pak pro libovolnou rozhodovací funkci $d^* : X \rightarrow \Theta$ platí

$$\begin{aligned} a P(d^*(X) = d_2 | \theta = \theta_1) + b P(d^*(X) = d_1 | \theta = \theta_2) &\leq \\ \leq a P(d^*(X) = d_2 | \theta = \theta_1) + b P(d^*(X) = d_1 | \theta = \theta_2). \end{aligned}$$

Důkaz. Pro libovolnou rozhodovací funkci δ platí:

$$\begin{aligned}
 a P(\delta(\tilde{x}) = d_2 | \Theta = \Theta_1) + b P(\delta(\tilde{x}) = d_1 | \Theta = \Theta_2) &= & (3.14) \\
 &= a \int_{\delta(\tilde{x})=d_2} r(\tilde{x} | \Theta_1) d\nu_n(\tilde{x}) + b \int_{\delta(\tilde{x})=d_1} r(\tilde{x} | \Theta_2) d\nu_n(\tilde{x}) = \\
 &= a + \int_{\delta(\tilde{x})=d_1} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}) = \\
 &= a + \int_{R_n} I\{\delta(\tilde{x})=d_1, \delta^*(\tilde{x})=d_1\} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}) \\
 &\quad + \int_{R_n} I\{\delta(\tilde{x})=d_1, \delta^*(\tilde{x})=d_2\} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}).
 \end{aligned}$$

Z definice δ^* plyne

$$\begin{aligned}
 &\int_{R_n} I\{\delta(\tilde{x})=d_1, \delta^*(\tilde{x})=d_1\} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}) \geq & (3.15) \\
 &\geq \int_{R_n} I\{\delta^*(\tilde{x})=d_1\} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}) \\
 &= \int_{R_n} I\{\delta^*(\tilde{x})=d_1, \delta^*(\tilde{x})=d_2\} (-a r(\tilde{x} | \Theta_1) + b r(\tilde{x} | \Theta_2)) d\nu_n(\tilde{x}) \geq 0 & (3.16)
 \end{aligned}$$

Z (3.14-3.16) snadno obdržíme tvrzení věty.

Q.E.D.

Poznamenejme, že při vedlejších podmínkách $P(\delta^*(\tilde{x})=d_1 | \Theta=\Theta_1) = \alpha$ je tvrzení věty shodné s tvrzením Neyman-Pearsonova lemma tu.

4. ÚLOHA ODHADU

4.1 ÚVOD

Nechť $\tilde{X} = (X_1, \dots, X_n)'$ má hustotu $r(\tilde{x}|\theta)$ vzhledem k σ -konečné míře γ_n , $\theta = (\theta_1, \dots, \theta_k) \in \Theta$ je parametr, $\emptyset \neq \Theta \in \mathcal{A}_X$. Nechť $q(\theta)$ je apriorní hustota parametru θ vzhledem k σ -konečné míře λ na $(\Theta, \mathcal{A}(\Theta))$ a $r(\theta|\tilde{x})$ příslušná aposteriorní hustota.

Úlohu odhadu parametru θ můžeme formulovat jako statistický rozhodovací problém (Θ, Δ, R) , kde množina možných rozhodnutí \mathcal{D} je shodná s množinou Θ . Rozhodovací funkce $d(\tilde{X})$ je pak odhad parametru θ a Δ je množina odhadů parametru θ . Nadále budeme používat termín odhad místo rozhodovací funkce. Ztrátová funkce $L(\theta, d(\tilde{X}))$ pak vyjadřuje odlišnost odhadu $d(\tilde{X})$ od skutečné hodnoty parametru θ .

4.2 BODOVÝ ODHAD, JEDNOROZMĚRNÝ PŘÍPAD

Typickými ztrátovými funkcemi jsou

$$L_{a,w}(\theta, d) = w(\theta) |\theta - d|^a \quad \theta \in \Theta, \quad d \in \mathcal{D}, \quad (4.1)$$

$$\begin{aligned} L_{k_0, k_1}(\theta, d) &= k_0 |\theta - d| & \theta - d \geq 0, \quad \theta \in \Theta, \quad d \in \mathcal{D}, & \quad (4.2) \\ &= k_1 (d - \theta) & \theta - d < 0, & \end{aligned}$$

kde $w(\theta)$ je nezáporná měřitelná funkce definovaná na Θ ; a, k_0, k_1 jsou kladné konstanty pevně zvolené. Nejčastěji používáme $w(\theta) = 1$, $a = 1$ nebo $a = 2$ a $k_0 = k_1 = 1$. Pro $w(\theta) = 1$, $a = k_0 = k_1 = 1$ jsou ztrátové funkce totožné.

Dále se budeme zabývat pouze $L_{2,w}$ a L_{k_0, k_1} . Pro $a = 2$, $w(\theta) = 1$ budeme používat pro ztrátovou funkci zkrácené značení L_2 .

Označme

$$q_1(\theta) = q(\theta)w(\theta) \left(\int q(\theta)w(\theta)d\lambda(\theta) \right)^{-1} \quad \text{pro } \int q(\theta)w(\theta)d\lambda(\theta) \neq 0$$

(4.3)

jinak,

$$= 0$$

$$\pi_1(\theta|\bar{x}) = \frac{r(\bar{x}|\theta) q_1(\theta)}{\int_{\Theta} r(\bar{x}|\theta) q_1(\theta) d\lambda(\theta)} \quad \text{pro } \int_{\Theta} q_1(\theta) r(\bar{x}|\theta) d\lambda(\theta) \neq 0$$

(4.4)

jinak,

$$= 0$$

$$r_1(\bar{x}) = \int_{\Theta} r(\bar{x}|\theta) q_1(\theta) d\lambda(\theta) \quad \bar{x} \in R_n. \quad (4.5)$$

Zřejmě q_1 je hustota (pro $\int q(\theta)w(\theta)d\lambda(\theta) \neq 0$) vzhledem k λ a považuje-
me-li ji za apriorní hustotu, pak příslušná aposteriorní hustota je
 $r_1(\theta|\bar{x})$; $r_1(\bar{x})$ je marginální hustota \bar{x} odpovídající $q_1(\theta)$.

Hlavní výsledky o bayesovských odhadech pro ztrátové funkce $L_{2,w}$
a L_{k_0, k_1} si zformulujeme ve větě:

Věta 4.1. a) Nechť

$$0 < \int_{\Theta} \theta^2 w(\theta) \pi(\theta|\bar{x}) d\lambda(\theta) < +\infty$$

pak pro ztrátovou funkci $L_{2,w}$ definovanou (4.1) je bayesovský odhad
 $\delta_{2,w}^*$ parametru θ dán vztahem

$$\delta_{2,w}^*(\bar{x}) = \frac{\int_{\Theta} \theta w(\theta) \pi(\theta|\bar{x}) d\lambda(\theta)}{\int_{\Theta} w(\theta) \pi(\theta|\bar{x}) d\lambda(\theta)} = E_{\pi}(\theta|\bar{x}=\bar{x}) \quad \text{pro } \int_{\Theta} w(\theta) \pi(\theta|\bar{x}) d\lambda(\theta) \neq 0$$

(4.6)

jinak

$$= 0$$

pro bayesovské riziko $\rho_{2,w}^*(q)$ platí

$$\rho_{2,w}^*(q) = E_{\pi}(\text{var}_{\pi}(\theta|\bar{X})) , \quad (4.7)$$

kde E_{ξ} označuje střední hodnotu \bar{x} vzhledem k hustotě (4.5) a var_{ξ} označuje podmíněný rozptyl θ vzhledem k hustotě (4.4).

b) Necht

$$0 < \int_{\Theta} |\theta| r(\bar{x}|\theta) q(\theta) d\lambda(\theta) < +\infty \quad (4.8)$$

pak pro ztrátovou funkci I_{k_0, k_1} definovanou (4.2) je bayesovský odhad δ_{k_0, k_1}^{**} parametru θ ($(100(k_0 + k_1))^{-1}$)% kvantil aposteriorního rozdělení $\pi(\theta|\bar{x})$ a pro bayesovské riziko $\rho_{k_0, k_1}^*(q)$ platí

$$\begin{aligned} \rho_{k_0, k_1}^*(q) &= (k_0 + k_1) \int_{R_{\Delta}} \int_{-\infty}^{+\infty} \delta_{k_0, k_1}^{**}(\bar{x}) (d_{k_0, k_1}^{**}(\bar{x}) - \theta) r(\bar{x}|\theta) q(\theta) d\nu_{\Delta}(\bar{x}) d\lambda(\theta) + \\ &+ k_0 \int_{R_{\Delta}} \int_{-\infty}^{+\infty} (\theta - d_{k_0, k_1}^{**}(\bar{x})) r(\bar{x}|\theta) q(\theta) d\nu_{\Delta}(\bar{x}) d\lambda(\theta). \end{aligned}$$

Důsledek 4.2. Je-li ve větě 4.1 a) $w(\theta) \equiv 1$, pak bayesovský odhad $\delta_2^{**}(\bar{x})$ parametru θ je

$$d_2^{**}(\bar{x}) = E(\theta|\bar{x}=\bar{x}) \quad (4.10)$$

a pro bayesovské riziko platí

$$\rho_2^*(q) = E\{\text{var}(\theta|\bar{x})\}. \quad (4.11)$$

Připomejme, že 100 α % kvantil $\hat{\theta}_{\alpha}(\bar{x})$ rozdělení $\pi(\theta|\bar{x})$ je definován vztahem

$$\int_{\Theta \cap \{\theta \leq \hat{\theta}_{\alpha}(\bar{x})\}} \pi(\theta|\bar{x}) d\lambda(\theta) < \alpha, \quad \int_{\Theta \cap \{\theta > \hat{\theta}_{\alpha}(\bar{x})\}} \pi(\theta|\bar{x}) d\lambda(\theta) > 1 - \alpha.$$

Poznámka 4.3. Bayesovský odhad $\delta_{2, w}^{**}(\bar{x})$ je podmíněná střední hodnota θ při daném $\bar{x} = \bar{x}$ vzhledem k hustotě $\pi_1(\theta|\bar{x})$.

Důkaz věty 4.1. a) Pro libovolné $\delta \in \Delta$ a libovolné $\bar{x} \in R_{\Delta}$ platí

$$E_1(\theta - \delta_{2, w}^{**}(\bar{x})) (\delta_{2, w}^{**}(\bar{x}) - \delta(\bar{x})) = 0$$

a tudíž

$$\begin{aligned}
q(q, \delta) &= E_1(\theta - \delta(\bar{x}))^2 = \int_{R_{\mu}} \int_{\Theta} (\theta - \delta(\bar{x}))^2 \tau_1^*(\bar{x}|\theta) d\lambda(\theta) r_1(\bar{x}) d\nu_{\mu}(\bar{x}) = \\
&= \int_{R_{\mu}} \int_{\Theta} \left\{ \int_{R_{\mu}} (\theta - \delta_{2,w}^*(\bar{x}))^2 \tau_1(\bar{x}|\theta) d\lambda(\theta) + \right. \\
&\quad \left. + \int_{\Theta} (\delta(\bar{x}) - \delta_{2,w}^*(\bar{x}))^2 \tau_1(\bar{x}|\theta) d\lambda(\theta) \right\} r_1(\bar{x}) d\nu_{\mu}(\bar{x}) \geq \\
&\geq \int_{R_{\mu}} \int_{\Theta} (\theta - \delta_{2,w}^*(\bar{x}))^2 \tau_1(\bar{x}|\theta) d\lambda(\theta) = E_1(\text{var}\{\theta|\bar{X}\}).
\end{aligned}$$

Odtud plyne, že $\delta_{2,w}^*(\bar{x})$ je bayesovský odhad a bayesovské riziko je dáno (4.7).

b) Necht δ je libovolný odhad parametru θ . Je-li $\delta(\bar{x}) > \delta_{k_0, k_1}^*(\bar{x})$, pak platí

$$\begin{aligned}
I_{k_0, k_1}(\theta, \delta(\bar{x})) - I_{k_0, k_1}(\theta, \delta_{k_0, k_1}^*(\bar{x})) &= k_0(\delta_{k_0, k_1}^*(\bar{x}) - \delta(\bar{x})) \quad \text{je-li} \\
&\quad \theta - \delta(\bar{x}) \geq 0 \\
&= k_1(-\theta + \delta(\bar{x})) - k_0(\theta - \delta_{k_0, k_1}^*(\bar{x})) \\
&\quad \text{je-li } \theta - \delta(\bar{x}) < 0 \leq \theta - \delta_{k_0, k_1}^*(\bar{x}) \\
&= k_1(\delta(\bar{x}) - \delta_{k_0, k_1}^*(\bar{x})) \quad \text{je-li} \\
&\quad \theta - \delta_{k_0, k_1}^*(\bar{x}) < 0,
\end{aligned}$$

což implikuje (pro jednodušší zápis klademe $q(\theta) = 0$ pro $\theta \notin \Theta$)

$$\begin{aligned}
\int_{-\infty}^{+\infty} (I_{k_0, k_1}(\theta, \delta(\bar{x})) - I_{k_0, k_1}(\theta, \delta_{k_0, k_1}^*(\bar{x}))) r(\bar{x}|\theta) q(\theta) d\lambda(\theta) &= \quad (4.11) \\
= k_0(\delta_{k_1, k_2}^*(\bar{x}) - \delta(\bar{x})) \left(\int_{-\infty}^{+\infty} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) - \int_{\delta_{k_0, k_1}^*(\bar{x})}^{\delta(\bar{x})} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) \right) - \\
\quad - \int_{-\infty}^{\delta_{k_0, k_1}^*(\bar{x})} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) + \\
\quad + (k_1 \delta(\bar{x}) + k_0 \delta_{k_0, k_1}^*(\bar{x})) \int_{\delta_{k_0, k_1}^*(\bar{x})}^{\delta(\bar{x})} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) -
\end{aligned}$$

$$\begin{aligned}
& - (k_0+k_1) \int_{\delta_{k_0, k_1}^*(\bar{x})}^{\delta^*(\bar{x})} \text{er}(\bar{x}|\theta)q(\theta)d\lambda(\theta) + \\
& \int_{\delta_{k_0, k_1}^{**}(\bar{x})}^{\delta^*(\bar{x})} \text{er}(\bar{x}|\theta)q(\theta)d\lambda(\theta) + \\
& + k_1 (\delta^*(\bar{x}) - \delta_{k_0, k_1}^{**}(\bar{x})) \int_{-\infty}^{\delta_{k_0, k_1}^{**}(\bar{x})} r(\bar{x}|\theta)q(\theta)d\lambda(\theta).
\end{aligned}$$

