# BAYESOVSKÉ METODY

Marie Hušková

UNIVERZITA KARLOVA

O B S A H

# P Ř E D M L U V A

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 skript je vyložit základy bayesovských metod v úlohách testování hypotéz a teorie odhadu. Skriptum je rozděleno do šesti kapitol. V první kapitole jsou vyloženy 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 používaných 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. Anděl: Matematická statistika, kap. 1-10, 13-15. Značení je převzato z této knihy.

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

# 1. Ú V O D

## 1.1 FORMULACE PROBLEMATIKY

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

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

Při bayesovském přístupu k závěrům o parametru $\underset{\sim}{\theta}$ použijeme kromě $\underset{\sim}{X}$ ještě informaci (byť neúplnou) o parametru $\underset{\sim}{\theta}$, kterou máme k dispozici nezávisle na realizaci $\underset{\sim}{X}$. 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 $\underset{\sim}{\theta}$ je náhodný vektor popř. náhodná veličina s rozdělením, které je více či méně známo v závislosti na tom, jak úplnou či neúplnou informaci o $\underset{\sim}{\theta}$ máme.

Výjimečně pracujeme s náhodným $\underset{\sim}{\theta}$ jako metodou získání závěrů o parametru $\underset{\sim}{\theta}$ (náhodnost tedy nevyjadřuje žádnou apriorní informaci, ale slouží jako prostředek jak obdržet závěry o $\underset{\sim}{\theta}$). V kapitole 3 uvidíme, že znáhodnění parametru $\underset{\sim}{\theta}$ je jedna z možností jak zavést uspořádání na množině rozhodovacích funkcí a definovat kritérium optimality.

Příklad 1.1. Uvažujme problém odhadu kvocientu inteligence $\theta$ u určitého dítěte na základě testu s výsledkem X. Dlouholeté výzkumy ukazují, že X má rozdělení $N(\theta,100)$, kde $\theta$ je kvocient inteligence, že $\theta$ 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 $\theta$. Fyzik má určitou představu o možných hodnotách $\theta$. Připouští několik možných hodnot $\theta$, přikládá jim obecně různé váhy (pravděpodobnosti) a tedy považuje je z tohoto hlediska za náhodné veličiny. Různí fyzikové mohou vyjádřit svou představu o hodnotě parametru $\theta$ 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 $\theta$ v jednotlivých příkladech. Zatímco v 1. a 2. příkladě bylo rozdělení parametru $\theta$ získáno z řady (objektivních) měření a $\theta$ lze skutečně pokládat za náhodnou veličinu, ve třetím příkladě rozdělení parametru $\theta$ vyjadřuje "stupeň víry" v jednotlivé hodnoty parametru $\theta$ (popř. víry, že náleží do určité množiny). V prvních dvou příkladech se jedná o běžné rozdělení pravděpodobností s četnostní interpretací tak jak se s ní běžně setkáváme v klasické statistice. Ve třetím příkladě rozdělení pravděpodobností parametru zachycuje "stupeň víry" v určité hodnoty $\theta$, 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 doplňkového jevu $E^c$. Přikládáme-li oběma jevům stejnou šanci, klademe $P(E) = P(E^c) = 1/2$. Přikládáme-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 $\sigma$-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. Přikládáme vlastně váhy možným výsledkům, obvykle říkáme, že ten či onen výsledek je nejpravděpodobnější, méně pravděpodobný či málo pravděpodobný.

Nyní obrátíme pozornost na výhody a nevýhody bayesovského přístupu. 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 $\underset{\sim}{\theta}$ na základě apriorní informace, použití subjektivní apriorní informace a v některých případech připustit, že $\underset{\sim}{\theta}$ je náhodná veličina. Část kritiky je filozofického rázu. Většina statistiků nemá námitek vůči postupu v příkladě 1, kde $\underset{\sim}{\theta}$ lze skutečně považovat za náhodnou veličinu a rozdělení parametru $\underset{\sim}{\theta}$ je konstruováno na základě předchozích objektivních měření. V příkladě 3 však může být obtížné považovat fyzikální konstantu $\theta$ za náhodnou veličinu s určitým rozdělením. Subjektivní informace může výrazně ovlivnit závěry o parametru $\underset{\sim}{\theta}$ (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 $\underset{\sim}{\theta}$ 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 $\underset{\sim}{\theta}$ nezávisí na volbě apriorního rozdělení parametru $\underset{\sim}{\theta}$ (viz Věta 2.1).

## 1.2 BAYESOVA VĚTA A JEJÍ POUŽITÍ

Nechť $\underset{\sim}{\theta} = (\theta_1,\dots,\theta_k)'$ je náhodný vektor s hustotou $q(\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\lambda$ na $(\Theta, \mathcal{B}(\Theta))$, kde $\Theta$ je neprázdná borelovská podmnožina $R_k$, $\mathcal{B}(\Theta)$ označuje borelovské podmnožiny $\Theta$. Nechť $\underset{\sim}{X} = (X_1,\dots,X_n)'$ je náhodný vektor s podmíněnou hustotou $r(\underset{\sim}{x}|\underset{\sim}{\theta})$ při daném $\underset{\sim}{\theta}$ vzhledem k $\sigma$-konečné míře $\nu_n$ na $(R_n, \mathcal{B}_n)$, kde $\mathcal{B}_n$ označuje borelovské podmnožiny $R_n$, tj.

$$P(\underset{\sim}{\theta}\in B, \underset{\sim}{X}\in C) = \int_B \left( \int_C r(\underset{\sim}{x}|\underset{\sim}{\theta})d\nu_n(\underset{\sim}{x})\right)q(\underset{\sim}{\theta})d\lambda(\theta), \tag{1.1}$$

kde B a C jsou libovolné měřitelné množiny.

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

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{x}) = \frac{q(\underset{\sim}{\theta})r(\underset{\sim}{x}|\underset{\sim}{\theta})}{\int_{\Theta} q(\underset{\sim}{\theta})r(\underset{\sim}{x}|\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})} \quad \text{je-li } \int q(\underset{\sim}{\theta})r(\underset{\sim}{x}|\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}) \neq 0,$$

$$= 0 \quad \text{jinak.} \tag{1.2}$$

Důkaz provedeme stejně jako důkaz věty III.3.14 v [1] (str. 54). Ze vztahu (1.1) je vidět, že $r(\underset{\sim}{x}|\underset{\sim}{\theta})q(\underset{\sim}{\theta})$ je sdružená hustota vektoru $(\underset{\sim}{x}', \underset{\sim}{\theta}')'$ vzhledem k $\nu_n \times \lambda$. Proto podle věty III.3.7 v [1] o marginální hustotě je

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

marginální hustota vektoru $\underset{\sim}{X}$. Tvrzení věty nyní plyne z věty III.3.13 v [1]. Q.E.D.

Tato věta má v bayesovských metodách klíčové postavení. Je-li $\underset{\sim}{\theta}$ parametr uvažovaný v minulém paragrafu, nazýváme $q(\underset{\sim}{\theta})$ apriorní hustotou, neboť vyjadřuje informaci o $\theta$ ještě před realizací $\underset{\sim}{X}$. Podmíněnou hustotu $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ parametru $\underset{\sim}{\theta}$ pak nazýváme hustotou aposteriorní, neboť jde o hustotu parametru $\underset{\sim}{\theta}$ po realizaci $\underset{\sim}{X}$. K závěrům o parametru $\underset{\sim}{\theta}$ pak použijeme aposteriorní hustotu $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$, která v sobě zahrnuje jak apriorní informaci o parametru $\underset{\sim}{\theta}$ tak informaci plynoucí z realizace $\underset{\sim}{X}$. Poznamenejme, že při klasickém přístupu použijeme k závěrům o parametru $\underset{\sim}{\theta}$ pouze $r(\underset{\sim}{x}|\underset{\sim}{\theta})$.

Příklad 1.1 (pokračování str. 8). Apriorní hustota parametru $\theta$ je $N(100, 225)$, podmíněné rozdělení $X$ při daném $\theta$ je $N(\theta, 100)$. Pak aposteriorní hustota je $N(\frac{225x+100.100}{225+100}; 69{,}23)$, je-li $x$ hodnota $X$. Kvocient inteligence můžeme odhadnout (viz str. 52) $\theta$ střední hodnotou příslušnou aposteriorní hustotě, tedy hodnotou $\frac{225}{325}.x + \frac{100}{325}.100$. Příslušný rozptyl je 69,23. Zatímco při klasickém přístupu bychom použili jako odhad $\theta$ 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 $h_1(\underset{\sim}{\theta},\underset{\sim}{x})$ a $h_2(\underset{\sim}{x})$ takové, že

$$q(\underset{\sim}{\theta})\ r(\underset{\sim}{x}|\underset{\sim}{\theta}) = h_1(\underset{\sim}{\theta},\underset{\sim}{x})\ h_2(\underset{\sim}{x}), \tag{1.3}$$

pak aposteriorní hustotu $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ můžeme přepsat následovně.

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

Odtud je vidět, že nahradíme-li v (1.2) $q(\underset{\sim}{\theta})$ funkcí $cq(\underset{\sim}{\theta})$, kde $c$ je kladná konstanta, aposteriorní hustota $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ se nezmění. Někdy dokonce za hustotu $q(\underset{\sim}{\theta})$ volíme tzv. <u>nevlastní hustotu</u>, 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 $X_i$.

$$EX_i = E(E(X_i|\underset{\sim}{\theta})) = \int ( \int x_i r(\underset{\sim}{x}|\underset{\sim}{\theta})d\nu(\underset{\sim}{x}))q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}), \tag{1.5}$$

$$\begin{aligned} \operatorname{var} X_i &= E(X_i-E(X_i|\underset{\sim}{\theta}))^2 + \operatorname{var}\{E(X_i|\underset{\sim}{\theta})\} = \\ &= E\operatorname{var}\{X_i|\underset{\sim}{\theta}\} + \operatorname{var}\{E(X_i|\underset{\sim}{\theta})\}. \end{aligned} \tag{1.6}$$

Nepodmíněná hustota náhodného vektoru $\underset{\sim}{X}$ vzhledem k $\nu_n$ je

$$r(\underset{\sim}{x}) = \int_{\Theta} q(\underset{\sim}{\theta})r(\underset{\sim}{x}|\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}). \tag{1.7}$$

# 2. VOLBA APRIORNÍHO ROZDĚLENÍ

## 2.1 ÚVOD

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

1. tvar hustoty $q(\underset{\sim}{\theta})$ (včetně hodnoty parametrů) vyplyne z apriorní informace;
2. jako hustotu použijeme histogram (při informaci objektivní $\Theta$ rozložíme na sjednocení disjunktních měřitelných množin a zjistíme četnosti jednotlivých množin; při subjektivní informaci také rozložíme $\Theta$ 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 $q(\underset{\sim}{\theta})$ určitého funkcionálního typu (ať už s parametry známými či neznámými). Běžně se používají jednak tzv. <u>systémy konjugovaných rozdělení</u>, jednak typy rozdělení, které odpovídají tzv. <u>principu neurčitosti</u>. Konjugovaným systémem rozdělení rozumíme takový systém, že apriorní i aposteriorní hustota do něj patří. Těmto systémům je věnován § 2.2. Princip neurčitosti používáme, pokud nemáme žádnou informaci o $\underset{\sim}{\theta}$.