Použijeme-li nerovnost

$$\int_{\delta_{k_0, k_1}^{**}(\bar{x})}^{\delta^*(\bar{x})} \text{er}(\bar{x}|\theta)q(\theta)d\lambda(\theta) \leq \delta^*(\bar{x}) \int_{\delta_{k_0, k_1}^{**}(\bar{x})}^{\delta^*(\bar{x})} r(\bar{x}|\theta)q(\theta)d\lambda(\theta)$$

(4.11)

a sečteme-li na pravé straně integrály s mezemi $(\delta_{k_0, k_1}^{**}(\bar{x}), \delta^*(\bar{x}))$, dostaneme, že jejich součet je nezáporný. Kromě toho z definice kvantilu plyne, že

$$\int_{-\infty}^{\delta_{k_0, k_1}^{**}(\bar{x})} r(\bar{x}|\theta)q(\theta)d\lambda(\theta) \geq \int_{-\infty}^{+\infty} r(\bar{x}|\theta)q(\theta)d\lambda(\theta) \cdot \frac{k_0}{k_1+k_0},$$

což dále zajišťuje, že součet zbylých členů na pravé straně (4.11) je nezáporný. Odtud již plyne naše tvrzení pro $\delta^*(\bar{x}) \geq \delta_{k_0, k_1}^{**}(\bar{x})$. V případě platnosti opačné nerovnosti postupujeme analogicky. Tedy $100 k_0(k_0+k_1)^{-1}\%$ kvantil aposteriorního rozdělení parametru θ je bayesovský odhad pro I_{k_0, k_1} , a bayesovské riziko je rovno

$$\begin{aligned}
\varrho_{k_0, k_1}^{\delta^*}(q) &= \int_{R_{\Pi}} \int_{-\infty}^{+\infty} I_{k_0, k_1}(\theta, \delta_{k_0, k_1}^{**}(\bar{x}))r(\bar{x}|\theta)q(\theta)d_{\Pi}(\bar{x})d\lambda(\theta) = \\
&= (k_0+k_1) \int_{R_{\Pi}} \int_{-\infty}^{\delta_{k_0, k_1}^{**}(\bar{x})} (\delta_{k_0, k_1}^{**}(\bar{x}) - \theta)r(\bar{x}|\theta)q(\theta)d_{\Pi}(\bar{x})d\lambda(\theta) \\
&+ k_0 \int_{R_{\Pi}} \int_{-\infty}^{+\infty} (\theta - \delta_{k_0, k_1}^{**}(\bar{x}))r(\bar{x}|\theta)q(\theta)d_{\Pi}(\bar{x})d\lambda(\theta).
\end{aligned}$$

q. e. d.

Z důsledku 4.2 a poznámky 4.3 plyne, že bayesovský odhad δ_2^{**} je podmíněná střední hodnota parametru θ při aposteriorní hustotě $r(\theta|\bar{x})$.

Tedy pro aposteriorní hustoty $\pi(\theta|x)$ uvedené v kapitole 2 snadno spočteme odhady δ_2^* , stačí nalézt v Apendixu příslušnou střední hodnotu. Za zmínku stojí, že pro limitní aposteriorní hustoty uvedené v 2.3 se většina bayesovských odhadů parametru θ shoduje s odhady běžně používanými v klasické statistice.

Příklad 1.1 (pokračování ze str. 8,45). Vyjádření (3.11) pro bayesovský odhad δ_2^* příslušný ztrátové funkci L_2 plyne z (4.6). Tento odhad je totožný s bayesovským odhadem $\delta_{1,1}^*$ odpovídajícím ztrátové funkci $L_{1,1}$. Záleží-li nám na podrobení dětí podprůměrných nebo nadprůměrných, můžeme volit místo ztrátové funkce L_2 funkci

$$L(\theta, \delta) = (\theta - \delta)^2 \exp \frac{1}{2r} (\theta - 100)^2,$$

kde $r > 225$. Podle (4.3) a (4.4) postupně dostaneme, že hustota

$q_1(\theta)$ je $N(100, (\frac{1}{225} - \frac{1}{r})^{-1})$ a hustota $\pi_1(\theta|x)$ je

$$N\left(\frac{225rx + 100^2(r - 225)}{225r + 100(r - 225)}, \frac{100 \cdot 225 \cdot r}{100(r - 225) + 225r}\right).$$

Z věty 4.1 plyne, že bayesovský odhad je

$$\delta_{2,w}^*(x) = \frac{225rx + 100^2(r - 225)}{225r + 100(r - 225)}$$

a bayesovské riziko je

$$R_{2,w}^*(q) = \frac{225 \cdot 100 \cdot r}{100(r - 225) + 225r},$$

kde $w(\theta) = \exp\{(\theta - 100)^2/2r\}$. Porovnáme-li tento bayesovský odhad s odhadem získaným při L_2 , zjistíme, že při $x < 100$ je

$$\delta_{2,w}^*(x) < \delta_2^*(x)$$

a při $x > 100$ platí opačná nerovnost. Tedy při $\delta_{2,w}^*$ pro $x < 100$ θ spíše podhodnotíme a pro $x > 100$ spíše nadhodnotíme ve srovnání s použitím odhadu δ_2^* .

Příklad 4.1. Chceme odhadnout pravděpodobnost θ výskytu určitého znaku u populace jedinců. Necht' X_1, \dots, X_n je náhodný výběr z této populace, kde

$$X_i = 1 \quad \text{u } i\text{-tého vybraného jedince zjištěn znak} \\ = 0 \quad \text{jinak,}$$

$i = 1, \dots, n$. Tedy $P(X_i=1) = \theta$ a $P(X_i=0) = 1 - \theta$.

Předpokládejme, že θ má beta rozdělení (α, β) , $\alpha > 0$, $\beta > 0$ a požadujeme co nejlepší odhad pro θ v okolí 0 a 1. Zvolíme-li

$$L(\theta, d^*(\bar{x})) = \frac{(\theta - d^*(\bar{x}))^2}{\theta(1-\theta)},$$

pak při téže hodnotě $(\theta - d^*(\bar{x}))^2$ je ztráta pro θ v okolí 0 a 1 mnohem větší než např. při θ v okolí 1/2. Podle (4.6) je pak bayesovský odhad dán vzorcem

$$d^*(\bar{x}) = \frac{\int_0^1 \theta^{\alpha + \sum_{i=1}^n x_i - 1} (1-\theta)^{\beta + n - \sum_{i=1}^n x_i - 2} d\theta}{\int_0^1 \theta^{\alpha + \sum_{i=1}^n x_i - 2} (1-\theta)^{\beta + n - \sum_{i=1}^n x_i - 2} d\theta} = \\ = \frac{B(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i - 1)}{B(\alpha + \sum_{i=1}^n x_i - 1, \beta + n - \sum_{i=1}^n x_i - 1)} = \\ = \frac{\alpha + \sum_{i=1}^n x_i - 1}{\alpha + \beta + n - 2},$$

není-li $0 < \alpha \leq 1$ a $\sum_{i=1}^n x_i = 0$ nebo $0 < \beta \leq 1$ a $\sum_{i=1}^n x_i = n$, kde $\sum_{i=1}^n x_i$ znamená $\sum_{i=1}^n x_i$.

Je-li $0 < \beta \leq 1$ a $\sum_{i=1}^n x_i = 0$, pak

$$d^*(\bar{x}) = 0.$$

Je-li $0 < \beta \leq 1$ a $\sum_{i=1}^n x_i = n$, pak d^* není definováno vztahem (4.6), ale můžeme ho dodefinovat pomocí limity:

$$d^*(\bar{x}) = \lim_{\alpha \rightarrow 1^-} \frac{\int_0^{\alpha} \theta^{\alpha+n-1} (1-\theta)^{\beta-2} d\theta}{\int_0^{\alpha} \theta^{\alpha+n-2} (1-\theta)^{\beta-2} d\theta} =$$

$$= \lim_{a \rightarrow 1} \frac{a^{\alpha+n-1} (1-a)^{\beta-2}}{a^{\alpha+n-2} (1-a)^{\beta-2}} = 1.$$

Pro bayesovské riziko platí ($\alpha > 1, \beta > 1$)

$$\begin{aligned} \varphi^*(q) &= \sum_{x=0}^n \binom{n}{x} \frac{(\alpha+x-1)(\beta+n-x-1)}{(\alpha+\beta+n-2)^2 (\alpha+\beta+n-1)} \cdot \frac{B(\alpha+x-1, \beta+n-x-1)}{B(\alpha-1, \beta-1)} = \\ &= \sum_{x=0}^n \binom{n}{x} \frac{B(\alpha+x, \beta+n-x)}{B(\alpha-1, \beta-1) (\alpha+\beta+n-2)} = \frac{(\alpha-1)(\beta-1)}{(\alpha+\beta-1)(\alpha+\beta-2)} (\alpha+n+\beta-2), \end{aligned}$$

kde jsme použili faktu, že marginální hustota $r_1(\bar{x})$ náhodného vektoru x je podle (4.5) rovna

$$r_1(\bar{x}) = \frac{B(\alpha+x-1, \beta+n-x-1)}{B(\alpha-1, \beta-1)} \quad \text{pro } x_i = 0 \text{ nebo } 1$$

x označuje $\sum_{i=1}^n x_i$ (jde o hustotu vzhledem k čítecí míře).

Při klasickém přístupu je $\bar{x} = \frac{1}{n} \sum_1 X_i$ eficientním odhadem.

Příklad 4.2. Elektronické součástky jsou zkoušeny za účelem odhadu střední životnosti. Předpokládejme, že doby životnosti jednotlivých součástek představují výběr X_1, \dots, X_n z exponenciálního rozdělení s parametrem $1/\theta$. Předchozí měření umožňují předpokládat, že θ má inverzní gama rozdělení s parametry (β, α) , $\alpha > 0$, $\beta > 0$ dané hustotou

$$\begin{aligned} q(\theta; \alpha, \beta) &= (\Gamma(\alpha))^{-1} \theta^{-\alpha} e^{-\theta/\beta} \theta^{-\alpha-1} e^{-(\theta/\beta)^{-1}} & \theta > 0 \\ &= 0 & \theta \leq 0. \end{aligned}$$

Aposteriorní hustota parametru θ je

$$\begin{aligned} \pi(\theta | \bar{x}) &= (\Gamma(\alpha+n))^{-1} (1/\beta)^{\alpha+n} \theta^{-\alpha-n-1} \exp \left\{ -\left(\frac{1}{\beta} + \sum_{i=1}^n x_i \right) / \theta \right\} & \theta > 0 \\ &= 0 & \theta \leq 0, \end{aligned}$$

t.j. inverzní gama rozdělení s parametry $\left(\left(\frac{1}{\beta} + \sum_{i=1}^n x_i \right)^{-1}, \alpha+n \right)$. Při levá-draťkové ztrátové funkci $L_{2,w}$ s $w = 1$ dané (4.1) je bayesovský odhad

dán

$$g^*(x) = \int_{-\infty}^{+\infty} g(\theta | \bar{x}) d\theta = \frac{1}{\beta} + \sum_{i=1}^n x_i$$

a pro bayesovské riziko platí pro $\alpha > 2$, $\beta > 0$

$$\begin{aligned} g^*(q) &= E\{\text{var}(\theta | \bar{X})\} = E\left(\frac{1}{\beta} + \sum_{i=1}^n x_i\right)^2 (\alpha + n - 1)^{-2} (\alpha + n - 2)^{-1} = \\ &= \frac{\beta^2}{(\alpha - 1)(\alpha - 2)(\alpha + n - 1)}, \end{aligned}$$

kde jsme použili faktu, že marginální hustota \bar{X} je (viz (1.7)) rovna

$$\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\beta^\alpha (1/\beta + \sum_{i=1}^n x_i)^{\alpha+n}} \quad \text{pro } x_i > 0, i = 1, \dots, n$$

a rovna nule jinak.

Příklad 1.2 (pokračování ze str. 8). Předpokládejme, že dva fyzikové vyjádří svou představu o sledované fyzikální konstantě θ následovně. Zkušenější z nich říká, že možné hodnoty θ mají rozdělení $N(900, 400)$. Zatímco druhý z nich (méně zkušený) říká, že možné hodnoty θ mají rozdělení $N(800, 6400)$ (větší rozptyl odráží menší zkušenost). Předpokládejme, že výsledek $X = x$ příslušného pokusu má rozdělení $N(\theta, 1600)$.

Pak aposteriorní rozdělení při použití apriorní informace zkušenějšího fyzika je

$$N\left(\frac{x+3600}{5}, 320\right)$$

a při použití apriorní informace fyzika méně zkušeného je

$$N\left(\frac{4 \cdot x + 800}{5}, 1750\right).$$

Je vidět, že u méně zkušeného fyzika došlo k výraznému snížení rozptylu ve srovnání s fyzikem zkušenějším.