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

Někdy máme apriorní informaci v následujícím tvaru: $Y_1,\dots,Y_N$ jsou nezávislé náhodné veličiny, které představují minulé výsledky, $Y_i$ má podmíněnou hustotu $r(y|\theta_i)$, $\theta_i \in \Theta$, $i = 1,\dots,N$; $\theta_1,\dots,\theta_N$ jsou nezávislé náhodné vektory, $\theta_i$ 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 $Y_i$, $i = 1,\dots,N$ nebo také nepodmíněnou střední hodnotu $EY_i$ a nepodmíněný rozptyl var $Y_i$, což nám za jistých předpokladů umožní odhadnout buď $q(\cdot)$ nebo aspoň některé momenty tohoto rozdělení. Takové metody jsou známy pod názvem <u>empirické bayesovské metody</u>. 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.

<u>Věta 2.1.</u> Nechť $\Theta \subset R_k$ je neprázdná borelovská množina. Mějme na borelovských podmnožinách $\Theta$ definovánu nějakou $\sigma$-konečnou míru $\lambda$. Nechť $r(x|\theta)$ je hustota náhodného vektoru $X = (X_1,\dots,X_n)'$ vzhledem k nějaké $\sigma$-konečné míře $\nu_n$ při dané hodnotě $\theta \in \Theta$. Nechť $q(\theta)$ je nezáporná omezená měřitelná funkce na $\Theta$. Předpokládejme, že pro daný $x \in R_n$ platí

$$0 < \int_\Theta r(x|\theta)d\lambda(\theta) < +\infty, \qquad 0 < \int_\Theta r(x|\theta)q(\theta)d\lambda(\theta) < +\infty.$$

Označme

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

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

$$\int_A \pi_0(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta}) > 1 - a, \tag{2.2}$$

$$m = \inf_{\underset{\sim}{\theta}\in A} q(\underset{\sim}{\theta}) > 0, \tag{2.3}$$

$$\sup_{\underset{\sim}{\theta}\in A} q(\underset{\sim}{\theta}) \leq (1+b)m,$$

$$\sup_{\underset{\sim}{\theta}\in \Theta - A} q(\underset{\sim}{\theta}) \leq (1+c)m.$$

Pak platí

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

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

Věta říká, že aposteriorní hustoty $\pi_0(\underset{\sim}{\theta}|\underset{\sim}{x})$ a $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ se nebudou příliš lišit, jestliže a a b budou dostatečně malá nezáporná čísla a c nebude příliš velké. Uvědomíme-li si význam čísel a,b,c, předpoklady znamenají, že na množině A je koncentrována velká část pravděpodobnosti odpovídající $\pi_0(\underset{\sim}{\theta}|\underset{\sim}{x})$, $q(\underset{\sim}{\theta})$ musí být na A prakticky konstantní a nenulová a omezená na $\Theta - A$. Je-li $X_1,\dots,X_n$ náhodný výběr z rozdělení $r(x|\underset{\sim}{\theta})$, pak při dostatečně velkém n je často aposteriorní hustota $\pi_0(\underset{\sim}{\theta}|\underset{\sim}{x})$ 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 $q(\underset{\sim}{\theta})$ splňuje (2.3), pak většinou stačí, aby $q(\underset{\sim}{\theta})$ byla hladká na A a omezená na $\Theta - A$.

## 2.2 KONJUGOVANÉ SYSTÉMY HUSTOT

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

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

Systém Q apriorních hustot $q(\theta)$ nazveme <u>systémem konjugovaným s hustotami $\{r(x|\theta), \theta \in \Theta\}$,</u> jestliže při dost velkém n a při libovolných hodnotách $\underset{\sim}{X} = \underset{\sim}{x}$, které splňují

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

patří aposteriorní hustota do systému Q.

Vezmeme-li za Q systém všech hustot vzhledem k $\sigma$-konečné míře $\lambda$, jde samozřejmě o systém hustot konjugovaných s $\{r(x|\underset{\sim}{\theta});\ \underset{\sim}{\theta} \in \Theta\}$. 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 postačující statistiky. Označme $T_n(X_1,\ldots,X_n)$ postačující statistiku příslušející systému hustot $\left\{\prod_{i=1}^{n} r(x_i|\underset{\sim}{\theta});\ \underset{\sim}{\theta} \in \Theta\right\}$. Předpokládejme, že existuje $n_0$ takové, že pro vš. $n > n_0$ existují nezáporné funkce $g_n$ a $h_n$ takové, že

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

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

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

Pak systém hustot

$$\{f_{n,\underset{\sim}{t}}(\underset{\sim}{\theta});\ \underset{\sim}{t} \in S_n,\ n > n_0\} \qquad (2.5)$$

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

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

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 $r(x|\underset{\sim}{\theta})$. Nechť $m > n_0$, $n > n_0$. Sdružená hustota obou výběrů má tvar

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

takže

$$g_{n+m}(T_{n+m}(\underset{\sim}{x},\underset{\sim}{y}),\underset{\sim}{\theta})h_{n+m}(\underset{\sim}{x},\underset{\sim}{y}) = g_n(T_n(\underset{\sim}{x}),\underset{\sim}{\theta})g_m(T_m(\underset{\sim}{y}),\underset{\sim}{\theta}).h_n(\underset{\sim}{x})h_n(\underset{\sim}{y}). \tag{2.7}$$

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

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{x}) = k \prod_{i=1}^{n} r(x_i|\underset{\sim}{\theta})\; f_{m,t}(\underset{\sim}{\theta}),$$

kde k je normující konstanta. Odtud dále dostaneme

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{x}) = k\; g_n(T_n(\underset{\sim}{x}),\underset{\sim}{\theta})h_n(\underset{\sim}{x})g_m(\underset{\sim}{t},\underset{\sim}{\theta})\Big(\int_{\Theta} g_m(\underset{\sim}{t},\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})\Big)^{-1} =$$

$$= \frac{k\; g_n(T_n(\underset{\sim}{x}))g_m(T_m(\underset{\sim}{y}))h_n(\underset{\sim}{x})h_m(\underset{\sim}{y})}{h_m(\underset{\sim}{y})\int_{\Theta} g_m(\underset{\sim}{t},\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})}.$$

Tedy z (2.7) vyplývá

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{x}) = c\; g_{n+m}(\underset{\sim}{u},\underset{\sim}{\theta})h_{n+m}(\underset{\sim}{x},\underset{\sim}{y}),$$

kde $\underset{\sim}{u} = T_{n+m}(\underset{\sim}{x},\underset{\sim}{y})$, $c = \dfrac{k}{\int_{\Theta} g_m(\underset{\sim}{t},\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})h_m(\underset{\sim}{y})}$.

Odtud je již vidět, že

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{x}) = g_{n+m}(\underset{\sim}{u},\underset{\sim}{\theta})\;\Big(\int_{\Theta} g_{n+m}(\underset{\sim}{u},\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})\Big)^{-1}, \tag{2.8}$$

kde $\underset{\sim}{u} \in S_{n+m}$. Tím jsme dokázali, že $\pi$ patří do systému (2.5).

Systém hustot (2.5) někdy nazýváme <u>přirozený konjugovaný systém</u>. 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 $\{r(x|\theta);\ \theta \in \Theta\}$. Podrobně probereme případ binomického a normálního (jedno- i vícerozměrného) rozdělení.

## Jednorozměrná rozdělení

Binomické rozdělení s parametry $(m,\theta)$; $m$ 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

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

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

Odtud a z (2.6) plyne

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

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

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

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

Je-li apriorní rozdělení beta rozdělení s parametry $(\alpha,\beta)$, je aposteriorní rozdělení beta rozdělení s parametry $(\alpha + \sum_{i=1}^{n} X_i, \beta + mn - \sum_{i=1}^{n} X_i)$. Marginální rozdělení $(X_1,\dots,X_n)$ je podle (1.7)

$$P(X_1=x_1,\dots,X_n=x_n) = \frac{B(\alpha + \sum x_i, \beta + mn - \sum x_i)}{B(\alpha,\beta)}, \quad \begin{array}{l} x_i = 0,\dots,m \\ i = 1,\dots,n. \end{array}$$

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

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 $(s,\theta)$; $s$ 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 $(\alpha,\beta)$, $\alpha > 0$, $\beta > 0$.

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

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

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 $\max_{1 \le i \le n} X_i$, 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 - Paretovo rozdělení $(m,t)$ odpovídá aposteriorní hustota - Paretovo 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 $(\min\limits_{1\le i\le n} X_i,\ \max\limits_{1\le i\le n} X_i)$, tedy $S_n = \{(t_1,t_2);\ t_1 < t_2\}$.

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

Obvyklý konjugovaný systém je systém dvourozměrných Paretových rozdělení s parametry $(r_1,r_2,\alpha)$, $r_1 < r_2$, $\alpha > 0$. 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 $\sum\limits_{i=1}^{n} X_i$, $S_n = R_1$.

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

$$\mu_1 = \frac{\sum\limits_i X_i b^2 + a\sigma_0^2}{nb^2 + \sigma_0^2}, \qquad \sigma_1^2 = \frac{b^2\sigma_0^2}{nb^2 + \sigma_0^2}. \tag{2.9}$$

Marginální rozdělení $X_i$ je podle (1.7)

$$r(\underline{x}) = \int_{-\infty}^{+\infty} (2\pi\sigma_0^2)^{-n/2} \exp\left\{\frac{-1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}(2\pi b^2)^{-1/2}\exp\left\{-\frac{1}{2b^2}(\mu-a)^2\right\}d\mu$$

$$= (2\pi)^{-n/2}(nb^2+\sigma_0^2)^{-1/2}(\sigma_0^{-2})^{(n-1)/2}\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2/\sigma_0^2 + n(\bar{x}-a)^2/(nb^2+\sigma_0^2)\right]\right\} \qquad (x_1,\dots,x_n)' \in R_n. \tag{2.10}$$

$$EX_i = E\mu = a \tag{2.11}$$

$$\operatorname{var} X_i = \sigma_0^2 + b^2 \tag{2.12}$$

$$\operatorname{cov}(X_i,X_j) = b^2, \quad i \ne j \tag{2.13}$$

**Normální rozdělení $N(\mu_0, \sigma^2)$**, $\mu_0$ 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 $(t + \sum_i (X_i-\mu_0)^2/2,\ (m+n)/2)$.

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

**Normální rozdělení $N(\mu, \sigma^2)$**, $\mu$ i $\sigma^2$ 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

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

Odtud a z (2.6) plyne

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

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

Přirozený konjugovaný systém pro parametry $(\mu, 1/\sigma^2)$ je systém rozdělení vymezený následovně: podmíněné rozdělení $\mu$ 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 $a \in R_1$, $r > 0$, $c > 0$, $d > 0$. Takovéto 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

$$\mu^* = \frac{ra + n\bar{X}}{n+r} \tag{2.15}$$

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

$$d^* = d + n/2 \tag{2.17}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i .$$

Tedy apriorní sdružená hustota $(\mu, 1/\sigma^2)$ (tj. hustota normálního - gama rozdělení s parametry $(a,r,c,d)$) je

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

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

což implikuje, že marginální hustota $\mu$ je až na násobící konstantu rovna

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

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

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

<u>Vícerozměrná rozdělení</u>

<u>Multinomické rozdělení s parametry $(s,\underset{\sim}{p})$</u>, $\underset{\sim}{p} = (p_1,\dots,p_k)$; s je známé,

$\Theta = \{\underset{\sim}{p} = (p_1,\dots,p_k);\ 0 \le p_i \le 1,\ i = 1,\dots,k,\ \sum_{i=1}^{k} p_i = 1\}.$

Postačující statistika je $\left\{ \sum_{j=1}^{n} X_{ij},\ i = 1,\dots,k \right\}$, tedy $S_n = \{\underset{\sim}{t} = (t_1,\dots,t_k)';\ t_i = 0,1,\dots,sn;\ i = 1,\dots,k;\ \sum_{i=1}^{k} t_i = sn\}$.

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

<u>Vícerozměrné normální rozdělení $N_k(\underset{\sim}{\mu},\underset{\sim}{\Sigma}_0)$</u>, kde $\underset{\sim}{\Sigma}_0$ je známá symetrická pozitivně definitní matice typu $k \times k$, $\Theta = R_k$.

Postačující statistika je $\underset{\sim}{\bar{X}} = (\bar{X}_1,\dots,\bar{X}_k)'$, $\bar{X}_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$. Podle (2.6) a (A.29)

$$g_n(\underset{\sim}{t};\underset{\sim}{\mu}) = \exp\left\{-\frac{n}{2}(\underset{\sim}{\mu}-\underset{\sim}{t})'\underset{\sim}{\Sigma}_0^{-1}(\underset{\sim}{\mu}-\underset{\sim}{t})\right\}$$

$$h_n(\underset{\sim}{x}) = (2\pi)^{-\frac{nk}{2}}(\det\underset{\sim}{\Sigma}_0)^{-n/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(\underset{\sim}{X}_i-\underset{\sim}{\bar{X}})'\underset{\sim}{\Sigma}_0^{-1}(\underset{\sim}{X}_i-\underset{\sim}{\bar{X}})\right\}.$$

Tedy přirozený konjugovaný systém je systém normálních rozdělení $N_k(\underset{\sim}{a},\underset{\sim}{\Sigma}_0/n)$, $\underset{\sim}{a}\in R_k$, $n = 1,2,\dots$ Obvykle bereme systém $N_k(\underset{\sim}{a},b\underset{\sim}{\Sigma}_0)$, $\underset{\sim}{a}\in R_k$, $b > 0$.

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

$$\underset{\sim}{a}^* = (n\underset{\sim}{\Sigma}_0^{-1} + \underset{\sim}{B}^{-1})^{-1}(n\underset{\sim}{\Sigma}_0^{-1}\underset{\sim}{\bar{X}} + \underset{\sim}{B}^{-1}\underset{\sim}{a}) \qquad (2.21)$$

$$\underset{\sim}{B}^* = (n\underset{\sim}{\Sigma}_0^{-1} + \underset{\sim}{B}^{-1})^{-1}. \qquad (2.22)$$

<u>Vícerozměrné normální rozdělení $N_k(\underset{\sim}{\mu}_0,\underset{\sim}{\Sigma})$</u>, $\underset{\sim}{\mu}_0$ je známé, $\underset{\sim}{\Sigma}$ pozitivně definitní symetrická neznámá matice $k \times k$, $\Theta = \{\underset{\sim}{B};$ symetrická pozitivně definitní matice $k \times k\}$.

Postačující statistika je $\underset{\sim}{V} = (v_{ij})_{i,j=1,\dots,k}$,

$$v_{ij} = \sum_{q=1}^{n}(X_{iq}-\mu_{i0})(X_{jq}-\mu_{j0}),$$

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

$$g_n(\underset{\sim}{V};\underset{\sim}{\Sigma}^{-1}) = (\det \underset{\sim}{\Sigma}^{-1})^{n/2} \exp\{-\tfrac{1}{2}\,\mathrm{tr}\{\underset{\sim}{\Sigma}^{-1}\underset{\sim}{V}\}\}$$

$$\int_{\Theta} g_n(\underset{\sim}{V};\underset{\sim}{\Sigma}^{-1})\, d\underset{\sim}{\Sigma}^{-1} = (\det \underset{\sim}{V})^{-(n+k+1)/2}\, c^{-1}_{k,n+k+1}$$

kde $c_{k,n}$ je dáno (A.32), $\mathrm{tr}\{.\}$ označuje stopu matice. Přirozený konjugovaný systém pro $\underset{\sim}{\Sigma}^{-1}$ je systém (centrálních) Wishartových rozdělení s r stupni volnosti a parametrickou maticí $\underset{\sim}{R}$, kde $r = k+1,\dots,\underset{\sim}{R}\in\Theta$ . Je-li apriorní rozdělení Wishartovo s a stupni volnosti a parametrickou maticí $\underset{\sim}{R}$, pak aposteriorní rozdělení je opět Wishartovo s a+n stupni volnosti a parametrickou maticí $\underset{\sim}{R}^*$ splňující

$$\underset{\sim}{R}^{*-1} = \underset{\sim}{R}^{-1} + \underset{\sim}{V} .$$

<u>Vícerozměrné normální rozdělení $N_k(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$</u>, $\underset{\sim}{\mu}$ i $\underset{\sim}{\Sigma} > 0$ neznámé, $\Theta =$ $= \{\underset{\sim}{a},\ \underset{\sim}{B} > 0;\ \underset{\sim}{a}\in R_k,\ \underset{\sim}{B}$ symetrická pozitivně definitní matice typu $k \times k\}$.

Postačující statistika je $(\underset{\sim}{X},\underset{\sim}{S})$, kde

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

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

$$g_n(\underset{\sim}{t},\underset{\sim}{S};\underset{\sim}{\mu},\underset{\sim}{\Sigma}) = (\det \underset{\sim}{\Sigma})^{-n/2} \exp\{-\tfrac{n}{2}(\underset{\sim}{t}-\underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{t}-\underset{\sim}{\mu})\}. \qquad (2.24)$$
$$.\ \exp\{-\mathrm{tr}\,\underset{\sim}{\Sigma}^{-1}\underset{\sim}{S}\}.$$

Odtud plyne, že přirozený konjugovaný systém pro $(\underset{\sim}{\mu},\underset{\sim}{\Sigma}^{-1})$ můžeme popsat následovně: podmíněné rozdělení $\underset{\sim}{\mu}$ při daném $\underset{\sim}{\Sigma}$ je $N_k(\underset{\sim}{a},r^{-1}\underset{\sim}{\Sigma})$, marginální rozdělení $\underset{\sim}{\Sigma}^{-1}$ je k-rozměrné Wishartovo rozdělení s q stupni volnosti a parametrickou maticí $\underset{\sim}{V}$, přičemž $\underset{\sim}{a}\in R_k$, $r=1,2,\dots$;

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

$$(2\pi)^{-k/2} (\det \underset{\sim}{\Sigma})^{-1/2} \exp\left\{- \frac{r}{2} \operatorname{tr}(\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\mu}-\underset{\sim}{a})(\underset{\sim}{\mu}-\underset{\sim}{a})')\right\} r^{k/2}. \quad (2.25)$$

$$\cdot\ c_{kq} ( \det \underset{\sim}{V})^{-q/2} (\det \underset{\sim}{\Sigma})^{-(q-k-1)/2} \exp\left\{- \frac{1}{2} \operatorname{tr}(\underset{\sim}{V}^{-1}\underset{\sim}{\Sigma}^{-1})\right\} .$$

Odtud lze integrací získat marginální rozdělení $\underset{\sim}{\mu}$ . Stačí vlastně vypočítat integrál

$$\int (\det \underset{\sim}{\Sigma})^{-(q-k)/2} \exp\left\{- \frac{1}{2} \operatorname{tr}(\underset{\sim}{\Sigma}^{-1}(r(\underset{\sim}{\mu}-\underset{\sim}{a})(\underset{\sim}{\mu}-\underset{\sim}{a})'+ \underset{\sim}{V}^{-1}))\right\} d\underset{\sim}{\Sigma}^{-1}$$

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

$$(\det(\underset{\sim}{V}^{-1} + r(\underset{\sim}{\mu}-\underset{\sim}{a})(\underset{\sim}{\mu}-\underset{\sim}{a})'))^{-(q+1)/2} =$$

$$=(\det \underset{\sim}{V}^{-1} \cdot (1 + r(\underset{\sim}{\mu}-\underset{\sim}{a})'\underset{\sim}{V}(\underset{\sim}{\mu}-\underset{\sim}{a})))^{-\frac{q+1}{2}}.$$

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

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

$$\underset{\sim}{V}^{*-1} = \underset{\sim}{V}^{-1} + \underset{\sim}{S} + \frac{rn}{r + n} (\underset{\sim}{a}-\bar{\underset{\sim}{X}})(\underset{\sim}{a}-\bar{\underset{\sim}{X}})'. \quad (2.27)$$

K odvození tohoto výsledku lze použít faktu, že aposteriorní hustota je až na násobící konstantu rovna součinu $g_n(\underset{\sim}{t};\underset{\sim}{S};\underset{\sim}{\mu},\underset{\sim}{\Sigma})$ dané (2.24) a pravé strany (A.29).