Při ztrátové funkci L_2 dostaneme odhad θ rovný $(x+3600)/5$ u zkušeného fyzika a $(4 \cdot x + 800)/5$ u méně zkušeného. Tedy odhady jsou různé. Obecně při n pokusech bude aposteriorní rozdělení při použití infor-

mače zkušnějšího fyzika

$$N\left(\frac{\bar{x}+3600/n}{1+4/n}, \frac{1600}{n+4}\right)$$

a méně zkušnějšího fyzika

$$N\left(\frac{4\bar{x}+800/n}{4+n}, \frac{6400}{4n+1}\right),$$

kde \bar{x} je aritmetický průměr výsledků pokusů. Tedy při provedení více pokusů bude vliv apriorní informace rychle klesat, což je v soulase s větou 2.1.

Vedle dvou uvedených typů odhadů se používá ještě bayesovský

odhad $\hat{\theta}_M$ maximálně věrohodného typu, který je definován následovně:

$$\max_{\theta \in \Theta} r(\bar{x}|\theta)q(\theta) = r(\bar{x}|\hat{\theta}_M)q(\hat{\theta}_M),$$

pokud maximum existuje. Někdy se též mluví o zobecněném maximálně věrohodném odhadu, neboť při $q(\theta)$ konstantním dostaneme obvyčejný maximálně věrohodný odhad. Poznamenejme, že zobecněný maximálně věrohodný odhad nemusí odpovídat žádné ztrátové funkci. Tato metoda se dá použít i ve vícerozměrném případě.

Příklad 4.2 (pokračování ze str. 57). Bayesovský maximálně věrohodný odhad $\hat{\theta}_M$ maximalizuje funkci

$$\theta^{\alpha-n-1} \exp\left\{-\left(\frac{1}{\theta} + \sum_1 x_i\right)/\theta\right\}$$

pro $\theta > 0$, výpočtem dostaneme

$$\hat{\theta}_M = \frac{\frac{1}{\theta} + \sum_1 x_i}{\alpha + n - 1},$$

$$\hat{\theta}_M = \theta_{2,1}^* (\bar{x}).$$

Nyní si uvedeme 2 příklady na odhad parametru, jestliže apriorní rozdělení bylo získáno empirickou bayesovskou metodou, která byla vyložena v § 2.4.

Příklad 4.3. Necht X_1, \dots, X_n jsou nezávislé náhodné veličiny, X_1 má rozdělení $N(\theta, \sigma_0^2)$, kde $\sigma_0^2 > 0$ je známé. Necht Y_1, \dots, Y_N jsou nezávislé náhodné veličiny, které představují výsledky z minulosti, Y_1 má jít rozdělení $N(\theta_q, \sigma_q^2)$, kde θ_q i $\sigma_q^2 > 0$ jsou neznámé.

Uvažujeme úlohu odhadu parametru θ při ztrátové funkci I_2 dané (4.1).

Podle (2.39) a (2.41) za odhady $\hat{\theta}_q$ a $\hat{\sigma}_q^2$ parametrů θ_q resp. σ_q^2 vezmeme

$$\begin{aligned} \hat{\theta}_q &= \bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i \\ \hat{\sigma}_q^2 &= \hat{\sigma}_0^2 - \sigma_0^2 \quad \text{je-li } \hat{\sigma}_q^2 > \sigma_0^2 \\ &= 0 \quad \text{jinak,} \end{aligned}$$

kde $\hat{\sigma}_0^2$ je dáno (2.40). Odhad $\hat{\theta}$ parametru θ je pak podle (2.9) a věty 4.1.a)

$$\hat{\theta} = \frac{\hat{\sigma}_0^2 n \bar{X} + \sigma_0^2 \bar{Y}}{n \hat{\sigma}_q^2 + \sigma_0^2},$$

přičemž $\hat{\theta} = \hat{Y}$ je-li $\hat{\sigma}_q^2 \leq \sigma_0^2$.

Je-li navíc $n = 1$, pak

$$\begin{aligned} \hat{\theta} &= X_1 + (\bar{Y} - X_1) \sigma_0^2 / \hat{\sigma}_q^2 \quad \text{je-li } \hat{\sigma}_q^2 > \sigma_0^2 \\ &= \bar{Y} \quad \text{jinak.} \end{aligned}$$

Příklad 4.4. Necht X má Poissonovo rozdělení s parametrem λ . Necht Y_1, \dots, Y_N představují výsledky z minulosti a jsou to nezávislé náhodné veličiny, Y_1 má Poissonovo rozdělení s parametrem λ_j , $\lambda_1, \dots, \lambda_N$ jsou nezávislé náhodné veličiny, λ_j má rozdělení gamma (α, β) , $\alpha > 0$, $\beta > 0$. Pro nepodmíněnou střední hodnotu a nepodmíněný rozptyl náhodné veličiny Y_1 platí

$$EY_1 = E\lambda_1 = \frac{\beta}{\alpha}$$

$$\text{var } Y_1 = E\lambda_1 + E(\lambda_1 \frac{\lambda}{\lambda})^2 = \lambda + \frac{\lambda}{\lambda} = \lambda \left(\frac{\lambda+1}{\lambda} \right)$$

$$\hat{\lambda} = \frac{\bar{Y}}{\hat{\sigma}^2_{-Y}} \quad \hat{\beta} = \frac{\bar{Y}^2}{\hat{\sigma}^2_{-Y}} \quad \text{je-11} \quad \hat{\sigma}^2_{-Y} > 0$$

kde $\hat{\sigma}^2$ je dáno (2.40). Odhad λ při kvadratické ztrátové funkci L_2 dané (4.1) je pak

$$\hat{\lambda} = \frac{\hat{\beta} + \bar{X}}{1 + \hat{\alpha}} \quad \text{je-11} \quad \hat{\sigma}^2_{-Y} > 0.$$

V případě $\hat{\sigma}^2_{-Y} < 0$ není tento odhad vhodný. Můžeme však postupovat jiným způsobem. Při ztrátové funkci L_2 je obecné vyjádření (při libovolné apriorní hustotě $q(\lambda)$ vzhledem k Lebesgueově míře) pro odhad $\tilde{\lambda}$ následující

$$\tilde{\lambda} = \frac{\int_0^{\infty} \lambda r(x|\lambda) q(\lambda) d\lambda}{r(x)} = \frac{(x+1)r(x+1)}{r(x)}.$$

Stačí tedy odhadnout $r(x)$ a $r(x+1)$, např.

$$\hat{r}(x) = \frac{\text{počet } Y_i = x}{N}$$

a dosazením do $\tilde{\lambda}$ dostaneme nový odhad

$$\tilde{\lambda} = \frac{(x+1)\hat{r}(x+1)}{\hat{r}(x)},$$

který je velmi jednoduchý, ale značně nestabilní při menších N .

4.3 BODOVÝ ODHAD: VÍCEROZMĚRNÝ PŘÍPAD

Ve vícerozměrném případě nejmáštěji používáme ztrátovou funkci

$$L_A(\theta, \hat{d}(x)) = (\theta - \hat{d})' A (\theta - \hat{d}),$$

kde A je symetrická pozitivně semidefiniční matice typu $k \times k$ a $\hat{d}(x) = (\hat{d}_1(x), \dots, \hat{d}_k(x))'$ je odhad parametru $\theta = (\theta_1, \dots, \theta_k)'$. Za předpo-

klaďu, že prvky varianční matice $\text{var}\{\bar{X}|\theta\}$ jsou konečné, splňuje bayesovský odhad $\hat{\delta}_A^*(\bar{x})$ parametru θ následující vztah

$$(\hat{\delta}_A^*(\bar{x}) - E(\theta|\bar{X}=\bar{x}))' \hat{A}(\hat{\delta}_A^*(\bar{x}) - E(\theta|\bar{X}=\bar{x})) = 0 \quad (4.13)$$

a bayesovské riziko je rovno

$$\rho_A^*(q) = E\{\text{tr}\{\hat{A} \text{var}(\theta|\bar{X})\}\} \cdot \int r(\bar{X}|\theta) q(\theta) d\lambda(\theta). \quad (4.14)$$

O správnosti se přesvědčíme výpočtem. Pro libovolný odhad $\hat{\delta}(\bar{x})$ dostaneme:

$$\begin{aligned} & \int_{\Theta} (\theta - \hat{\delta}(\bar{x}))' \hat{A}(\theta - \hat{\delta}(\bar{x})) r(\bar{x}|\theta) q(\theta) d\lambda(\theta) = \\ & = \int_{\Theta} \{(\theta - \hat{\delta}_A^*(\bar{x}))' \hat{A}(\theta - \hat{\delta}_A^*(\bar{x})) + 2(\hat{\delta}_A^*(\bar{x}) - \hat{\delta}(\bar{x}))' \hat{A}(\theta - \hat{\delta}_A^*(\bar{x})) + \\ & + (\hat{\delta}_A^*(\bar{x}) - \hat{\delta}(\bar{x}))' \hat{A}(\hat{\delta}_A^*(\bar{x}) - \hat{\delta}(\bar{x}))\} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) = \\ & = \text{tr}\{\hat{A} \text{var}\{\theta|\bar{X}=\bar{x}\}\} + \int_{\Theta} r(\bar{x}|\theta) q(\theta) d\lambda(\theta) \cdot \\ & \cdot (\hat{\delta}_A^*(\bar{x}) - \hat{\delta}(\bar{x}))' \hat{A}(\hat{\delta}_A^*(\bar{x}) - \hat{\delta}(\bar{x})) \geq \text{tr}\{\hat{A} \text{var}\{\theta|\bar{X}=\bar{x}\}\}. \end{aligned}$$

Je-li matice \hat{A} regulární, pak platí

$$\hat{\delta}_A^*(\bar{x}) = E(\theta|\bar{X}=\bar{x}). \quad (4.15)$$

Je-li matice \hat{A} singulární, pak existuje řešení více.

Matici \hat{A} volíme singulární pokud nás zajímají odhady jen některých složek parametru $\theta = (\theta_1, \dots, \theta_k)'$ nebo jejich lineárních kombinací.

Nechť X_1, \dots, X_n jsou nezávislé náhodné veličiny s $N(\mu, \sigma^2)$, kde μ a σ^2 jsou neznámé parametry s a prioriním rozdělením: podmíněná a prioriní rozdělení μ při daném σ^2 je $N(a, r^{-1}\sigma^2)$ a marginální a priori rozdělení σ^2 je gama rozdělení (c, d) . Při ztrátové funkci

$$L(\mu, \sigma^{-2}; d_1(\bar{x}), d_2(\bar{x})) = (\mu - d_1(\bar{x}))^2 + (\sigma^{-2} - d_2(\bar{x}))^2,$$

t.j. v (4.12) klademe $\hat{A} = \hat{I}_2$ a podle (4.15) a (2.15) je bayesovský odhad μ roven

$$\frac{r\alpha + \sum_{i=1}^n X_i}{r+n} \quad (4.16)$$

a podle (4.15), (2.16-2.17) a (A.10) je bayesovský odhad σ^{-2} roven

$$\frac{d+n/2}{c+1/2 \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{rn(\bar{X}-a)^2}{2(r+n)}}.$$

Pedy klademe-li $r = 0$, $c = 0$, $d = -1/2$, dostáváme běžný odhad μ a σ^{-2} uživený v klasické statistice. Zobecněný maximálně věrohodný odhad je shodný s (4.16) a zobecněný maximálně věrohodný odhad σ^{-2} je

$$\frac{d+n/2-1}{c+1/2 \sum_{i=1}^n (X_i - \bar{X})^2 + \frac{rn(\bar{X}-a)^2}{2(r+n)}}.$$

Příklad 4.5. Necht $\tilde{X} = (X_1, \dots, X_k)'$ má multinomické rozdělení s parametry (m, θ) , kde m je známé a $\theta = (\theta_1, \dots, \theta_k)'$ má Dirichletovo rozdělení s parametry $\alpha = (\alpha_1, \dots, \alpha_k)'$. Označme $\alpha_0 = \sum_{i=1}^k \alpha_i$. Úkolem je najít bayesovský odhad parametru θ a bayesovské riziko $\rho^*(q)$ při zkrácené funkci

$$L(\theta, \tilde{d}(\tilde{x})) = \sum_{i=1}^k (\theta_i - d_i(\tilde{x}))^2,$$

kde $\tilde{d}(\tilde{x}) = (d_1(\tilde{x}), \dots, d_k(\tilde{x}))'$. Aposteriorní rozdělení parametru θ je Dirichletovo $(\alpha_1 + x_1, \dots, \alpha_k + x_k)$. Dále platí

$$\begin{aligned} \rho(q, \tilde{d}) &= \sum_{i=1}^k \int_{R_m} \left(\int_{\Theta} (\theta_i - d_i(\tilde{x}))^2 r(\theta | \tilde{x}) d\theta \int_{\mathbb{R}^k} r(x_i | z) q(z) dz \right) \\ &\quad dx_1 \dots dx_n \\ &= \sum_{i=1}^k E\{E((\theta_i - d_i(\tilde{X}))^2 | \tilde{X})\}, \end{aligned} \quad (4.17)$$

kde $\Theta = \{\theta = (\theta_1, \dots, \theta_k)'; 0 \leq \theta_1 < 1, 1=1, \dots, k; \sum_{i=1}^k \theta_i = 1\}$.