Je-li $\underset{\sim}{\Sigma} = \sigma^2 \underset{\sim}{I}_k$, kde $\underset{\sim}{I}_k$ je jednotková matice typu $k \times k$, je obvyklý konjugovaný systém pro $(\underset{\sim}{\mu}, \sigma^{-2})$ tvořen hustotami

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

$$\underset{\sim}{\mu} \in R_k,\ \sigma^{-2} > 0, \qquad (2.28)$$

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

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

Apriornímu rozdělení (2.28) odpovídá aposteriorní rozdělení

$$\exp\left\{-\frac{r+n}{2\sigma^2}(\underset{\sim}{\mu}-\underset{\sim}{\mu}^*)'(\underset{\sim}{\mu}-\underset{\sim}{\mu}^*)\right\} ((r+n)(2\pi\sigma^2)^{-1})^{1/2} \,. \qquad (2.30)$$

$$.\ (\sigma^{-2})^{d+n/2-1}\, c^{**\,d+n/2}\, (\Gamma(d+\tfrac{n}{2}))^{-1} \exp\{c^*\sigma^{-2}\},\ \underset{\sim}{\mu} \in R_k,\ \sigma^{-2} > 0,$$

kde $\underset{\sim}{\mu}^*$ je dáno (2.26)

$$c^{**} = c + \frac{1}{2}\sum_i (\underset{\sim}{X}_i-\bar{\underset{\sim}{X}})'(\underset{\sim}{X}_i-\bar{\underset{\sim}{X}}) + \frac{rn}{2(r+n)}(\bar{\underset{\sim}{X}}-\underset{\sim}{a})'(\bar{\underset{\sim}{X}}-\underset{\sim}{a}) \qquad (2.31)$$

## 2.3 PRINCIP NEURČITOSTI, JEFFREYSOVA HUSTOTA, LIMITNÍ APOSTERIORNÍ HUSTOTY

Nechť $r(\underset{\sim}{x}|\underset{\sim}{\theta})$ je podmíněná hustota náhodného vektoru $\underset{\sim}{X}=(X_1,\dots,X_n)'$ při dané hodnotě parametru $\underset{\sim}{\theta} \in \Theta \in \mathcal{B}_k$, $\Theta \neq \emptyset$. Pokládáme-li $\underset{\sim}{\theta}$ za náhodný vektor, o němž víme jen to, že $\underset{\sim}{\theta} \in \Theta$, vzniká problém jak volit apriorní rozdělení. Některým řešením tohoto problému se budeme věnovat v tomto paragrafu.

Podle <u>principu neurčitosti</u> bereme za apriorní rozdělení $\underset{\sim}{\theta}$ rovnoměrné rozdělení na $\Theta$. Příslušnou hustotu budeme značit $q_0(\underset{\sim}{\theta})$ a příslušnou aposteriorní hustotu $\pi_0(\underset{\sim}{\theta}|\underset{\sim}{x})$. Pro $\Theta$ nanejvýš spočetnou půjde o hustotu vzhledem k čítací míře. Je-li Lebesgueova míra $\Theta$ 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 $\Theta$ nekonečná spočetná nebo je-li Lebesgueova míra $\Theta$ nekonečná, je hustota $q_0(\underset{\sim}{\theta})$ 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 $\theta$ žádné informace, volíme apriorní hustotu (vzhledem k Lebesgueově míře) rovnu

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

Příslušná aposteriorní hustota je

$$\pi_0(\theta|\underset{\sim}{x}) = \Big(B\big(\sum_i x_i+1,\ n-\sum_i x_i+1\big)\Big)^{-1}\ \theta^{\sum_i x_i}\ (1-\theta)^{n-\sum_i x_i} \qquad \theta\in(0,1),$$

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

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

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

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

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

$$\operatorname{var} \bar{X} = \operatorname{var} \mu + \operatorname{var} Z$$

$$\operatorname{var} \mu = \operatorname{var} \bar{X} + \operatorname{var} Z,$$

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 $\theta$ 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

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

přísluší aposteriorní hustota gama rozdělení s parametry $(n, \sum_i X_i + 1)$. Definujme nový parametr $\lambda = \theta^{1/2}$, pak apriorní hustota parametru (odpovídající (2.14)) je

$$\begin{aligned} q(\lambda) &= \lambda \qquad \lambda > 0 \\ &= 0 \qquad \lambda \le 0. \end{aligned} \tag{2.33}$$

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

$$\begin{aligned} \pi(\lambda|\underline{x}) &= \frac{2n^{\sum_i x_i + 1}}{\Gamma(\sum x_i + 1)}\, e^{-\lambda^2 n}\, \lambda^{2\sum x_i + 1} \qquad \lambda > 0 \\ &= 0 \qquad \lambda \le 0 \end{aligned}$$

Zatímco apriorní hustotě

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

přísluší aposteriorní hustota

$$\pi^*(\lambda|\underset{\sim}{x}) = \frac{2n^{\sum_i x_i + \frac{1}{2}}}{\Gamma(\sum_i x_i + \frac{1}{2})} e^{-\lambda^2 n} \lambda^{2\sum_i x_i} \qquad \lambda > 0$$
$$= 0 \qquad \lambda \leq 0.$$

Tedy aposteriorní hustoty $\pi$ a $\pi^*$ 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 $\lambda$ je Lebesgueova míra a $\lambda(\Theta) > 0$.

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

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

a) $\Theta$ je neprázdná otevřená množina interval v $R_k$.

b) Množina $M = \{\underset{\sim}{x};\ r(\underset{\sim}{x}|\underset{\sim}{\theta}) > 0\}$ nezávisí na $\underset{\sim}{\theta}$.

c) Pro skoro všechna $\underset{\sim}{x} \in M$ (vzhledem k $\sigma$-konečné míře $\nu_n$) existuje konečná parciální derivace $r_i'(\underset{\sim}{x}|\underset{\sim}{\theta}) = \frac{\partial r(\underset{\sim}{x}|\underset{\sim}{\theta})}{\partial \theta_i}$ $i=1,\dots,k$.

d) Pro každé i a pro vš. $\underset{\sim}{\theta} \in \Theta$ platí $\int_M r_i'(\underset{\sim}{x}|\underset{\sim}{\theta}) d\nu_n(\underset{\sim}{x}) = 0$.

e) Pro každou dvojici (i, j) existuje konečný integrál

$$J_{ij}(\underset{\sim}{\theta}) = \int_M \frac{r_i'(\underset{\sim}{x}|\underset{\sim}{\theta})\, r_j'(\underset{\sim}{x}|\underset{\sim}{\theta})}{r^2(\underset{\sim}{x}|\underset{\sim}{\theta})} r(\underset{\sim}{x}|\underset{\sim}{\theta})\, d\nu_n(\underset{\sim}{x}) .$$

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

Matici $J(\underset{\sim}{\theta})$ nazýváme $\boxed{\text{Fisherovou informační maticí.}}$

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

$$c = \int_{\Theta} r(\underset{\sim}{x}|\underset{\sim}{\theta})(\det J(\underset{\sim}{\theta}))^{1/2}\, d\underset{\sim}{\theta} \tag{2.34}$$

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

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

$$\int_B c^{-1} r(\underset{\sim}{x}|\underset{\sim}{\theta})\ (\det J(\underset{\sim}{\theta}))^{1/2}\, d\underset{\sim}{\theta} = \tag{2.35}$$

$$= \int_{H(B)} c^{-1}\ r^*(\underset{\sim}{x}|\underset{\sim}{\eta})\ (\det J^*(\underset{\sim}{\eta}))^{1/2}\, d\underset{\sim}{\eta}\,.$$

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

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

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

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

Označíme-li

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



(*)

můžeme (2.36) napsat pomocí matic jako

$$J(\underset{\sim}{\theta}) = \underset{\sim}{D}' J^*(\underset{\sim}{\eta})\underset{\sim}{D}.$$

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

Podle věty o substituci v mnohonásobných integrálech platí

$$\int_B c\, r(\underset{\sim}{x}|\underset{\sim}{\theta})\, (\det J(\underset{\sim}{\theta}))^{1/2} d\underset{\sim}{\theta} =$$

$$= \int_{H(B)} c\, r(\underset{\sim}{x}|H^{-1}(\underset{\sim}{\eta}))(\det(\underset{\sim}{D}' J^*(\underset{\sim}{\eta})\underset{\sim}{D}))^{1/2}(\det \underset{\sim}{D})^{-1} d\underset{\sim}{\eta}\, .$$

Tvrzení věty nyní plyne, použijeme-li vztah

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

Q.E.D.

Je vidět, že za předpokladů věty je apriorní hustota parametru $\underset{\sim}{\theta}$ rovna funkci $(\det J(\underset{\sim}{\theta}))^{1/2}$ (nebo jakémukoliv kladnému násobku této funkce) a aposteriorní hustota parametru $\underset{\sim}{\theta}$ je rovna $cr(\underset{\sim}{x}|\underset{\sim}{\theta})\cdot$ $\cdot(\det J(\underset{\sim}{\theta}))^{1/2}$ a že je to pravděpodobnostní hustota. Funkci $k.(\det J(\underset{\sim}{\theta}))^{1/2}$, kde k je libovolné kladné číslo, budeme říkat Jeffreysova apriorní hustota. Z tvrzení věty plyne, že při Jeffreysově volbě apriorní hustoty parametrů $\underset{\sim}{\theta}$ a $\underset{\sim}{\eta}$ 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 $(\frac{1}{2}+\sum_{i=1}^{n} x_i,\ \frac{1}{2}+mn-\sum_{i=1}^{n} x_i)$.

Pro Poissonovo rozdělení s parametrem $\lambda$ má Jeffreysova hustota tvar

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

Je tedy nevlastní 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

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

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

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 $N(\mu_0, \sigma^2)$, $\mu_0$ známé, je Jeffreysova hustota rovna $\sigma^2$ pro $\sigma^{-2} > 0$. Jde o nevlastní hustotu. Aposteriorní hustota $\sigma^{-2}$ je gama rozdělení $(\sum_i (x_i - \mu_0)^2/2, n/2)$.

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

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

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

Další možná volba apriorní hustoty parametru $\theta$ je vyjít z konjugovaného systému hustot (většinou tzv. obvyklého) $\{q(\theta;\lambda);\lambda\in\Lambda\}$ příslušného $\{r(x|\theta);\theta\in\Theta\}$ ($\Lambda$ je většinou otevřená borelovská množina) a volíme přímo aposteriorní hustotu $\pi^*(\theta|x)$, kterou dostaneme jako limitu aposteriorních hustot z $\{q(\theta;\lambda),\lambda\in\Lambda\}$ pro $\lambda$ konvergující k nějakému bodu na hranici množiny $\Lambda$. Formálně odpovídá tento způsob tomu, že za apriorní hustotu $q^*(\theta)$ vezmeme limitu apriorních hustot z $\{q(\theta,\lambda),\lambda\in\Lambda\}$ pro $\lambda$ konvergující k nějakému bodu na hranici množiny $\Lambda$. 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 <u>limitní aposteriorní hustoty</u>.