Vzhledem k (4.16) bude minima dosaženo pro

$$\begin{aligned} \delta_i^*(\bar{x}) &= (\delta_1^*(\bar{x}), \dots, \delta_k^*(\bar{x}))' \\ \delta_i^*(\bar{x}) &= E(\theta_i | x) = \frac{\alpha_i + x_i}{\alpha_0 + m}, \end{aligned}$$

kde jsme použili faktu, že a posteriorní marginální rozdělení θ_i je beta rozdělení s parametry $(\alpha_i + x_i, \alpha_0 + m - \alpha_i - x_i)$. Odtud zároveň plyne, že

$$\begin{aligned} E((\theta_i - \delta_i^*(\bar{x}))^2 | \bar{X} = \bar{x}) &= E((\theta_i - E(\theta_i | \bar{X} = \bar{x}))^2 | \bar{X} = \bar{x}) = (4.18) \\ &= \frac{(\alpha_i + x_i)(\alpha_0 + m - \alpha_i - x_i)}{(\alpha_0 + m)^2 (\alpha_0 + m + 1)}. \end{aligned}$$

K nalezení bayesovského rizika stačí tedy určit

$$E(\alpha_i + X_i)(\alpha_0 + m - \alpha_i - X_i).$$

Nepodmíněné sdružené rozdělení (X_1, \dots, X_k) (viz (1.7)) má tvar

$$\begin{aligned} \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} \cdot \frac{m!}{x_1! \dots x_k!} \int \dots \int \prod_{i=1}^k \theta_i^{\alpha_i + x_i - 1} d\theta_1 \dots d\theta_k = \\ = \frac{1}{\binom{m + \alpha_0}{m}} \prod_{i=1}^k \binom{\alpha_i + x_i - 1}{\alpha_i - 1} \end{aligned}$$

a tudíž marginální rozdělení X_i je

$$\frac{1}{\binom{m + \alpha_0}{m}} \binom{\alpha_i + x_i - 1}{\alpha_i - 1} \binom{m + \alpha_0 - \alpha_i - x_i - 1}{\alpha_0 - \alpha_i - 1} \quad \text{pro } x_i = 0, \dots, m, \quad \sum_{i=1}^k x_i = m$$

a odtud dále obdržíme

$$E(\alpha_i + X_i)(\alpha_0 + m - \alpha_i - X_i) = \alpha_i (\alpha_0 - \alpha_i) \frac{(\alpha_0 + m + 1)(\alpha_0 + m)}{(\alpha_0 + 1)\alpha_0},$$

což spolu s (4.17), (4.18) implikuje

$$\varphi^*(q) = \frac{\alpha_0^2 - \sum_{i=1}^k \alpha_i^2}{(\alpha_0 + m)(\alpha_0 + 1)k_0}$$

4.4 VĚROHODNOSTNÍ MNOŽINY

Kromě bodového odhadu se při klasickém přístupu k úloze odhadu setkáváme s problémem najít konfidenční množinu pro parametr ϑ . Přesněji řečeno najít borelovskou množinu $D_\alpha(\bar{X}) \subset \Theta$, které s předepsanou pravděpodobností $1 - \alpha$ pokrývá skutečnou hodnotu parametru ϑ , tj. platí pro ni

$$P(\vartheta \in D_\alpha(\bar{X}) | \vartheta) = 1 - \alpha \quad \text{pro v\u0161. } \vartheta \in \Theta \quad (4.19)$$

Při bayesovském přístupu konfidenčním množinám odpovídají tzv. 100(1- α)% věrohodnostní množiny (credible region v angličtině) parametru ϑ . 100(1- α)% věrohodnostní množina parametru ϑ je definovaná jako libovolná množina $C_\alpha(\bar{X}) \subset \Theta$ taková, že

$$P(\vartheta \in C_\alpha(\bar{X}) | \bar{X}) = \int_{C_\alpha(\bar{X})} \pi(\vartheta | \bar{X}) d\lambda(\vartheta) = 1 - \alpha; \quad (4.20)$$

$1 - \alpha$ nazýváme věrohodnost. Někdy nahrazujeme poslední rovnost nerovností \geq . Protože $\pi(\vartheta | \bar{X})$ je hustota na Θ , můžeme mluvit o pravděpodobnosti, že ϑ náleží do $C_\alpha(\bar{X})$. Na rozdíl od klasického přístupu, kdy konfidenční množinu $D_\alpha(\bar{X})$ lze interpretovat jen v termínech pravděpodobnosti pokrytí. Jak uvidíme na příkladech v řadě případů jsou množiny $C_\alpha(\bar{X})$ a $D_\alpha(\bar{X})$ totožné.

Věrohodnostní množina $C_\alpha(\bar{X})$ není obvykle předpisem (4.20) jednoznačně určena. Snažíme se najít množinu $C_\alpha^*(\bar{X})$ takovou, že

$$C_{\alpha}^*(\bar{X}) = \{\theta \in \Theta; \pi(\theta|\bar{X}) \geq K_{\alpha}\}, \quad (4.21)$$

kde K_{α} je největší konstanta taková, že

$$\int_{C_{\alpha}^*(\bar{X})} \pi(\theta|\bar{X}) d\lambda(\theta) = 1 - \alpha \quad (\text{popř. } \geq 1 - \alpha). \quad (4.22)$$

Nařít $C_{\alpha}^*(\bar{X})$ je mnohdy značně obtížné. Poznámenejme, že $C_{\alpha}^*(\bar{X})$ je věro-
hodnostní množina splňující (4.21) a navíc

$$\int_{C_{\alpha}^*(\bar{X})} d\lambda(\theta) \leq \int_{C_{\alpha}(\bar{X})} d\lambda(\theta) \quad \text{pro vř. } C_{\alpha}(\bar{X}). \quad (4.23)$$

Je-li θ jednorozměrný parametr, hledáme obvykle věrohodnou množinu
ve tvaru intervalu a mluvíme o 100(1- α)% věrohodnostním intervalu.

Nechť X_1, \dots, X_n je náhodný výběr z $N(\mu, \sigma^2)$, kde μ a $\sigma^2 > 0$ jsou
neznámé parametry s apriorním rozdělením normálním-gama s parametry
(a, r, c, d), $a \in R_1$, $r > 0$, $c > 0$, $2d$ je přirozené číslo. Tedy podle str.

22 je marginální aposteriorní rozdělení náhodné veličiny $(\mu - \mu^*)$
($d^* r / c^*$)^{1/2}, kde μ^* , c^* , d^* jsou dány, (2.15), (2.16) resp. (2.17),
je t -rozdělení o $2d^*$ stupních volnosti. Z vlastností t -rozdělení ply-
ne, že interval

$$\left(\mu^* - t_{1-\alpha/2} (2d^*) (c^* (d^* r)^{-1})^{1/2}, \mu^* + t_{1-\alpha/2} (d^*) (c^* (d^* r)^{-1})^{1/2} \right), \quad (4.24)$$

kde $t_{1-\alpha/2}(d)$ je 100(1- $\alpha/2$)% kvantil t -rozdělení o d stupních vol-
nosti, je 100(1- α)% věrohodnostní interval a je nejkratší, tj. má
vlastnost (4.23).

Marginální aposteriorní σ^{r-2} je gama rozdělení s parametry (c^*, d^*),
kde c^* je dáno (2.16) a $d^* = d + n/2$. Toto rozdělení je jednovrho-
lové s maximem v bodě

$$\sigma^{r-2} = \frac{d^* - 1}{c^*},$$

ale není symetrické kolem tohoto bodu. Tudiž zkonstruovat věrohod-
nostní interval je numericky velmi obtížné a tak většinou bereme

následující interval za $100(1-\alpha)\%$ věrohodnostní pro σ^{-2} .

$$(H^{-1}(\alpha/2); c^*, d^*), H^{-1}(1-\alpha/2; c^*, d^*) \quad (4.25)$$

kde $H(y; c^*, d^*)$ a $H^{-1}(y; c^*, d^*)$ jsou distribuční resp. kvantilové funkce gamma rozdělení (c^*, d^*) .

Není-li o rozdělení (μ, σ^{-2}) nic známo, můžeme na základě metod vyložených v 2.3 volit za apriorní hustotu q^* danou (2.37). V tomto případě získáváme věrohodnostní intervaly pro μ a σ^{-2} , položíme-li v (4.24) resp. v (4.25) $c^* = \frac{1}{2} S_n^2$, $d^* = (n-1)/2$, kde S_n^2 je dáno (2.38), tj. $100(1-\alpha)\%$ věrohodnostní interval pro μ má tvar

$$(\bar{x} - t_{1-\alpha/2}(n-1)S_n/\sqrt{n}, \bar{x} + t_{1-\alpha/2}(n-1)S_n/\sqrt{n})$$

a $100(1-\alpha)\%$ věrohodnostní interval pro γ^2 má tvar

$$(S_n^2(n-1)/\gamma_{1-\alpha/2}^2(n-1), S_n^2(n-1)/\gamma_{1-\alpha/2}^2(n-1)),$$

kde $\gamma^2(n-1)$ je 100% kvantil χ^2 -rozdělení o $(n-1)$ stupních volnosti. Oba tyto věrohodnostní intervaly jsou shodné s odpovídajícími $100(1-\alpha)\%$ intervaly spolehlivosti.

Příklad 1.1 (pokračování) ^{str. 8, 45}. Chceme-li najít 95% věrohodný interval pro

kvocient inteligence θ , vyjdeme z aposteriorního rozdělení

$N(\frac{100}{325} \cdot 100 + \frac{225}{325} \cdot x; 69, 23)$. Vzhledem k tomu, že jde o jednovrcholové rozdělení s maximem v bodě $\frac{100}{325} \cdot 100 + \frac{225}{325} \cdot x$; bude 95% věrohodný interval s vlastností (4.23), tj. nejkratší interval, mít následující tvar

$$(\frac{100}{325} \cdot 100 + \frac{225}{325} \cdot x - 16, 3; \frac{100}{325} \cdot 100 + \frac{225}{325} \cdot x + 16, 3).$$

Klasický 95% interval spolehlivosti má tvar

$$(x-19,6, x+19,6).$$

Nedy bayesovský interval je kratší.

Příklad 4.4. Při kontrole jakosti výrobků bylo náhodně vybráno a zkontrolováno n výrobků. Předpokládáme, že podíl vadných θ v základním souboru má beta rozdělení (a, b) , $a > 0$, $b > 0$. Chceme nalézt $100(1-\alpha)\%$ věrohodnostní interval. Aposteriorní rozdělení θ je beta rozdělení s parametry $(a+x, b+n-x)$, kde x je počet vadných výrobků ve výběru.

Zkonstruovat nejkratší věrohodnostní interval s předepsanou věrohodností je obtížné z důvodů stejných jako σ^{-2} . Za $100(1-\alpha)\%$ věrohodnostní interval lze vzít interval $(G^{-1}(\frac{\alpha}{2}; a+x, b+n-x), G^{-1}(1-\frac{\alpha}{2}; a+x, b+n-x))$, kde $G^{-1}(y; a+x, b+n-x)$ je kvantilová funkce beta rozdělení s parametry $(a+x, b+n-x)$.

Při $a = 1$, $x = 0$ bereme však spíše interval $(0, 1-\alpha^{\frac{1}{n+b}})$ místo $(1-(1-\alpha/2)^{-1}(b+n)^{-1}, 1-(\alpha/2)^{-1}(b+n)^{-1})$. Je-li $a = 0$, $b = 10$, $n = 10$, $x = 0$, $\alpha = 0,05$, pak je věrohodnostní interval $(0; 0,1391)$.

Příklad 4.5. Počet požárů ve městě za týden se řídí Poissonovým rozdělením s parametrem θ . O hodnotě θ nejsou dostupné žádné informace a tak za apriorní hustotu zvolíme

$$q(\theta) = 1/\theta \quad \theta > 0$$

$$= 0 \quad \theta \leq 0.$$

Jsou-li X_1, \dots, X_n (aspoň jedno X_i je různé od 0) počty požárů v jednotlivých týdnech, pak aposteriorní rozdělení je gama s parametry $(n, \sum_{i=1}^n X_i)$ (je o limitní aposteriorní rozdělení). Úkolem je konstruovat $100(1-\alpha)\%$ interval věrohodnosti pro parametr θ . Za $100(1-\alpha)\%$ věrohodnostní interval obvykle bereme

$$(H^{-1}(\alpha/2, n, \sum_1^n x_i), H^{-1}(1-\alpha/2, n, \sum_1^n x_i)),$$

kde $H^{-1}(x;c,d)$ je kvantilová funkce gama rozdělení s parametry (c,d) .
Je-li $\sum_1^I x_i = 1$, pak za $100(1-\alpha)\%$ věrohodnostní interval bereme
 $(0, H^{-1}(1-\alpha;n,1))$, což při $n = 4$, $\alpha = 0,05$ dává věrohodnostní interval
 $(0;0,7489)$.

5. ТЕСТОВАНИИ ГИПОТЕЗ

5.1 УВОД

Нехť náhodný vektor $\bar{X} = (X_1, \dots, X_n)'$ má hustotu $r(\bar{x}|\vartheta)$ vzhledem k 6-konečné míře γ_n , kde ϑ je parametr náležející do neprázdné borelovské množiny $\Theta \subset R_k$. Parametr ϑ má apriorní hustotu $q(\vartheta)$ vzhledem k 6-konečné míře λ .

Úlohu testovat hypotézu $H_0: \vartheta \in \Theta_0$ proti $H_1: \vartheta \in \Theta - \Theta_0 = \Theta_1$, Θ_1 jsou neprázdné borelovské podmnožiny Θ , lze formulovat jako statistický rozhodovací problém (Θ, Δ, R) , kde Δ je množina rozhodovacích funkcí, které nabývají pouze dvou hodnot d_0, d_1 , kde d_i označuje rozhodnutí, že platí hypotéza H_i (popř. přijmout hypotézu H_i), $i = 0, 1$.