Uvedeme si několik příkladů:

Pro binomické rozdělení (m,p) je konjugovaný systém systémem beta rozdělení $(\alpha,\beta)$, $\alpha>0$, $\beta>0$. Limita aposteriorního rozdělení pro $\alpha\to0$ a $\beta\to0$ je beta rozdělení $(\sum_{i=1}^{n}x_i, mn-\sum_{i=1}^{n}x_i)$, což odpovídá apriornímu rozdělení

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

Pro Poissonovo rozdělení s parametrem $\theta$ je systém konjugovaných hustot tvořen systémem gama rozdělení $(a,t)$, $a>0$, $t>0$. Limita aposteriorního rozdělení pro $a\to0$, $t\to0$ je gama rozdělení $(n,\sum_{i=1}^{n}x_i)$. Totéž aposteriorní rozdělení dostaneme, jestliže za apriorní rozdělení zvolíme

$$q^*(\theta) = \theta^{-1} \qquad \theta>0$$
$$= 0 \qquad \theta\le0.$$

Stejnou úvahou dospějeme pro negativně binomické rozdělení (s,p) k tomu, že za aposteriorní rozdělení $\pi^*(p|x)$ vezmeme beta rozdělení $(sn, \sum_i x_i)$, které odpovídá apriorní hustotě

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

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

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

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

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

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

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

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

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

$$\frac{\mu - \bar{x}}{\sqrt{S_n^2}} \sqrt{n} ,$$

kde

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

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 $q(\underset{\sim}{\theta})$, jestliže máme k dispozici výsledky z minulosti v následujícím tvaru. $Y_1,\dots,Y_N$ jsou nezávislé náhodné veličiny, $Y_i$ má podmíněnou hustotu $r(y|\underset{\sim}{\theta}_i)$ (vzhledem k $\sigma$-konečné míře $\nu$), $\underset{\sim}{\theta}_i \in \Theta$, $i = 1,\dots,N$, $\underset{\sim}{\theta}_1,\dots,\underset{\sim}{\theta}_N$ jsou nezávislé náhodné vektory, $\underset{\sim}{\theta}_i$ má hustotu $q(\underset{\sim}{\theta})$ (vzhledem k $\sigma$-konečné míře $\lambda$). Tedy nepodmíněné rozdělení $Y_i$ je

$$r_q(y) = \int_{\Theta} r(y|\underset{\sim}{\theta})\, q(\underset{\sim}{\theta})\, d\lambda(\underset{\sim}{\theta}) \tag{2.39}$$

a odtud plyne pro nepodmíněnou střední hodnotu a rozptyl

$$EY_i = E(E(Y_i|\underset{\sim}{\theta})) \tag{2.40}$$

var $E(Y_i|\theta)$

$$\operatorname{var} Y_i = E(\operatorname{var}(Y_i|\underset{\sim}{\theta})) + E(E(Y_i|\underset{\sim}{\theta}) - EY_i)^2. \tag{2.41}$$

za předpokladu konečnosti $EY_i$ resp. var $Y_i$.

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 $Y_i$ (popsané $r(y)$), která vznikne složením variability $\theta$ (popsané $q(\theta)$) a variability $Y_i$ při pevném $\theta_i$ (popsané $r(y|\theta)$) je způsobená hlavně variabilitou $\theta_i$. 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 $\mathrm{var}(Y_i|\theta)$ malá ve srovnání s var $\theta$, pak bereme za odhad $q(\theta)$ funkci $r(y)$ (pro $\theta$ 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 $q(\underset{\sim}{\theta})$ má určitý funkcionální tvar, u kterého neznáme jen parametry, tj. předpokládáme, že apriorní hustota je $q(\underset{\sim}{\theta};\underset{\sim}{\alpha})$, kde $\underset{\sim}{\alpha} = (\alpha_1,\dots,\alpha_s)'$ je vektor konstant, který neznáme, víme jen, že $\underset{\sim}{\alpha} \in \Lambda$, kde $\emptyset \neq \Lambda \in \mathcal{B}_s$. Tedy také marginální hustotu $Y_i$ známe až na vektor $\underset{\sim}{\alpha}$, který můžeme odhadnout pomocí $Y_1,\dots,Y_n$ některou klasickou metodou odhadu např. metodou momentů (viz [1]). Při ní klademe

$$EY^j = \frac{1}{N}\sum_{j=1}^{N} Y_i^j \qquad j = 1,2,\dots,s$$

(za předpokladu konečnosti příslušných momentů) a řešíme vzhledem k $\underset{\sim}{\alpha}$. Řešení označíme $\hat{\underset{\sim}{\alpha}}$. Za apriorní hustotu vezmeme $q(\underset{\sim}{\theta},\hat{\underset{\sim}{\alpha}})$.

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 $r(y|\underset{\sim}{\theta})$ 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, $Y_i$ s rozdělením $N(\theta_i,\sigma_0^2)$ a představují výsledky z minulosti. $\theta_1,\dots,\theta_N$ jsou nezávislé náhodné veličiny. $\theta_i$ má rozdělení $N(\mu_q,\sigma_q^2)$, kde $\mu_q$ a $\sigma_q^2>0$ neznáme. Nepodmíněné rozdělení $Y_i$ je $N(\mu_q,\sigma^2)$, kde $\sigma^2=\sigma_q^2+\sigma_0^2$. Metodou momentů získáme odhady $\hat{\mu}_q$ a $\hat{\sigma}^2$ pro $\mu$ a $\sigma^2$, a to

$$\hat{\mu}_q = \frac{1}{N}\cdot\sum_{i=1}^{N} Y_i = \bar{Y} \tag{2.39}$$

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

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

$$\begin{aligned} \hat{\sigma}_q^2 &= \hat{\sigma}^2-\sigma_0^2 && \text{je-li } \hat{\sigma}^2>\sigma_0^2 \\ &= 0 && \text{je-li } \hat{\sigma}^2\leq\sigma_0^2 \end{aligned} \tag{2.41}$$

(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 $\underset{\sim}{X} = (X_1,\dots,X_n)'$ je náhodný vektor s hustotou $r(\underset{\sim}{x}\,|\,\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\nu_n$ na $(R_n,\mathcal{B}_n)$, $\underset{\sim}{\theta} \in \Theta$ je parametr, $\Theta$ je neprázdná borelovská podmnožina $R_k$. Označme $\mathcal{D}$ množinu možných rozhodnutí (závěrů) o parametru $\underset{\sim}{\theta}$ a $d$ prvek množiny $\mathcal{D}$.

Dále označme $L(\underset{\sim}{\theta},d)$ ztrátovou funkci, která nám udává číselně jakou ztrátu utrpíme, jestliže skutečná hodnota parametru je $\underset{\sim}{\theta}$ 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 $L(\underset{\sim}{\theta},d) > k$ pro vš. $\underset{\sim}{\theta} \in \Theta$ a vš. $d \in \mathcal{D}$. Pokud budeme uvažovat $\sigma$-algebry podmnožin $\Theta$ a $\mathcal{D}$, budeme předpokládat, že L je měřitelná funkce vzhledem k součinové $\sigma$-algebře, přičemž na $R_1$ bereme $\sigma$-algebru borelovských množin $\mathcal{B}_1$.

Dále budeme pracovat s rozhodovacími funkcemi $\delta$, kterými budeme rozumět zobrazení z $R_n$ do $\mathcal{D}$. Jinak řečeno hodnota $\delta(\underset{\sim}{x})$, $\underset{\sim}{x} \in R_n$ představuje rozhodnutí (závěr) o parametru $\theta$, je-li $\underset{\sim}{X} = \underset{\sim}{x}$.
Riziko $R(\underset{\sim}{\theta},\delta)$ příslušné rozhodovací funkci $\delta$, je-li skutečná hodnota parametru $\underset{\sim}{\theta}$, definujeme následovně

$$R(\underset{\sim}{\theta},\delta) = E(L(\underset{\sim}{\theta},\delta(\underset{\sim}{X}))/\underset{\sim}{\theta}) = \int L(\underset{\sim}{\theta},\delta(\underset{\sim}{x}))\, r(\underset{\sim}{x}\,|\,\underset{\sim}{\theta})\, d\nu_n(\underset{\sim}{x}). \qquad (3.1)$$

Značení $E(.\,|\,\underset{\sim}{\theta})$ jsme použili, přestože v právě uvažované situaci je $\underset{\sim}{\theta}$ vektor konstant. Vzorec však budeme používat i při bayesovském přístupu, kde $\underset{\sim}{\theta}$ považujeme za náhodný vektor a použití nepodmíněné střední hodnoty by mohlo vést k nejasnostem. Riziko $R(\underset{\sim}{\theta},\delta)$ představuje střední ztrátu způsobenou volbou rozhodovací funkce $\delta$, je-li skutečná hodnota parametru $\theta$.

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

Funkci $R: \Theta \times \Delta \rightarrow R_1$ nazýváme <u>rizikovou funkcí</u>.

<u>Statistickým rozhodovacím problémem</u> budeme rozumět trojici $(\Theta, \Delta, R)$.

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

$$L_1(\theta, \delta(X)) = (\theta - \delta(X))^2$$

a tedy pro rizikovou funkci máme

$$R_1(\theta, \delta) = E\,((\theta - \delta(X))^2 | \theta).$$

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

$$\begin{aligned} L_2(\theta, d_i) &= 0 \qquad \theta \in \Theta_i \quad i = 0,1 \\ &= a \qquad a \notin \Theta_i \end{aligned} \tag{3.2}$$

a pro rizikovou funkci máme:

$$\begin{aligned} R_2(\theta,\delta) &= a\, E(I\{\delta(X) = d_1\}; \theta) = a \,.\, P(\delta(X) = d_1; \theta) \qquad \theta \in \Theta_0 \\ &= a\, E(I\{\delta(X) = d_0\}; \theta) = a \,.\, P(\delta(X) = d_0; \theta) \qquad \theta \in \Theta_1 \end{aligned} \tag{3.3}$$

kde $I\{A\}$ je indikátor množiny $A$.

Odtud je vidět, že riziková funkce je a násobek pravděpodobnosti chybného rozhodnutí a množina $\{\underset{\sim}{x}; \delta(\underset{\sim}{x})=d_i\}$ je oblast přijetí $H_i$ příslušná rozhodovací funkci $\delta$.