Podobně úlohu diskriminace mezi hypotézami $H_i: \vartheta \in \Theta_i$, $i = 1, \dots, q$, $\Theta_1, \dots, \Theta_q$ jsou neprázdné disjunktí borelovské podmnožiny Θ , $\bigcup_{i=1}^q \Theta_i = \Theta$, můžeme formulovat jako statistický rozhodovací problém (Θ, Δ, R) , kde Δ je množina rozhodovacích funkcí, které nabývají pouze hodnot d_1, \dots, d_k , kde d_i označuje rozhodnutí, že platí hypotéza H_i , $i = 1, \dots, k$.

5.2 ZTRÁTOVÉ FUNKCE POUŽÍVANÉ PŘI TESTOVÁNÍ HYPOTÉZ

Uvedeme si některé nejčastěji používané typy ztrátových funkcí pro úlohu testu hypotézy H_0 proti H_1 (pro úlohu diskriminace se používají obdobně). Typ velice často souvisí s tvarem hypotéz.

Obecně volíme ztrátovou funkci L s vlastnostmi

$$\begin{aligned} L(\vartheta, d_1) &= 0 & \vartheta \in \Theta_1, \quad i = 0, 1 \\ L(\vartheta, d_1) &\gtrsim 0 & \vartheta \notin \Theta_1 \end{aligned}$$

Jsou-li obě hypotézy jednoduché, volíme ztrátovou funkci

$$\begin{aligned} L(\theta_1, d_1) &= 0 & i = 0, 1 \\ L(\theta_0, d_1) &= a_1 \\ L(\theta_1, d_0) &= a_0 \end{aligned}$$

kde $a_1 > 0$, $\theta_1 = \theta_1$, $i = 0, 1$. Z věty 3.3 pak plyne, že se rozhodneme pro platnost hypotézy H_0 , jestliže

$$a_1 \int r(\bar{x} | \theta_0) > a_0 (1 - \int) r(\bar{x} | \theta_1), \quad (5.1)$$

pro platnost hypotézy H_1 , jestliže platí nerovnost opačná a v případě rovnosti se můžeme rozhodnout libovolně.

V obecném případě používáme většinou jeden ze dvou následujících typů ztrátových funkcí:

$$\begin{aligned} L_*(\theta, d_1) &= 0 & \theta \in \mathbb{M}_1 & (5.2) \\ &= a_1 & \theta \notin \mathbb{M}_1 & i = 0, 1, \end{aligned}$$

$$\begin{aligned} L_{**}(\theta, d_1) &= 0 & \theta \in \mathbb{M}_1 & (5.3) \\ &= K_1 d(\theta, \theta_1) & \theta \notin \mathbb{M}_1 & i = 0, 1 \end{aligned}$$

kde $d(\theta, \theta_1)$ je vzdálenost (obvykle Eukleidova) θ od množiny \mathbb{M}_1 , $a_1 \geq 0$, $K_1 \geq 0$, $i = 0, 1$. Zatímco ztrátová funkce L_* závisí pouze na tom, zda θ náleží do \mathbb{M}_0 nebo \mathbb{M}_1 , ztrátová funkce L_{**} nabývá tím větší hodnoty, čím je skutečná hodnota θ vzdálenější od hypotézy, pro kterou jsme se rozhodli.

Nyní se budeme zabývat vlastnostmi těchto ztrátových funkcí. Pro rizikovou funkci odpovídající L_* platí

$$\begin{aligned} R_*(\tilde{\theta}, \delta) &= a_0 P(\delta(\bar{X}) = d_0 / \tilde{\theta}) & \tilde{\theta} \in \mathbb{M}_1 \\ &= a_1 P(\delta(\bar{X}) = d_1 / \tilde{\theta}) & \tilde{\theta} \in \mathbb{M}_0 \end{aligned}$$

a odtud dále obdržíme pro bayesovskou rizikovou funkci

$$q_*(q, \delta) = a_0 \int_{\Theta_1} P(\delta(\bar{X})=d/\theta) q(\theta) d\lambda(\theta) + a_1 \int_{\Theta_0} P(\delta(\bar{X})=d_1/\theta) q(\theta) d\lambda(\theta). \quad (5.4)$$

Pravděpodobnosti $K(\delta(\bar{X})=d_i/\theta)$ pro $\theta \in \Theta_i$, $i = 0, 1$ jsou vlastně pravděpodobnosti chybných rozhodnutí.

Pro libovolnou rozhodovací funkci δ' platí

$$P(\delta'(\bar{X})=d_0/\theta) + P(\delta'(\bar{X})=d_1/\theta) = 1 \text{ pro vš. } \theta \in \Theta \quad (5.5)$$

$$\int_{\Theta_1} L_*(\theta, d_0) \tau(\theta|\bar{x}) d\lambda(\theta) = a_0 P(\theta \in \Theta_1/\bar{X}=\bar{x}) \quad (5.7)$$

$$\int_{\Theta_0} L_*(\theta, d_1) \tau(\theta|\bar{x}) d\lambda(\theta) = a_1 P(\theta \in \Theta_0/\bar{X}=\bar{x}) \quad (5.8)$$

Vzhledem k (3.10) se rozhodneme pro d_0 , jestliže

$$a_0 P(\theta \in \Theta_1/\bar{X}=\bar{x}) < a_1 P(\theta \in \Theta_0/\bar{X}=\bar{x}). \quad (5.9)$$

V případě platnosti opačné nerovnosti se rozhodneme pro d_1 , při rovnosti se můžeme rozhodnout pro d_0 nebo d_1 . Tudiž bayesovská rozhodovací funkce δ^* je vymezená následovně:

$$\delta^*(\bar{x}) = d_0 \quad \text{pro } P(\theta \in \Theta_1/\bar{X}=\bar{x}) < \frac{a_1}{a_0+a_1} \quad (5.10)$$

$$= d_1 \quad \text{pro } P(\theta \in \Theta_1/\bar{X}=\bar{x}) > \frac{a_1}{a_0+a_1}$$

$$= \text{libovolně pro } P(\theta \in \Theta_1/\bar{X}=\bar{x}) = \frac{a_1}{a_0+a_1}.$$

Čísla a_0, a_1 připisujeme hypotézám H_0, H_1 váhy, které mohou odrážet závažnost té které hypotézy.

Příklad 5.1. Doba čekání na autobus na určité zastávce v určitém denní dobu má rovnoměrné rozdělení $(0, \theta)$. Chceme testovat hypotézu

$H_0: \theta \leq d$, kde d je dáno, proti alternativě $H_1: \theta > d$. Ze situace na jiných tratiích plyne, že θ můžeme považovat za náhodnou veličinu s Paretovým rozdělením (a, x_0) .

Jsou-li zjištěné čekací doby X_1, \dots, X_n , je a posteriori rozdělení parametru θ opět Paretovo s parametry $(a+1, \max(x_0, X_1, \dots, X_n))$.

Tedy platí

$$P(\theta \leq d | \bar{X} = \bar{x}) = \int_0^d (1+e) \left(\max_{0 \leq i \leq n} x_i \right)^{a+1} \theta^{-a-2} I\{\theta > \max_{0 \leq i \leq n} x_i\} d\theta = \quad (5.11)$$

$$= 1 - \left(\max_{0 \leq i \leq n} x_i / d \right)^{a+1} \quad \text{pro } d > \max_{0 \leq i \leq n} x_i,$$

$$= 0 \quad \text{pro } d \leq \max_{0 \leq i \leq n} x_i,$$

$$P(\theta > d | \bar{X} = \bar{x}) = \left(\max_{0 \leq i \leq n} x_i / d \right)^{a+1} \quad \text{pro } d > \max_{0 \leq i \leq n} x_i, \quad (5.12)$$

$$= 1 \quad \text{pro } d \leq \max_{0 \leq i \leq n} x_i,$$

kde $\bar{X} = (X_1, \dots, X_n)$.

Při ztrátové funkci

$$L(\theta, d_0) = 0 \quad \text{pro } \theta \leq d$$

$$= K \quad \text{pro } \theta > d$$

$$L(\theta, d_1) = 0 \quad \text{pro } \theta > d$$

$$= K \quad \text{pro } \theta \leq d$$

kde d_1 označuje rozhodnutí, že platí hypotéza H_1 , $1 = 0, 1$; $K > 0$, je bayesovská rozhodovací funkce δ^* následující (viz (5.9)):

$$\delta^*(\bar{x}) = d_1 \quad \text{pro } P(\theta \leq d | \bar{X} = \bar{x}) < P(\theta > d | \bar{X} = \bar{x})$$

$$= d_0 \quad \text{pro } P(\theta \leq d | \bar{X} = \bar{x}) > P(\theta > d | \bar{X} = \bar{x})$$

$$= \text{libovolně} \quad \text{pro } P(\theta \leq d | \bar{X} = \bar{x}) = P(\theta > d | \bar{X} = \bar{x}).$$

Vzhledem k (5.11) a (5.12) se rozhodneme pro d_1 , jestliže

$$d \leq \max_{0 \leq i \leq n} x_i \quad \text{nebo} \quad d > \max_{0 \leq i \leq n} x_i \quad \text{a} \quad \left(\max_{0 \leq i \leq n} x_i / d \right)^{a+1} > 1/2.$$

Máme-li např. $d = 15$, $x_0 = 5$, $a = 3$, $x_1 = 10$, $x_2 = 3$, $x_3 = 2$,
 $x_4 = 5$, $x_5 = 14$, pak

$$\left(\max_{0 \leq i \leq 5} x_i \mid d \right)^{a+1} = (14 \mid 15)^4 > \frac{1}{2}.$$

Rozhodneme se tedy pro d_1 .

Ztrátová funkce L_{**} daná (5.3) je užívána především, je-li jedna z hypotéz jednoduchá a $\textcircled{4}$ je otevřený k -rozměrný interval (konečný nebo nekonečný). Používáme ji též v případě, že θ je jednorozměrný parametr a máme-li hypotézy $H_0: \theta \leq \theta_0$, $H_1: \theta > \theta_0$, nebo je-li $\theta = (\theta_1, \dots, \theta_k)'$ k rozměrný parametr a máme-li hypotézu např. $H_{01}: \theta_1 \leq \theta_0$, $H_{11}: \theta_1 > \theta_0$.

Je-li θ jednorozměrný parametr, $\textcircled{4}$ otevřený interval, $\theta_0 \in \textcircled{4}$,
 $H_0: \theta \leq \theta_0$, $H_1: \theta > \theta_0$ a je-li ztrátová funkce L_{**} dána předpisem

$$\begin{aligned} L_{**}(\theta, d_0) &= 0 & \theta \leq \theta_0 & \\ &= \theta - \theta_0 & \theta > \theta_0 & \\ L_{**}(\theta, d_1) &= \theta_0 - \theta & \theta \leq \theta_0 & \\ &= 0 & \theta > \theta_0 & \end{aligned} \quad (5.13)$$

pak bayesovská rozhodovací funkce je rovna d_0 , jestliže

$$E(\theta \mid \bar{x} = \bar{x}) \leq \theta_0 \quad (5.14)$$

a je d_1 , jestliže platí nerovnost $>$.

Převědčíme se, že tato rozhodovací funkce je bayesovská. Podle (3.10) stačí spočítat pro libovolnou rozhodovací funkci d' podmíněnou střední hodnotu ztrátové funkce:

$$\begin{aligned} E(L_{d'}(\theta, d(\bar{x})) \mid \bar{x} = \bar{x}) &= \int_{-\infty}^{+\infty} (\theta - \theta_0) \tau(\theta \mid \bar{x}) d\lambda(\theta) & \text{pro } d(\bar{x}) = d_0 \\ &= \int_{\theta_0}^{\theta_0} (\theta_0 - \theta) \tau(\theta \mid \bar{x}) d\lambda(\theta) & \text{pro } d(\bar{x}) = d_1. \end{aligned}$$

Podle (3.10) se rozhodneme pro d_0 , jestliže

$$\int_{\theta_0}^{+\infty} (\theta - \theta_0) r(\theta | \bar{x}) d\lambda(\theta) < \int_{-\infty}^{\theta_0} (\theta_0 - \theta) r(\theta | \bar{x}) d\lambda(\theta),$$

což je ekvivalentní

$$E(\theta | \bar{X} = \bar{x}) = \int_{-\infty}^{+\infty} \theta (\theta | \bar{x}) d\lambda(\theta) < \theta_0 \int_{-\infty}^{+\infty} (\theta | \bar{x}) d\lambda(\theta) = \theta_0.$$

Pro bayesovské riziko postupně dostaneme:

$$\begin{aligned} q^*(d) &= \int_{R_n} \int_{-\infty}^{+\infty} L_{**}(\theta, d | \bar{x}) r(\bar{x} | \theta) q(\theta) d\lambda(\theta) = \\ &= \int_{\{x; E(\theta | \bar{X} = \bar{x}) \leq \theta_0\}} \left(\int_{\theta_0}^{+\infty} (\theta - \theta_0) r(\bar{x} | \theta) q(\theta) d\lambda(\theta) \right) d\nu_n(\bar{x}) \\ &\quad - \int_{\{x; E(\theta | \bar{X} = \bar{x}) > \theta_0\}} \left(\int_{-\infty}^{\theta_0} (\theta - \theta_0) r(\bar{x} | \theta) q(\theta) d\lambda(\theta) \right) d\nu_n(\bar{x}) = \\ &= \int_{\{x; E(\theta | \bar{X} = \bar{x}) \leq \theta_0\}} (E(\theta | \bar{X} = \bar{x}) - \theta_0) r(\bar{x}) d\nu_n(\bar{x}) - \int_{-\infty}^{\theta_0} (\theta - \theta_0) q(\theta) d\lambda(\theta). \end{aligned}$$

Příklad 5.1 (pokračování). Při ztrátové funkci

$$\begin{aligned} L_{**}(\theta, d_0) &= 0 && \text{pro } \theta \leq d \\ &= \theta - d && \text{pro } \theta > d \\ L_{**}(\theta, d_1) &= 0 && \text{pro } \theta > d \\ &= d - \theta && \text{pro } \theta \leq d \end{aligned}$$

je bayesovská rozhodovací funkce rovna d_0 , jestliže

$$E(\theta | \bar{X} = \bar{x}) = \frac{\max(x_0, \dots, x_n)}{n-1} \cdot a \leq d.$$

Pro data uvedená v první části příkladu se i při této volbě ztrátové funkce rozhodneme pro d_1 .