Na množině $\Delta$ rozhodovacích funkcí zavedeme uspořádání a ekvivalenci. Rozhodovací funkce $\delta_1$ je <u>R-lepší</u> (popř. lepší vzhledem k rizikové funkci R) než rozhodovací funkce $\delta_2$, jestliže

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

a jestliže ostrá nerovnost platí pro aspoň jedno $\underset{\sim}{\theta}$.

Rozhodovací funkce $\delta_1$ <u>je R-ekvivalentní</u> rozhodovací funkci $\delta_2$, jestliže

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

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

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

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

Při bayesovském přístupu předpokládáme, že parametr $\underset{\sim}{\theta}$ je náhodný vektor s hustotou $q(\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\lambda$ na $(\Theta,\mathcal{B}(\Theta))$ a definujeme <u>bayesovskou rizikovou funkci</u> následovně

$$\varrho(q,\delta) = E\, R(\underset{\sim}{\theta},\delta) = \tag{3.4}$$

$$= \int_{\Theta} \Big( \int_{R_n} L(\underset{\sim}{\theta},\delta(\underset{\sim}{x}))\, r(\underset{\sim}{x}|\underset{\sim}{\theta})\, d\nu_n(\underset{\sim}{x})\Big) q(\underset{\sim}{\theta})\, d\lambda(\underset{\sim}{\theta}),$$

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

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

Číslo $\varrho^*(q) = \varrho(q,\delta^*)$ nazýváme <u>bayesovské riziko</u>.

Další řešení může poskytnout tzv. <u>invariančnı́ princip</u>. Nechť je systém hustot $\{r(\underset{\sim}{x}|\underset{\sim}{\theta});\theta\in\Theta\}$ invariantní vůči grupě transformací $\mathcal{G}$, tj. pro každé $g\in\mathcal{G}$ a $\underset{\sim}{\theta}\in\Theta$ existuje jediné řešení $\underset{\sim}{\theta}^o\in\Theta$ takové, že $\underset{\sim}{Y} = g(\underset{\sim}{X})$ má hustotu $r(\underset{\sim}{y}|\underset{\sim}{\theta}^o)$ ($\underset{\sim}{X}$ má hustotu $r(\underset{\sim}{x}|\underset{\sim}{\theta})$). Označme

$$\underset{\sim}{\theta}^o = \bar{g}(\underset{\sim}{\theta}).$$

Pak je rozumné uvažovat <u>ztrátové funkce invariantní vůči</u> $\mathcal{G}$, tj. takové ztrátové funkce $L(\underset{\sim}{\theta},d)$, že pro každé $g\in\mathcal{G}$ a $d\in\mathcal{D}$ existuje $d^o\in\mathcal{D}$ takové, že

$$L(\underset{\sim}{\theta},d) = L(\bar{g}(\underset{\sim}{\theta}),d^o) \quad \text{pro vš. } \theta\in\Theta .$$

Označme

$$d^o = \tilde{g}(d).$$

<u>Rozhodovací funkci</u> $\delta(\underset{\sim}{x})$ nazveme <u>invariantní vůči grupě</u> $\mathcal{G}$, jestliže pro vš. $\underset{\sim}{x}\in R_n$ a $g\in\mathcal{G}$ platí

$$\delta(g(\underset{\sim}{x})) = \tilde{g}(\delta(\underset{\sim}{x})).$$

Označme $\Delta_I$ množinu takovýchto rozhodovacích funkcí.
Pro rizikovou funkci můžeme dokázat následující tvrzení:

<u>Věta 3.1.</u> Pro rizikovou funkci $R(\underset{\sim}{\theta},\delta)$ odpovídající invariantní ztrátové funkci L (vůči grupě $\mathcal{G}$) platí

$$R(\underset{\sim}{\theta},\delta) = R(\bar{g}(\underset{\sim}{\theta}),\delta) \quad \text{pro vš. } \underset{\sim}{\theta}\in\Theta, \text{ vš. } g\in\mathcal{G} \text{ a vš. } \delta\in\Delta_I.$$

<u>Důkaz.</u> Použijeme-li postupně invariance ztrátové funkce, invariance rozhodovací funkce a invariance rozdělení $\{r(\underset{\sim}{x}|\underset{\sim}{\theta});\ \underset{\sim}{\theta}\in\Theta\}$, obdržíme

$$\begin{aligned} R(\underset{\sim}{\theta},\delta) &= E(L(\underset{\sim}{\theta},\delta(\underset{\sim}{X}))/\underset{\sim}{\theta}) = \\ &= E(L(\bar{g}(\underset{\sim}{\theta}),\tilde{g}(\delta(\underset{\sim}{X})))/\underset{\sim}{\theta}) = \\ &= E(L(\bar{g}(\underset{\sim}{\theta}),\delta(g(\underset{\sim}{X})))/\underset{\sim}{\theta}) = \\ &= E(L(\bar{g}(\underset{\sim}{\theta}),\delta(\underset{\sim}{X}))/\bar{g}(\underset{\sim}{\theta})) = \\ &= R(\bar{g}(\underset{\sim}{\theta}),\delta(\underset{\sim}{X})). \end{aligned}$$

Q.E.D.

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

Optimální bayesovskou rozhodovací funkcí invariantní vůči $\mathcal{G}$ budeme rozumět rozhodovací funkci $\delta^+$ takovou, že

$$\min_{\delta \in \Delta_I} \varphi(q,\delta) = \varphi(q,\delta^+), \tag{3.6}$$

kde $\varphi$ je definováno (3.4) a odpovídá invariantní ztrátové funkci. V řadě případů vede tento postup k tomu, že riziková funkce $R(\underset{\sim}{\theta},\delta)$ nezávisí na $\underset{\sim}{\theta}$ a optimální bayesovská rozhodovací funkce $\delta^+$ invariantní vůči $\mathcal{G}$ splňuje

$$\min_{\delta \in \Delta_I} R(\underset{\sim}{\theta},\delta) = R(\underset{\sim}{\theta},\delta^+),$$

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

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

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 $d_c \in \mathcal{D}$ takové, že

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

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

$$L(\theta,d) = |\theta-d|^2. \tag{3.7}$$

Pak $\tilde{g}(d) = d - c$ a množina $\Delta_I$ rozhodovacích funkcí invariantních vůči $\mathcal{G}$ je tvořena funkcemi $\delta(\underset{\sim}{x})$ splňujícími

$$\delta(x_1-c,\dots,x_n-c) = \delta(x_1,\dots,x_n)-c \text{ pro vš. } c \in R_1 \text{ a } \underset{\sim}{x} \in R_n.$$

Tedy věty 3.1 implikuje, že riziková funkce nezávisí na $\theta$ a je rovna

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

V dalším se budeme zabývat pouze bayesovským přístupem.

Jiné řešení rozhodovacího problému poskytuje např. <u>princip minimaxu</u>, při kterém za optimální rozhodovací funkci budeme považovat rozhodovací funkci $\delta^0$ s vlastností

$$\min_{\delta \in \Theta} \max_{\underset{\sim}{\theta} \in \Theta} R(\underset{\sim}{\theta},\delta) = \max_{\underset{\sim}{\theta} \in \Theta} R(\underset{\sim}{\theta},\delta^0).$$

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

## 3.2 BAYESOVSKÉ ROZHODOVACÍ FUNKCE

Bayesovská rozhodovací funkce $\delta^*$ definovaná (3.5) závisí na volbě ztrátové funkce a volbě rozdělení parametru $\underset{\sim}{\theta}$. Volbě rozdělení parametru $\underset{\sim}{\theta}$ 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 následujících 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 $\varrho(q,\delta)$ (při bayesovském přístupu) pak pracujeme s tzv. <u>totálním rizikem</u>

$$\varrho(q,\delta,c) = \varrho(q,\delta) + E\, c(\underset{\sim}{\theta},\underset{\sim}{X}),$$

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

Mějme dvě ztrátové funkce $L_1$ a $L_2$ mezi nimiž platí vztah

$$L_1(\underset{\sim}{\theta},d) = a\, L_2(\underset{\sim}{\theta},d) + b$$

pro vš. $(\underset{\sim}{\theta},d) \in \Theta \times \mathcal{D}$ 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 zdola, vede k tomu, že bez újmy na obecnosti můžeme předpokládat nezápornost ztrátové funkce.

Nyní si zformulujeme a dokážeme tvrzení o konkávitě bayesovského rizika $\varrho^*(q)$ jako funkce rozdělení q parametru $\underset{\sim}{\theta}$.

<u>Věta 3.2.</u> Pro libovolné hustoty $q_1, q_2$ parametru $\underset{\sim}{\theta}$ vzhledem k $\sigma$-konečné míře $\lambda$ a pro libovolné $\alpha \in \langle 0,1\rangle$ platí

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

za předpokladu, že příslušná bayesovská rizika existují.

<u>Důkaz.</u> Z definice bayesovské rizikové funkce máme:

$$\varrho(\alpha q_1+(1-\alpha)q_2,\delta) = \alpha\varrho(q_1,\delta) + (1-\alpha)\,\varrho(q_2,\delta)$$

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

$$\inf_{\delta\in\Delta} \varrho(\alpha q_1+(1-\alpha)q_2,\delta) \geq \alpha \inf_{\delta\in\Delta} \varrho(q_1,\delta) + (1-\alpha) \inf_{\delta\in\Delta} \varrho(q_2,\delta).$$

Q.E.D.

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

$$\varrho(q,\delta) = \int \Big\{\Big(\int_{\Theta} q(\underset{\sim}{\theta})r(\underset{\sim}{x}|\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})\Big) \,\cdot \tag{3.8}$$

$$\cdot\, \Big(\int_{\Theta} L(\underset{\sim}{\theta},\delta(\underset{\sim}{x}))\pi(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta})\Big)\Big\}\, d\nu_n(\underset{\sim}{x}),$$

kde $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ je aposteriorní hustota parametru $\underset{\sim}{\theta}$.

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

$$\min_{\delta\in\Delta} \int_{R_n} \left\{ \int_{\Theta} L(\underset{\sim}{\theta},\delta(\underset{\sim}{x}))\pi(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta})\right\} d\nu(\underset{\sim}{x}) = \tag{3.9}$$

$$= \int_{R_n} \left\{ \int_{\Theta} L(\underset{\sim}{\theta},\delta^*(\underset{\sim}{x}))\pi(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta})\right\} d\nu(\underset{\sim}{x}).$$

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

$$\min_{\delta(\underset{\sim}{x})\in\mathcal{D}} \int_{\Theta} L(\underset{\sim}{\theta},\delta(\underset{\sim}{x}))\pi(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta}) = \tag{3.10}$$

$$= \int_{\Theta} L(\underset{\sim}{\theta},\delta^0(\underset{\sim}{x}))\pi(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta}),$$

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

Nahradíme-li nyní v (3.10) aposteriorní hustotu $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ hustotou apriorní $q(\underset{\sim}{\theta})$, dostaneme tzv. <u>bayesovské rozhodnutí bez provedení pokusu</u>, tj. k rozhodnutí použijeme pouze informace plynoucí z apriorního rozdělení.