5.3 TESTY PŘI $\lambda(\Theta_0) = 0$

Oba typy ztrátových funkcí uvedené v předchozím paragrafu však mají jednu nevýhodu. Je-li $q(\Theta) = 0$ pro sk. vš. (vzhledem k λ) $\Theta \in \Theta_1$, kde $l = 0$ nebo 1, pak bychom se nikdy nemohli rozhodnout pro hypotézu H_1 , neboť nulovost apriorní hustoty implikuje nulovost aposteriorní hustoty. Takováto situace nastane např. je-li $\Theta_1 \in \mathcal{B}_r$, $\Theta \in \mathcal{B}_k$, $r < k$, λ Lebesguova míra a $\lambda(\Theta) > 0$ (u všech apriorních hustot uvažovaných ve 2. kapitole byla λ Lebesguova míra).

V tomto případě se dá postupovat jedním ze dvou způsobů.

Předpokládejme, že $\lambda(\Theta_0) = 0$. Místo hypotéz $H_0: \Theta \in \Theta_0$,

$H_1: \Theta \in \Theta_1$ budeme uvažovat hypotézy $H_0^*: \Theta \in \Theta_0^*$, $H_1^*: \Theta \in \Theta_0^* - \Theta_1^* = \Theta_1^*$, kde $\Theta_0^* \in \mathcal{B}_k$ taková, že $\Theta_0^* \supset \Theta_0$

$$0 < \int_{\Theta_0^*} q(\Theta) d\lambda(\Theta) < 1$$

a míra $\lambda(\Theta_0^* - \Theta_0)$ je z praktického hlediska malá. Je-li např. Θ_0 jednobodová, bereme za Θ_0^* vhodné okolí tohoto bodu.

Příklad 1.1 (pokračování)^{str. 8, 45}. Úlohu testovat hypotézu $H_0: \theta = 100$ proti hypotéze $H_1: \theta \neq 100$ (apriorní pravděpodobnost, že nastane H_0 je rovna nule) nahradíme úlohou testovat hypotézu $H_0^*: \theta \in (100-a, 100+a)$ proti $H_1^*: \theta \notin (100-a, 100+a)$, kde $a > 0$ volíme podle konkrétní situace, např. $a = 5$. Pak při ztrátové funkci (5.2) s $a_0 = a_1$ (vzhledem k (5.10)) se rozhodneme pro H_0 , jestliže

$$1 - \Phi\left(\frac{105 - \frac{225}{325}x - \frac{100^2}{325}}{\sqrt{69,23}}\right) + \Phi\left(\frac{95 - \frac{225}{325}x - \frac{100^2}{325}}{\sqrt{69,23}}\right) < \frac{1}{2}.$$

Je-li např. $x = 110$, pak se rozhodneme pro H_1^* , při $x = 90$ se rozhodneme pro H_1^* .

Při $H_0: \Theta = \Theta_0$ (jednoduchá hypotéza) používáme též jiný postup. Definujeme nové apriorní rozdělení. Bodu Θ_0 přiřadíme apriorní prav-

děpodobnost $q > 0$ a borelovským množinám $B \subset \Theta - \{\theta_0\}$ přiřadíme pravděpodobnost

$$P(\theta \in B) = (1-q) \int_B q(\theta) d\lambda(\theta).$$

Pak aposteriorní rozdělení $\tilde{\theta}$ je

$$P(\tilde{\theta} = \theta_0 | \bar{x} = \bar{x}) = \frac{q r(\bar{x} | \theta_0)}{q r(\bar{x} | \theta_0) + (1-q) \int_{\Theta} r(\bar{x} | \theta) q(\theta) d\lambda(\theta)} \quad (5.15)$$

$$P(\theta \in B | \bar{x} = \bar{x}) = \frac{(1-q) \int_B r(\bar{x} | \theta) q(\theta) d\lambda(\theta)}{q r(\bar{x} | \theta_0) + (1-q) \int_{\Theta} r(\bar{x} | \theta) q(\theta) d\lambda(\theta)}.$$

Marginální hustota X je rovna

$$r^*(\bar{x}) = r(\bar{x} | \theta_0)q + (1-q) \int_{\Theta} r(\bar{x} | \theta) q(\theta) d\lambda(\theta). \quad (5.16)$$

Na $\Theta - \{\theta_0\}$ existují apriorní i aposteriorní hustoty q^* a r^* vzhledem k λ :

$$q^*(\theta) = (1-q)q(\theta) \quad \theta \in \Theta - \{\theta_0\} \quad (5.17)$$

$$r^*(\theta | \bar{x}) = \frac{(1-q)r(\bar{x} | \theta)q(\theta)}{r^*(\bar{x})} \quad \theta \in \Theta - \{\theta_0\} \quad (5.18)$$

Obdobně lze postupovat i v obecnějších případech, např. je-li

$\Theta_0 = \{\theta_0\} \times \Theta_0^*$, $\theta_0 \in R_1$, $\Theta_0^* \in \mathcal{B}_{k-1}$. Tento postup nemusí být vždy vhodný, jak nyní uvidíme.

Lindleyův paradox. Nechť X má rozdělení $N(\theta, \sigma^2)$, $\sigma^2 > 0$ známé,

a testujeme hypotézu $H_0: \theta = \theta_0$ proti $H_1: \theta \neq \theta_0$. Předpokládejme, že apriorní pravděpodobnost, že $\theta = \theta_0$ je $q \in (0, 1)$ a apriorní hustota θ na $R_1 - \{\theta_0\}$ je $(1-q)(2\pi b^2)^{-1/2} \exp\{-(\theta-a)^2/(2b^2)\}^{-1}$, $b^2 > 0$, $a \in R_1$.

Z (2.10) a (5.16) obdržíme pro marginální hustotu X

$$r^*(x) = \frac{q}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\theta_0)^2\right\} + \frac{1-q}{\sqrt{2\pi(\sigma^2+b^2)}} \exp\left\{-\frac{1}{2(b^2+\sigma^2)}(x-a)^2\right\}, \quad x \in R_1 \quad (5.19)$$

Dále z (5.15) plyne, že aposteriorní rozdělení je

$$P(\theta = \theta_0 | X=x) = \frac{q}{\sqrt{2\pi}\sigma^2} \exp\left\{-\frac{1}{2\sigma^2} (x-\theta_0)^2\right\} (r^*(x))^{-1} \quad (5.20)$$

$$r^*(\theta | x) = \frac{1-q}{\sqrt{2\pi}(b^2+\sigma^2)} \exp\left\{-\frac{1}{2(b^2+\sigma^2)} (x-a)^2\right\} (r^*(x))^{-1} \quad (5.21)$$

kde $r^*(x)$ je dáno (5.19).

Je-li ztrátová funkce dána (5.2) s $a_0 = a_1$, pak podle (5.10) je bayesovská rozhodovací funkce rovna d_0 , jestliže

$$P(\theta = \theta_0 | X=x) > 1/2.$$

Dosaďme-li na levé straně z (5.20), dostaneme po jednoduché úpravě ekvivalentní nerovnost

$$\frac{q}{1-q} \cdot \sqrt{\frac{\sigma^2+b^2}{\sigma^2}} > \exp\left\{-\frac{1}{2(b^2+\sigma^2)} (x-a)^2 + \frac{1}{2\sigma^2} (x-\theta_0)^2\right\}.$$

Položme dále pro jednoduchost $b^2 = 1$, $a = \theta_0$, $q = 1/2$, $\sigma^2 = \exp\{-25\}$. Pak je poslední nerovnost ekvivalentní nerovnosti

$$\left| \frac{x-\theta_0}{\sigma} \right| < (1 + e^{-25})^{1/2} (\log(1 + e^{25}))^{1/2},$$

příčemž výraz na pravé straně je větší než 5. Docházíme tedy k paradoxiálnímu závěru. Pro hypotézu H_0 se totiž rozhodneme i v případech

$$\left| \frac{x-\theta_0}{\sigma} \right| \approx 5,$$

zatímco při klasickém přístupu bychom H_0 zamítli i na hladině $5,1 \cdot 10^{-7}$.

Z toho lze soudit, že buď ztrátová funkce nebo apriorní rozdělení nebyly vhodně zvoleny (např. hodnoty blízké θ_0 jsou mnohem pravděpodobnější než hodnoty vzdálenější).

Tento paradox publikovaný Lindleyem vyvolal řadu diskusí, z nichž některé byly publikovány, např. v [8].

Na závěr paragrafu se zmíníme ještě o jednom typu testů, a to testech podílem a posterioriních hustot. Pro test hypotézy $H_0: \theta \in \Theta_0$ proti hypotéze $H_1: \theta \notin \Theta_0$ sestavíme podíl

$$\frac{\sup_{\theta \in \Theta_0} \pi(\theta | \bar{x})}{\sup_{\theta \in \Theta} \pi(\theta | \bar{x})} \quad (5.22)$$

Tento podíl je vždy ≤ 1 . Hodnoty podílu blízké jedné indikují platnost hypotézy H_0 , zatímco malé hodnoty indikují platnost hypotézy

H_1 .

Tato metoda je doporučována, pokud $\int_{\Theta_0} q(\theta) d\lambda(\theta) = 0$ nebo je nula blízký. Neodpovídá obecně žádné ztrátové funkci.

Je analogií testu podílem věrohodností v klasické statistice

Je analogií testu podílem věrohodností v klasické statistice

a používá se též jako v klasické statistice pro test lineárních hypotéz v lineárním modelu. Dosažením (2.30) do (5.22) obdržíme po dalším výpočtu obecný tvar testové statistiky, který je ve speciálním případě roven funkci F-statistiky používané v klasické statistice. Další podrobnosti o této metodě a jejím využití při testech v lineárních modelech může čtenář najít např. v [5], [4].

5.4 TESTY O STŘEDNÍ HODNOTĚ NORMÁLNÍHO ROZDĚLENÍ

Z látky vyložené v předchozích dvou odstavcích můžeme získat řadu testů o střední hodnotě normálního rozdělení. Uvedeme si některé z nich.

A. Nechť X_1, \dots, X_n je náhodný výběr z normálního rozdělení $N(\mu, \sigma_0^2)$, $\sigma_0^2 > 0$ známé a μ je neznámý parametr.

Uvažujme úlohu testovat $H_0: \mu = \mu_0$ proti $H_1: \mu \neq \mu_0$. Předpokládejme, že apriorní rozdělení μ je následující: $P(\mu = \mu_0) = q$ a na $R_1 - \{\mu_0\}$ má hustotu

$$(1-q)(2rb^2)^{-1/2} \exp\{-(\mu-a)^2(2b^2)^{-1}\},$$

$a \in R_1$, $b^2 > 0$. Pak platí

$$\frac{P(\mu = \mu_0 | \bar{X} = \bar{x})}{P(\mu \neq \mu_0 | \bar{X} = \bar{x})} = \frac{q(2r\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2}(\bar{x}_1 - \mu_0)^2(2\sigma_0^2)^{-1}\right\}}{(1-q) \cdot r(\bar{x})} \quad (5.23)$$

kde $r(x)$ je dáno (2.10). Při ztrátové funkci

$$L_*(\mu, d_0) = 0 \quad \mu = \mu_0 \quad (5.24)$$

$$= a_0 \quad \mu \neq \mu_0$$

$$L_*(\mu, d_1) = 0 \quad \mu \neq \mu_0$$

$$= a_1 \quad \mu = \mu_0,$$

kde $a_0 > 0$, $a_1 > 0$, se rozhodneme pro d_0 , jestliže

$$-\frac{n}{nb^2 + \sigma_0^2} (\bar{x} - a)^2 + \frac{n}{\sigma_0^2} (\bar{x} - \mu_0)^2 < 2 \log \left(\frac{a_0}{1-q} \cdot a_1 \right) + \log \frac{nb^2 + \sigma_0^2}{\sigma_0^2}$$

(použili jsme (5.9), (5.23) a (2.10)).

Při ztrátové funkci

$$L_{**}(\mu, d_0) = k_1(\mu - \mu_0)^2 \quad \mu \in R_1 \quad (5.25)$$

$$L_{**}(\mu, d_1) = k_2 \quad \mu = \mu_0$$

$$= 0 \quad \mu \neq \mu_0,$$

kde $k_1 > 0$, $k_2 > 0$, se rozhodneme pro d_0 , jestliže

$$k_1 E(\mu - \mu_0)^2 | \bar{X} = \bar{x} < k_2 P(\mu = \mu_0 | \bar{X} = \bar{x})$$

t.j.

$$k_1(1-q)(\sigma_1^2 + (\mu_1 - \mu_0)^2) < k_2 P(\mu = \mu_0 | \bar{X} = \bar{x}),$$

kde μ_1 a σ_1^2 jsou dány (2.9).

Mějme hypotézu $H_0: \mu \leq \mu_0$ proti $H_1: \mu > \mu_0$ a předpokládáme, že a priori rozdělení μ je $N(a, b^2)$. Při ztrátové funkci

$$I_0(\mu, d_0) = 0 \quad \mu \leq \mu_0 \quad (5.26)$$

$$= \mu - \mu_0 \quad \mu > \mu_0$$

$$I_0(\mu, d_1) = 0 \quad \mu > \mu_0$$

$$= \mu_0 - \mu \quad \mu \leq \mu_0$$

se vzhledem k (5.14) a (2.9) rozhodneme pro d_0 , jestliže

$$\frac{\sum_1 X_i b^2 + a \sigma_0^2}{nb^2 + \sigma_0^2} \leq \mu_0.$$

B. Necht X_1, \dots, X_n je náhodný výběr z $N(\mu, \sigma^2)$, kde μ a $\sigma^2 > 0$ jsou neznámé.

Mějme úlohu testovat $H_0: \mu = \mu_0$ proti $H_1: \mu \neq \mu_0$. Předpokládejme, že apriorní rozdělení (μ, σ^{-2}) je následující: $P(\mu = \mu_0) = q$, podmíněná apriorní hustota σ^{-2} při daném $\mu = \mu_0$ je gama rozdělení (c, d) , podmíněná apriorní hustota (μ, σ^{-2}) při podmínce $\mu \neq \mu_0$ je normální-gama s parametry (a, r, c, d) (hustota je dána (2.18)). Pak pro podmíněnou hustotu $r(\bar{x} | \mu = \mu_0)$ náhodného vektoru \bar{x} při podmínce $\mu = \mu_0$ platí

$$\begin{aligned} r(\bar{x} | \mu = \mu_0) &= \int_0^{+\infty} (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_1 (x_i - \mu_0)^2\right\} \cdot \\ &\cdot \frac{c^d}{\Gamma(d)} (\sigma^{-2})^{d-1} \exp\{-c/\sigma^2\} d\sigma^{-2} = \\ &= (2\pi)^{-n/2} \frac{c^d}{\Gamma(d)} \cdot \frac{\Gamma(\frac{n}{2} + d)}{(c+1/2 \sum_1 (x_i - \mu_0)^2)^{n/2+d}} \end{aligned}$$

a tedy podle (5.15)

$$P(\mu = \mu_0 | \bar{x} = \bar{x}) = \frac{q r(\bar{x} | \mu = \mu_0)}{q r(\bar{x} | \mu = \mu_0) + (1-q) r(\bar{x})},$$

kde $r(\bar{x})$ je dáno (2.20). Při ztrátové funkci (5.24), která nezávisí na $\sigma^{-2} > 0$ se pak rozhodneme pro d_0 , jestliže

$$\left(\frac{a_1 q}{a_0(1-q)} \sqrt{\frac{r+\eta}{r}} \right)^{\frac{2}{r+2d}} < \frac{c + \frac{1}{2} \sum_1 (x_i - \mu_0)^2}{c + \frac{1}{2} \sum_1 (x_i - \bar{x})^2 + \frac{r\eta}{2(r+\eta)} (\bar{x} - a)^2}.$$

Nyní uvažujme ztrátovou funkci

$$L_{**}(\mu, \sigma^{-2}; d_0) = k_0 \sigma^{-2} (\mu - \mu_0)^2 \quad \mu \in R_1, \quad \sigma^{-2} > 0$$

$$L_{**}(\mu, \sigma^{-2}; d_1) = k_1 \quad \mu = \mu_0, \quad \sigma^{-2} > 0$$

$$= 0 \quad \mu \neq \mu_0, \quad \sigma^{-2} > 0,$$

kde $k_0 > 0$, $k_1 > 0$. Pro libovolnou rozhodovací funkci $d(\bar{x})$ platí

$$\begin{aligned} E L_{**}(\mu, \sigma^{-2}; d(\bar{x})) | \bar{X} = \bar{x} &= (1-q) k_0 E \{ \sigma^{-2} (\mu - \mu_0)^2 | \bar{X} = \bar{x} \} \text{ je-li } d(\bar{x}) = d_0 \\ &= q k_1 P(\mu = \mu_0 | \bar{X} = \bar{x}) \quad \text{je-li } d(\bar{x}) = d_1. \end{aligned}$$

Využitím (2.15), (2.16) obdržíme

$$\begin{aligned} E(\sigma^{-2} (\mu - \mu_0)^2 | \bar{X} = \bar{x}, \mu \neq \mu_0) &= E(\sigma^{-2} [(\mu_0 - \mu^*)^2 + \sigma^2 (n+r)^{-1}] | \bar{X} = \bar{x}) = \\ &= \frac{n/2+d}{c^*} (\mu_0 - \mu^*)^2 + (n+r)^{-1}, \end{aligned}$$

kde μ^* a c^* je dáno (2.15) resp. (2.16). Odtud plyne, že se rozhodneme pro d_0 , jestliže

$$(1-q) k_0 \left(\frac{n/2+d}{c^*} (\mu_0 - \mu^*)^2 + (n+r)^{-1} \right) < q k_1 P(\mu = \mu_0 | \bar{X} = \bar{x}).$$

Při úloze testovat $H_0: \mu \leq \mu_0$ proti $H_1: \mu > \mu_0$ a ztrátové funkci

(5.26) pro vš. $\sigma^{-2} > 0$ se rozhodneme pro d_0 , jestliže

$$\frac{r a + n \bar{x}}{r + n} \leq \mu_0.$$

(stačí dosadit (2.15) do (5.14)).

C. Uvažujme 2 nezávislé náhodné výběry X_1, \dots, X_n z $N(\mu_1, \sigma^2)$ a Y_1, \dots, Y_{n_2} z $N(\mu_2, \sigma^2)$, $\mu_1, \mu_2, \sigma^2 > 0$ neznámé a úlohu testovat hypoté-

zu $H_0: \mu_1 \leq \mu_2$ proti $H_1: \mu_1 > \mu_2$. Předpokládejme, že apriorní rozdělení $(\mu_1, \mu_2, \sigma^{-2})$ lze popsat následovně. Podmíněné apriorní rozdělení (μ_1, μ_2) při daném σ^{-2} je $N(a_1, a_2), \sigma^2 \begin{pmatrix} r_1^{-1} & 0 \\ 0 & r_2^{-1} \end{pmatrix}$, marginální apriorní rozdělení σ^{-2} je gama s parametry (c, d) . Pak aposteriorní rozdělení $\pi(\mu_1, \mu_2, \sigma^{-2} | \bar{x}, \bar{y})$ parametrů $(\mu_1, \mu_2, \sigma^{-2})$ je následující. Podmíněné aposteriorní rozdělení (μ_1, μ_2) při daném σ^{-2} je $N(a_1^*, a_2^*)$, $\sigma^2 \begin{pmatrix} r_1 + n_1 & 0 \\ 0 & r_2 + n_2 \end{pmatrix}^{-1}$), kde

$$a_1^* = \frac{r_1 a_1 + \sum_{i=1}^{n_1} X_i}{r_1 + n_1}, \quad a_2^* = \frac{r_2 a_2 + \sum_{i=1}^{n_2} Y_i}{r_2 + n_2}.$$

Marginální aposteriorní rozdělení σ^{-2} je gama s parametry (c^*, d^*) ,

kde $d^* = d + (n_1 + n_2)/2$ a

$$c^* = c + \frac{1}{2} \sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \frac{1}{2} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2 + \frac{r_1 n_1 (\bar{X} - a_1)^2}{r_1 + n_1} + \frac{r_2 n_2 (\bar{Y} - a_2)^2}{r_2 + n_2},$$

$$\bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i.$$

Při ztrátové funkci

$$\begin{aligned} L(\mu_1, \mu_2, \sigma^{-2}; d_0) &= 0 & \mu_1 \leq \mu_2, \quad \sigma^{-2} > 0 \\ &= \mu_1 - \mu_2 & \mu_1 > \mu_2, \quad \sigma^{-2} > 0 \\ L(\mu_1, \mu_2, \sigma^{-2}; d_1) &= \mu_2 - \mu_1 & \mu_1 \leq \mu_2, \quad \sigma^{-2} > 0 \\ &= 0 & \mu_1 > \mu_2, \quad \sigma^{-2} > 0 \end{aligned}$$

platí pro lib. rozhodovací funkci $\delta(\bar{x}, \bar{y})$

$$\begin{aligned} E(L(\mu_1, \mu_2, \sigma^{-2}; \delta(\bar{x}, \bar{y})) | \bar{x} = \bar{x}, \bar{y} = \bar{y}) &= \\ &= \iint \sum_{\mu_1 > \mu_2}^{\infty} (\mu_1 - \mu_2) (\mu_1, \mu_2, \sigma^{-2} | \bar{x}, \bar{y}) d\mu_1 d\mu_2 d\sigma^{-2} \text{ pro } \delta(\bar{x}, \bar{y}) = d_0, \\ &= \iint \sum_{\mu_1 \leq \mu_2}^{\infty} (\mu_2 - \mu_1) (\mu_1, \mu_2, \sigma^{-2} | \bar{x}, \bar{y}) d\mu_1 d\mu_2 d\sigma^{-2} \text{ pro } \delta(\bar{x}, \bar{y}) = d_1. \end{aligned}$$

Po úpravě zjistíme, že bayesovská rozhodovací funkce bude rovna d_0 , jestliže

$$E(\mu_1 | \bar{X} = \bar{x}) \leq E(\mu_2 | \bar{X} = \bar{x})$$

tj. jestliže

$$\frac{r_1 a_1 + n_1 \bar{x}}{r_1 + n_1} \leq \frac{r_2 a_2 + n_2 \bar{y}}{r_2 + n_2} .$$

Bayesovská rozhodovací funkce bude rovna d_1 , jestliže platí ne-
rovnost \succ .

str. 8, 45

Příklad 1.1 (pokračování). Uvažujme úlohu diskriminace mezi hypotézami $H_1: \theta \leq 90$, $H_2: 90 < \theta < 110$, $H_3: \theta \geq 110$ (které odpovídají podprůměrné, průměrné a nadprůměrné inteligenci) a předpokládejme, že ztrátová funkce je dána předpisem (d_i označuje rozhodnutí, že H_1 platí):

$$\begin{aligned} L(\theta, d_1) &= 0 & \theta \leq 90 \\ &= \theta - 90 & 90 < \theta < 110 \\ &= 2(\theta - 90) & \theta \geq 110. \end{aligned}$$

$$\begin{aligned} L(\theta, d_2) &= 90 - \theta & \theta \leq 90 \\ &= 0 & 90 < \theta < 110 \\ &= \theta - 90 & \theta \geq 110 \end{aligned}$$

$$\begin{aligned} L(\theta, d_3) &= 2(110 - \theta) & \theta \leq 90 \\ &= 110 - \theta & 90 < \theta < 110 \\ &= 0 & \theta \geq 110. \end{aligned}$$

Nedy ztráta závisí na vzdálenosti od hypotetické množiny, jestliže jde o θ ze "sousední" množiny a na jejím dvojnásobku, jestliže θ náleží ani do hypotetické ani do "sousední" množiny. Příným výpočtem dostaneme pro lib. rozhodovací funkci

$$E(L(\theta, d^*(X)) | X = x) = \int_{90}^{110} (\theta - 90) \pi(\theta | x) d\theta + 2 \int_{110}^{+\infty} (\theta - 90) \pi(\theta | x) d\theta \text{ jestliže } d^*(x) = d_1,$$

$$E(L(\theta, d(x)) | \bar{x} = \bar{x}) = \int_{-\infty}^{90} (90 - \theta) \tau(\theta | x) d\theta + \int_{110}^{+\infty} (\theta - 110) \tau(\theta | x) d\theta \quad \text{jestliže} \\ f(x) = d_2,$$

$$E(L(\theta, d(x)) | \bar{x} = \bar{x}) = \int_{-\infty}^{90} 2(110 - \theta) \tau(\theta | x) d\theta + \int_{90}^{110} (110 - \theta) \tau(\theta | x) d\theta \quad \text{jestliže} \\ f(x) = d_3.$$

Specielně při $x = 115$ platí

$$\begin{aligned} E(L(\theta, d(x)) | x=115) &= 34,32 & \text{při } d(115) &= d_1 \\ &= 3,55 & &= d_2 \\ &= 3,24 & &= d_3. \end{aligned}$$

Tedy bayesovské rozhodnutí je d_3 .

Dále platí, že

$$\begin{aligned} P(\theta \leq 90 | x=115) &= 0,005 \\ P(90 < \theta < 110 | x=115) &= 0,475 \\ P(110 > \theta | x=115) &= 0,520. \end{aligned}$$

Odtud plyne, že při ztrátové funkci

$$\begin{aligned} L^*(\theta, d_1) &= 0 & \theta \in \mathbb{M}_1 \\ &= 1 & \theta \notin \mathbb{M}_1 \quad l = 1, 2, 3 \end{aligned}$$

přičemž $\mathbb{M}_1 = (-\infty, 90)$, $\mathbb{M}_2 = (90, 110)$, $\mathbb{M}_3 = (110, +\infty)$, je bayesovské rozhodnutí opět d_3 .

APPENDIX; PŘEHLED POUŽITÝCH ROZDĚLENÍ

Níže uvedené hustoty jsou hustoty buď vzhledem k číselní míře (pro diskrétní rozdělení) nebo vzhledem k Lebesguově míře (pro spojité rozdělení).

Binomické rozdělení s parametry (m, p) ($m = 1, 2, \dots; p \in \langle 0, 1 \rangle$) má hustotu

$$r(x|m, p) = \binom{m}{x} p^x (1-p)^{m-x} \quad x = 0, 1, \dots, m. \quad (\text{A.1})$$

Platí

$$EX = mp, \quad \text{var } X = mp(1-p), \quad J(p) = \frac{m}{p(1-p)}; \quad (\text{A.2})$$

při $m = 1$ mluvíme o alternativním rozdělení.