Povšimněme si, že bayesovská rozhodovací funkce $\delta^*$ je založena na informaci plynoucí jak z $\underset{\sim}{X}$, tak z apriorního rozdělení parametru $\underset{\sim}{\theta}$.

<u>Příklad 1.1</u> (pokračování ze str. 8). Uvažujme problém odhadu kvocientu inteligence $\theta$. Ztrátové funkci

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

odpovídá bayesovská riziková funkce (viz (3.8))

$$\varrho(q,\delta) = \int_{-\infty}^{+\infty}\Big(\int_{-\infty}^{+\infty}(\theta-\delta(x))^2\pi(\theta|x)d\theta\Big(\int_{-\infty}^{+\infty} q(\theta)r(x|\theta)d\theta\Big)\Big)dx,$$

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 vyplyne z úvah v následujícím paragrafu

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

Přímým výpočtem pak dostaneme

$$\varrho(q,\delta^*(x)) = 69,23\int_{-\infty}^{+\infty}(2\pi.325)^{-1/2}\exp\Big\{-\frac{1}{2}.\frac{(x-100)^2}{325}\Big\}dx = $$
$$= 69,23.$$

Zatímco bayesovské rozhodnutí $\delta^+$ bez provedení pokusu je

$$\delta^+ = 100$$

a

$$\varrho(q,\delta^+) = \int(\theta-\delta^+)^2\, q(\theta)\, d\theta = 225\ . \tag{3.12}$$

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

$$f(y) = e^{-y} \qquad y \geqslant 0,$$
$$\quad = 0 \qquad y < 0,$$

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

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

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

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

**klasifikace dle výsledku měření**

| | | $d_1$ | $d_2$ | $d_3$ | $d_4$ |
|---|---|---|---|---|---|
| skutečnost | $d_1$ | 0 | 1 | 1 | 2 |
| | $d_2$ | 1 | 0 | 2 | 2 |
| | $d_3$ | 1 | 2 | 0 | 2 |
| | $d_4$ | 3 | 3 | 3 | 0 |

**Aposteriorní rozdělení $\theta$ je rovnoměrné na $(0,X)$. Dále platí, že**

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

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

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

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

**Tedy 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 $\Theta$ i $\mathcal{D}$ obsahují právě 2 body. Toto je případ, se kterým se nejčastěji setkáváme při testování jednoduché nulové hypotézy proti jednoduché alternativní hypotéze. Předpokládejme, že ztrátová funkce je definována následovně:

$$L(\underset{\sim}{\theta}_i, d_i) = 0, \qquad i = 1,2$$
$$L(\underset{\sim}{\theta}_1, d_2) = a_1,$$
$$L(\underset{\sim}{\theta}_2, d_1) = a_2,$$

kde $a_i > 0$, $\mathcal{D} = \{d_1, d_2\}$, $\Theta = \{\underset{\sim}{\theta}_1, \underset{\sim}{\theta}_2\}$.

Dále předpokládejme, že

$$P(\underset{\sim}{\theta} = \underset{\sim}{\theta}_1) = \xi \;, \; P(\underset{\sim}{\theta} = \underset{\sim}{\theta}_2) = 1 - \xi \;, \; \xi \in (0,1) \text{ dáno.}$$

Pak podle (3.4) pro bayesovskou rizikovou funkci platí

$$\varphi(\xi, \delta) = a_1 \xi P(\delta(\underset{\sim}{X}) = d_2 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_1) + a_2(1-\xi) P(\delta(\underset{\sim}{X}) = d_1 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_2). \qquad (3.13)$$

Pravděpodobnosti $P(\delta(\underset{\sim}{X}) = d_2 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_1)$ a $P(\delta(\underset{\sim}{X}) = d_1 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_2)$ jsou podmíněné pravděpodobnosti chybných rozhodnutí. Najít bayesovskou rozhodovací funkci $\delta^*$ vlastně znamená najít takovou rozhodovací funkci $\delta$, která minimalizuje jistou lineární kombinaci podmíněných pravděpodobností chybných rozhodnutí. Řešení problému je obsaženo v následující větě.

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

$$\begin{aligned} \delta^*(x) &= d_1 \quad \text{je-li} \quad a\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_1) > b\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_2) \\ &= d_2 \quad \text{je-li} \quad a\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_1) < b\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_2) \\ &= \text{libovolně je-li } a\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_1) = b\, r(\underset{\sim}{x} | \underset{\sim}{\theta}_2). \end{aligned}$$

Pak pro libovolnou rozhodovací funkci $\delta : \mathcal{X} \to \mathcal{D}$ platí

$$a\, P(\delta^*(\underset{\sim}{X}) = d_2 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_1) + b\, P(\delta^*(\underset{\sim}{X}) = d_1 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_2) \leq$$

$$\leq a\, P(\delta(\underset{\sim}{X}) = d_2 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_1) + b\, P(\delta(\underset{\sim}{X}) = d_1 | \underset{\sim}{\theta} = \underset{\sim}{\theta}_2).$$

Důkaz. Pro libovolnou rozhodovací funkci $\delta$ platí:

$$a\,P(\delta(\underset{\sim}{X}) = d_2 | \underset{\sim}{\theta}=\underset{\sim}{\theta}_1) + b\,P(\delta(\underset{\sim}{X}) = d_1 | \underset{\sim}{\theta}=\underset{\sim}{\theta}_2) = \tag{3.14}$$

$$= a \int_{\delta(\underset{\sim}{x})=d_2} r(\underset{\sim}{x}|\underset{\sim}{\theta}_1)d\nu_n(\underset{\sim}{x}) + b \int_{\delta(\underset{\sim}{x})=d_1} r(\underset{\sim}{x}|\underset{\sim}{\theta}_2)d\nu_n(\underset{\sim}{x}) =$$

$$= a + \int_{\delta(\underset{\sim}{x})=d_1} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))d\nu_n(\underset{\sim}{x}) =$$

$$= a + \int_{R_n} I\{\delta(\underset{\sim}{x})=d_1, \delta^*(x)=d_1\} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))d\nu_n(\underset{\sim}{x})$$

$$+ \int_{R_n} I\{\delta(\underset{\sim}{x})=d_1, \delta^*(\underset{\sim}{x})=d_2\} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))d\nu_n(\underset{\sim}{x}).$$

Z definice $\delta^*$ plyne

$$\int_{R_n} I\{\delta(\underset{\sim}{x})=d_1, \delta^*(\underset{\sim}{x})=d_1\} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))d\nu_n(\underset{\sim}{x}) \geqslant \tag{3.15}$$

$$\geqslant \int_{R_n} I\{\delta^*(\underset{\sim}{x})=d_1\} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))\, d\nu_n(\underset{\sim}{x})$$

a

$$\int_{R_n} I\{\delta(\underset{\sim}{x})=d_1, \delta^*(\underset{\sim}{x})=d_2\} (-a\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1) + b\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_2))d\nu_n(\underset{\sim}{x}) \geqslant 0 \tag{3.16}$$

Z (3.14-3.16) snadno obdržíme tvrzení věty.
Q.E.D.

Poznamenejme, že při vedlejší podmínce $P(\delta^*(\underset{\sim}{X})=d_1 | \underset{\sim}{\theta}=\underset{\sim}{\theta}_1) = \alpha$ je tvrzení věty shodné s tvrzením Neyman-Pearsonova lemmatu.

# 4. ÚLOHA ODHADU

## 4.1 ÚVOD

Nechť $\underset{\sim}{X} = (X_1,\dots,X_n)'$ má hustotu $r(\underset{\sim}{x}|\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\nu_n$, $\underset{\sim}{\theta} = (\theta_1,\dots,\theta_k)' \in \Theta$ je parametr, $\emptyset \neq \Theta \in \mathcal{B}_k$. Nechť $q(\underset{\sim}{\theta})$ je apriorní hustota parametru $\underset{\sim}{\theta}$ vzhledem k $\sigma$-konečné míře $\lambda$ na $(\Theta,\mathcal{B}(\Theta))$ a $\pi(\underset{\sim}{\theta}|\underset{\sim}{x})$ příslušná aposteriorní hustota.

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

## 4.2 BODOVÝ ODHAD, JEDNOROZMĚRNÝ PŘÍPAD

Typickými ztrátovými funkcemi jsou

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

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

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

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

Označme

$$q_1(\theta) = q(\theta)w(\theta)\left(\int_{\Theta} q(\theta)w(\theta)d\lambda(\theta)\right)^{-1} \quad \text{pro} \int_{\Theta} q(\theta)w(\theta)d\lambda(\theta) \neq 0$$
$$= 0 \quad \text{jinak}, \tag{4.3}$$

$$\pi_1(\theta|\underset{\sim}{x}) = \frac{r(\underset{\sim}{x}|\theta)\; q_1(\theta)}{\int_{\Theta} r(\underset{\sim}{x}|\theta)q_1(\theta)d\lambda(\theta)} \quad \text{pro} \int_{\Theta} q_1(\theta)\; r(\underset{\sim}{x}|\theta)d\lambda(\theta) \neq 0$$
$$= 0 \quad \text{jinak}, \tag{4.4}$$

$$r_1(\underset{\sim}{x}) = \int_{\Theta} r(\underset{\sim}{x}|\theta)\; q_1(\theta)d\lambda(\theta) \qquad \underset{\sim}{x} \in R_n. \tag{4.5}$$

Zřejmě $q_1$ je hustota (pro $\int_{\Theta} q(\theta)w(\theta)d\lambda(\theta) \neq 0$) vzhledem k $\lambda$ a považujeme-li ji za apriorní hustotu, pak příslušná aposteriorní hustota je $\pi_1(\theta|\underset{\sim}{x})$; $r_1(\underset{\sim}{x})$ je marginální hustota $\underset{\sim}{X}$ 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

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

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

$$\delta^*_{2,w}(x) = \frac{\int_{\Theta} \theta w(\theta)\pi(\theta|\underset{\sim}{x})d\lambda(\theta)}{\int_{\Theta} w(\theta)\; \pi(\theta|\underset{\sim}{x})d\lambda(\theta)} = E_1(\theta|\underset{\sim}{X}=\underset{\sim}{x}) \quad \text{pro} \int_{\Theta} w(\theta)\pi(\theta|\underset{\sim}{x})d\lambda(\theta) \neq 0$$
$$= 0 \quad \text{jinak} \tag{4.6}$$