Poissonovo rozdělení s parametrem λ ($\lambda > 0$) má hustotu

$$r(x|\lambda) = e^{-\lambda} \lambda^x (x!)^{-1}, \quad x = 0, 1, 2, \dots \quad (\text{A.3})$$

Platí

$$EX = \lambda, \quad \text{var } X = \lambda, \quad J(\lambda) = \lambda^{-1} \quad (\text{A.4})$$

Negativně binomické rozdělení s parametry (a, p) ($a > 0, p \in \langle 0, 1 \rangle$)

má hustotu

$$r(x|a, p) = \binom{a+x-1}{x} p^a (1-p)^x \quad x = 0, 1, 2, \dots \quad (\text{A.5})$$

Platí

$$EX = a(1-p)p^{-1}, \quad \text{var } X = a(1-p)p^{-2}, \quad J(p) = \frac{a}{p^2(1-p)} \quad (\text{A.6})$$

Při $a = 1$ mluvíme o geometrickém rozdělení.

Normální rozdělení s parametry μ, σ^2 (ozn. $N(\mu, \sigma^2)$, $\mu \in R_1, \sigma^2 > 0$) má hustotu

$$r(x|\mu, \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad x \in R_1. \quad (\text{A.7})$$

Platí

$$EX = \mu, \quad \text{var } X = \sigma^2, \quad J(\mu, \sigma) = \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & 2\sigma^{-2} \end{pmatrix}. \quad (\text{A.8})$$

Gamma rozdělení s parametry (a, p) (a > 0), p > 0) má hustotu

$$r(x|a, p) = a^p (\Gamma(p))^{-1} x^{p-1} e^{-ax} \quad x \geq 0, \quad (\text{A.9})$$

$$= 0 \quad x < 0,$$

$$\text{kde } \Gamma(p) = \int_0^{+\infty} t^{p-1} e^{-t} dt.$$

Platí

$$EX = p a^{-1}, \quad \text{var } X = p a^{-2}. \quad (\text{A.10})$$

Při p = 1 mluvíme o exponenciálním rozdělení. Při p = n/2 a a = 1/2 mluvíme o χ^2 -rozdělení (centrálním) o n stupňích volnosti.

Beta rozdělení s parametry (a, b) (a > 0, b > 0) má hustotu

$$r(x|a, b) = (B(a, b))^{-1} x^{a-1} (1-x)^{b-1} \quad x \in (0, 1), \quad (\text{A.11})$$

$$= 0 \quad x \notin (0, 1),$$

$$\text{kde } B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx.$$

Platí

$$EX = a(a+b)^{-1}, \quad \text{var } X = ab \{(a+b)^2(a+b+1)\}^{-1}. \quad (\text{A.12})$$

Rovnoměrné rozdělení s parametry (a, b) (a < b) má hustotu

$$r(x|a, b) = (b-a)^{-1} \quad x \in (a, b) \quad (\text{A.13})$$

$$= 0 \quad x \notin (a, b)$$

Platí

$$EX = (a+b)/2, \quad \text{var } X = (b-a)^2/12. \quad (\text{A.14})$$

Paretovo rozdělení s parametry (a, b) ($a > 0, b > 0$) má hustotu

$$r(x|a, b) = (a/b) (b/x)^{a+1} \quad x \geq b \quad (\text{A.15})$$

$$= 0 \quad x < b.$$

Pro $a > 2$ platí

$$EX = ab(a-1)^{-1}, \quad \text{var } X = ab^2\{(a-1)^2(a-2)\}^{-1} \quad (\text{A.16})$$

Studentovo (t-) rozdělení o n stupních volnosti s parametrem μ

($\mu \in R_1$) má hustotu

$$r(x|\mu, \mu) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi} \Gamma(\frac{n}{2})} \left(1 + \frac{(x-\mu)^2}{n}\right)^{-\frac{n+1}{2}} \quad x \in R_1. \quad (\text{A.17})$$

Má-li náhodná veličina X rozdělení $N(\mu, 1)$, Y rozdělení χ^2 o n stupních volnosti a jsou-li X a Y nezávislé, pak náhodná veličina

$$T = \frac{X}{\sqrt{Y}} \sqrt{n}$$

má t-rozdělení o n stupních volnosti s parametrem μ .

Platí

$$EX = \mu \quad (\text{A.18})$$

a pro $n > 2$ platí

$$\text{var } X = \frac{n}{n-2}.$$

F-rozdělení s n_1 a n_2 stupni volnosti má hustotu

$$r(x|n_1, n_2) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2})} \left(\frac{n_1}{n_2}\right)^{n_1/2} \left(1 + \frac{n_1}{n_2} x\right)^{-\frac{n_1+n_2}{2}} \frac{n_1/2-1}{x} \quad x \in R_1. \quad (\text{A.19})$$

Jsou-li náhodné veličiny X a Y nezávislé a mají-li χ^2 -rozdělení o n_1 resp. n_2 stupních volnosti, má náhodná veličina

$$F = \frac{X/n_1}{Y/n_2}$$

F-rozdělení s n_1 a n_2 stupni volnosti.

Pro $n_2 > 2$ platí

$$EF = \frac{n_2}{n_2 - 2} \quad (\text{A.20})$$

a pro $n_2 > 4$

$$\text{var } F = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 4)(n_2 - 2)^2} \cdot \quad (\text{A.21})$$

Multinomialné rozdělení s parametry (n, p) ($n = 1, 2, \dots; \mathbf{p} = (p_1, \dots, \dots, p_k)'$, $p_i \in (0, 1)$, $i = 1, \dots, k$; $\sum_{i=1}^k p_i = 1$, $k \geq 2$) má hustotu

$$r(\mathbf{x}|n, p) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} \quad \mathbf{x} = (x_1, \dots, x_k)', \quad (\text{A.22})$$

$$x_i = 0, \dots, n, \quad i = 1, \dots, k, \quad \sum_{i=1}^k x_i = n$$

= 0 jinak.

Platí

$$EX_i = np_i, \quad \text{var } X_i = np_i(1-p_i) \quad i = 1, \dots, k, \quad (\text{A.23})$$

$$\text{cov}(X_i, X_j) = -np_i p_j \quad i, j = 1, \dots, k; \quad i \neq j, \quad (\text{A.24})$$

$$j(p_1, \dots, p_{k-1}) = \frac{n}{p_k} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} + n \begin{pmatrix} 1/p_1 & & 0 \\ & \dots & \\ 0 & & 1/p_{k-1} \end{pmatrix}. \quad (\text{A.25})$$

Dirichletovo rozdělení (mnohorozměrné beta rozdělení) s parametry

$\mathbf{\alpha} = (\alpha_1, \dots, \alpha_k)'$ ($\alpha_i > 0$, $i = 1, \dots, k$) má hustotu

$$r(\mathbf{x}|\mathbf{\alpha}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_k)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_k)} x_1^{\alpha_1 - 1} \dots x_k^{\alpha_k - 1} \quad \mathbf{x} = (x_1, \dots, x_k)' \quad (\text{A.26})$$

$$x_i > 0, \quad i = 1, \dots, k, \quad \sum_{i=1}^k x_i = 1$$

= 0 jinak.

Platí

$$EX_i = \frac{\alpha_i}{\alpha_0}, \quad \text{var } X_i = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}, \quad i = 1, \dots, k, \quad (\text{A.27})$$

$$\text{cov}(X_i, X_j) = -\frac{\alpha_1 \alpha_j}{\alpha_0(\alpha_0 + 1)}, \quad j, i = 1, \dots, k; \quad i \neq j, \quad (\text{A.28})$$

$$\text{kde } \alpha_0 = \sum_{i=1}^k \alpha_i.$$

Marginální rozdělení X_i je beta rozdělení s parametry α_i ,

$$\alpha_0 - \alpha_i.$$

k-rozměrné normální rozdělení s parametry μ a Σ (ozn. $N_k(\mu, \Sigma)$) ($\mu \in R_k, \Sigma$ - symmetrická pozitivně definitní matice typu $k \times k$) má hustotu

$$r(\underline{x} | \mu, \Sigma) = (2\pi)^{-n/2} (\det \Sigma)^{-1/2} \exp\left\{-\frac{1}{2}(\underline{x}-\mu)' \Sigma^{-1}(\underline{x}-\mu)\right\}, \quad (\text{A.29})$$

$\underline{x} \in R_k.$

Platí

$$E\underline{x} = \mu, \quad \text{var } \underline{x} = \Sigma. \quad (\text{A.30})$$

k-rozměrné Wishartovo rozdělení (centrální) s n stupni volnosti s parametrickou maticí Σ (Σ - symmetrická pozitivně definitní matice typu $k \times k$) má hustotu

$$r(\underline{x} | n, \Sigma) = c_{k,n} (\det \Sigma)^{-n/2} (\det \underline{x})^{(n-k-1)/2} \cdot \exp\left\{-\frac{1}{2} \text{tr}(\underline{\Sigma}^{-1} \underline{x})\right\} \quad \text{pro vš. } \underline{x} = \begin{pmatrix} x_{11} & \dots & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{k1} & \dots & x_{kk} \end{pmatrix}, \quad (\text{A.31})$$

kde

$$c_{k,n}^{-1} = 2^{nk/2} \pi^{k(k-1)/4} \prod_{j=1}^k \Gamma\left(\frac{n+1-j}{2}\right). \quad (\text{A.32})$$

Je-li $\underline{x}_1, \dots, \underline{x}_n$ náhodný výběr z $N_k(\underline{0}, \Sigma)$, pak náhodná matice

$$\underline{S} = \sum_{i=1}^n \underline{x}_i \underline{x}_i'$$

má k-rozměrné Wishartovo rozdělení s n stupni volnosti s parametrickou maticí Σ . Náhodná veličina

$$\underline{a}' \underline{S} \underline{a} \quad (\underline{a}' \Sigma \underline{a})^{-1},$$

kde $g = (a_1, \dots, a_k)$ je vektor konstant, má χ^2 -rozdělení s n stupni volnosti.

k-rozměrné t-rozdělení (omnitrélní) s n stupni volnosti s parametry

$\mu \in R_k$, Σ - symetrická pozitivně definitní matice typu $k \times k$ má hustotu

$$r(\underline{x}|\mu, \Sigma) = d_{kn}^{-1} (1 + \frac{1}{n}(\underline{x}-\mu)' \Sigma^{-1}(\underline{x}-\mu))^{-\frac{k+n}{2}}, \quad \underline{x} = (x_1, \dots, x_k)' \in R_k \quad (\text{A.33})$$

$$d_{kn}^{-1} = \frac{\Gamma(\frac{n+k}{2}) (\det \Sigma)^{-1/2}}{\Gamma(\frac{n}{2}) (n\pi)^{k/2}}. \quad (\text{A.34})$$

Platí

$$E \underline{X} = \mu \quad (\text{A.35})$$

a pro $n > 2$

$$\text{var } \underline{X} = \frac{\mu}{n-2} \Sigma. \quad (\text{A.36})$$

Nechť má náhodný vektor $\underline{Y} = (Y_1, \dots, Y_k)'$ rozdělení $N_k(0, \Sigma)$, kde Σ je regulární, nechť má náhodná veličina Z χ^2 -rozdělení s n stupni volnosti a \underline{Y} a Z jsou nezávislé. Definujme náhodný vektor $\underline{X} = (X_1, \dots, X_k)'$ předpisem

$$X_1 = \frac{Y_1}{\sqrt{Z}} \sqrt{n} + \mu_1 \quad 1 = 1, \dots, k. \quad (\text{A.37})$$

Pak náhodný vektor \underline{X} má k-rozměrné t-rozdělení s n stupni volnosti a parametry $\mu = (\mu_1, \dots, \mu_k)$ a Σ .

Dvojirozměrné Paretovo rozdělení s parametry (r_1, r_2, a) ($r_1 < r_2$, $a > 0$)

má hustotu

$$r(x_1, x_2 | r_1, r_2, a) = \frac{a(a+1)(r_2 - r_1)^a}{(x_2 - x_1)^{a+2}} \quad (x_1, x_2) \in R_2, \quad (\text{A.34})$$

$$= 0 \quad x_1 < r_1, \quad x_2 < x_1, \quad \text{jinak.}$$

Pro $a > 1$ plati

$$EX_1 = \frac{ar_1 - r_2}{a-1}, \quad EX_2 = \frac{ar_2 - r_1}{a-1} \quad (\text{A.35})$$

a pro $a > 2$

$$VAR X_1 = VAR X_2 = \frac{a(r_2 - r_1)^2}{(a-1)^2(a-2)} \quad (\text{A.36})$$

L I T E R A T U R A

- [1] Anděl, J.: *Matematická statistika*, SNTL, Praha 1978.
- [2] Berger, J. O.: *Statistical decision theory*, New York Inc., Springer-Verlag, 1980.
- [3] Blackwell, D. a Ghshick, M. A.: *Teorie her a statistického rozhodování*, Praha, Academia 1964.
- [4] Box, G. E. P. a Tiao, G. C.: *Bayesian inference in statistical analysis*, Reading, Massachusetts, Addison-Wesley, 1973.
- [5] De Groot, M. H.: *Optimal statistical decisions*, New York, Mc-Graw-Hill Company, 1970 (ruský překlad: *Optimalnyje statističeskije rešenija*, Moskva, Mir, 1974).
- [6] Lindley, D. V.: *Introduction to probability and statistics from a Bayesian viewpoint*, Part 1. Probability, Part 2. Inference, Cambridge, Cambridge University Press, 1965.
- [7] Maritz, J. S.: *Empirical Bayes methods*. London, Methuen and Co, 1970.
- [8] Shafer, G.: *Lindley's paradox*. Technical Report No. 125, Department of Statistics, Stanford University, Stanford 1975.
- [9] Winkler, R. L.: *Introduction to Bayesian inference and decision*. New York, Holt, Rinehart and Winston, Inc., 1972.

20398