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

$$\varrho^*_{2,w}(q) = E_1(\mathrm{var}_1(\theta|\underset{\sim}{X}))\;, \tag{4.7}$$

kde $E_1$ označuje střední hodnotu $\underset{\sim}{X}$ vzhledem k hustotě (4.5) a $var_1$ označuje podmíněný rozptyl $\theta$ vzhledem k hustotě (4.4).

b) Nechť

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

pak pro ztrátovou funkci $L_{k_0,k_1}$ definovanou (4.2) je bayesovský odhad $\delta^*_{k_0 k_1}$ parametru $\theta$ $(100(k_0(k_0+k_1)^{-1})\%$ kvantil aposteriorního rozdělení $\pi(\theta|\underset{\sim}{x})$a pro bayesovské riziko $\varrho^*_{k_0,k_1}(q)$ platí

$$\varrho^*_{k_0,k_1}(q) = (k_0+k_1) \int_{R_n} \int_{-\infty}^{\delta^*_{k_0,k_1}(\underset{\sim}{x})} (\delta^*_{k_0,k_1}(\underset{\sim}{x})-\theta) r(\underset{\sim}{x}|\theta) q(\theta) d\nu_n(\underset{\sim}{x}) d\lambda(\theta) +$$

$$+ k_0 \int_{R_n} \int_{-\infty}^{+\infty} (\theta-\delta^*_{k_0,k_1}(\underset{\sim}{x})) r(\underset{\sim}{x}|\theta) q(\theta) d\nu_n(\underset{\sim}{x}) d\lambda(\theta).$$

<u>Důsledek 4.2.</u> Je-li ve větě 4.1 a) $w(\theta) \equiv 1$, pak bayesovský odhad $\delta^*_2(\underset{\sim}{x})$ parametru $\theta$ je

$$\delta^*_2(\underset{\sim}{x}) = E(\theta|\underset{\sim}{X}=\underset{\sim}{x}) \qquad (4.10)$$

a pro bayesovské riziko platí

$$\varrho^*_2(q) = E\{var(\theta|\underset{\sim}{X})\}. \qquad (4.11)$$

Připomeňme, že $100\alpha\,\%$ kvantil $\hat{\theta}_\alpha(\underset{\sim}{x})$ rozdělení $\pi(\theta|\underset{\sim}{x})$ je definován vztahem $\int_{\Theta\cap\{\theta\le\hat{\theta}_\alpha(\underset{\sim}{x})\}} \pi(\theta|\underset{\sim}{x}) d\lambda(\theta) < \alpha$ , $\int_{\Theta\cap\{\theta\ge\hat{\theta}_\alpha(\underset{\sim}{x})\}} \pi(\theta|\underset{\sim}{x}) d\lambda(\theta) \geq 1-\alpha$ .

<u>Poznámka 4.3.</u> Bayesovský odhad $\delta^*_{2,w}(\underset{\sim}{x})$ je podmíněná střední hodnota $\theta$ při daném $\underset{\sim}{X} = \underset{\sim}{x}$ vzhledem k hustotě $\pi_1(\theta|\underset{\sim}{x})$.

<u>Důkaz věty 4.1.</u> a) Pro libovolné $\delta\in\Delta$ a libovolné $\underset{\sim}{x}\in R_n$ platí

$$E_1(\theta-\delta^*_{2,w}(\underset{\sim}{x}))(\delta^*_{2,w}(\underset{\sim}{x})-\delta(\underset{\sim}{x})) = 0$$

a tudíž

$$\varrho(q,\delta) = E_1(\theta-\delta(\underset{\sim}{x}))^2 = \int_{R_n}(\int_{\Theta}(\theta-\delta(\underset{\sim}{x}))^2\pi_1^*(\underset{\sim}{x}|\theta)d\lambda(\theta)r_1(\underset{\sim}{x})d\nu_n(\underset{\sim}{x}) =$$

$$= \int_{R_n}\Big\{\int_{\Theta}(\theta-\delta^*_{2,w}(\underset{\sim}{x}))^2\pi_1(\underset{\sim}{x}|\theta)d\lambda(\theta) +$$

$$+ \int_{\Theta}(\delta(\underset{\sim}{x})-\delta^*_{2,w}(\underset{\sim}{x}))^2\pi_1(\underset{\sim}{x}|\theta)d\lambda(\theta)\Big\} r_1(\underset{\sim}{x})d\nu_n(\underset{\sim}{x}) \geq$$

$$\geq \int_{R_n}\int_{\Theta}(\theta-\delta^*_{2,w}(\underset{\sim}{x}))^2\pi_1(\underset{\sim}{x}|\theta)d\lambda(\theta) = E_1(\mathrm{var}\{\theta|\underset{\sim}{X})).$$

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

b) Nechť $\delta$ je libovolný odhad parametru $\theta$. Je-li $\delta(\underset{\sim}{x}) > \delta^*_{k_0,k_1}(\underset{\sim}{x})$, pak platí

$$L_{k_0,k_1}(\theta,\delta(\underset{\sim}{x}))-L_{k_0,k_1}(\theta,\delta^*_{k_0,k_1}(\underset{\sim}{x})) = k_0(\delta^*_{k_0,k_1}(\underset{\sim}{x})-\delta(\underset{\sim}{x})) \quad \text{je-li } \theta-\delta(\underset{\sim}{x}) \geq 0$$

$$= k_1(-\theta+\delta(\underset{\sim}{x}))-k_0(\theta-\delta^*_{k_0,k_1}(\underset{\sim}{x})) \quad \text{je-li } \theta-\delta(\underset{\sim}{x}) < 0 \leq \theta-\delta^*_{k_0,k_1}(\underset{\sim}{x})$$

$$= k_1(\delta(\underset{\sim}{x})-\delta^*_{k_0,k_1}(\underset{\sim}{x})) \quad \text{je-li } \theta-\delta^*_{k_0,k_1}(\underset{\sim}{x}) < 0,$$

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

$$\int_{-\infty}^{+\infty}(L_{k_0,k_1}(\theta,\delta(\underset{\sim}{x}))-L_{k_0,k_1}(\theta,\delta^*_{k_0,k_1}(\underset{\sim}{x})))r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) = \tag{4.11}$$

$$= k_0(\delta^*_{k_1,k_2}(\underset{\sim}{x})-\delta(\underset{\sim}{x}))(\int_{-\infty}^{+\infty}r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) - \int_{\delta^*_{k_0,k_1}(\underset{\sim}{x})}^{\delta(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) -$$

$$- \int_{-\infty}^{\delta^*_{k_0,k_1}(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta)) +$$

$$+ (k_1\delta(\underset{\sim}{x}) + k_0\delta^*_{k_0,k_1}(\underset{\sim}{x}))\int_{\delta^*_{k_0,k_1}(\underset{\sim}{x})}^{\delta(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) -$$

$$- (k_0+k_1) \int_{\delta^*_{k_0,k_1}(\underset{\sim}{x})}^{\delta(\underset{\sim}{x})} \theta r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) + $$
$$+ k_1(\delta(\underset{\sim}{x})-\delta^*_{k_0,k_1}(\underset{\sim}{x})) \int_{-\infty}^{\delta^*_{k_0,k_1}(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta).$$

Použijeme-li nerovnost

$$\int_{\delta^*_{k_0,k_1}(\underset{\sim}{x})}^{\delta(\underset{\sim}{x})} \theta r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) \le \delta(\underset{\sim}{x}) \int_{\delta^*_{k_0,k_1}(\underset{\sim}{x})}^{\delta(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta)$$

a sečteme-li na pravé straně (4.11) integrály s mezemi $(\delta^*_{k_0,k_1}(\underset{\sim}{x}), \delta(\underset{\sim}{x}))$, dostaneme, že jejich součet je nezáporný. Kromě toho z definice kvantilu plyne, že

$$\int_{-\infty}^{\delta^*_{k_0,k_1}(\underset{\sim}{x})} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta) \geq \int_{-\infty}^{+\infty} r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta).\frac{k_0}{k_1+k_0},$$

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

$$\varrho^*_{k_0,k_1}(q) = \int_{R_n}\int_{-\infty}^{+\infty} L_{k_0,k_1}(\theta,\delta^*_{k_0,k_1}(\underset{\sim}{x}))r(\underset{\sim}{x}|\theta)q(\theta)d\nu_n(\underset{\sim}{x})d\lambda(\theta) =$$
$$= (k_0+k_1)\int_{R_n}\int_{-\infty}^{\delta^*_{k_0,k_1}(\underset{\sim}{x})} (\delta^*_{k_0,k_1}(\underset{\sim}{x})-\theta)r(\underset{\sim}{x}|\theta)q(\theta)d\nu_n(\underset{\sim}{x})d\lambda(\theta)$$
$$+ k_0 \int_{R_n}\int_{-\infty}^{+\infty} (\theta-\delta^*_{k_0,k_1}(\underset{\sim}{x}))r(\underset{\sim}{x}|\theta)q(\theta)d\nu_n(\underset{\sim}{x})d\lambda(\theta).$$

Q.E.D.

Z důsledku 4.2 a poznámky 4.3 plyne, že bayesovský odhad $\delta^*_2$ je podmíněná střední hodnota parametru $\theta$ při aposteriorní hustotě $\pi(\theta|\underset{\sim}{x})$.

Tedy pro aposteriorní hustoty $\pi(\theta|\underline{x})$ uvedené v kapitole 2 snadno spočteme odhady $\delta_2^*$, stačí nalézt v Apendixu příslušnou střední hodnotu. Za zmínku stojí, že pro limitní aposteriorní hustoty uvedené v 2.3 se většina bayesovských odhadů parametru $\theta$ shoduje s odhady běžně používanými v klasické statistice.

Příklad 1.1 (pokračování ze str. 8,45). Vyjádření (3.11) pro bayesovský odhad $\delta_2^*$ 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 podchycení 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

$$N(\frac{225rx+100^2(r-225)}{225r+100(r-225)}, \frac{100.225.r}{100(r-225)+225r}).$$

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

$$\varrho_{2,w}^*(q) = \frac{225.100.r}{100(r-225)+225r} ,$$

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 $x < 100$ $\theta$ spíše podhodnotíme a pro $x > 100$ spíše nadhodnotíme ve srovnání s použitím odhadu $\delta_2^*$.

**Příklad 4.1.** Chceme odhadnout pravděpodobnost $\theta$ výskytu určitého znaku u populace jedinců. Nechť $X_1,\dots,X_n$ je náhodný výběr z této populace, kde

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

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

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

$$L(\theta,\delta(\underset{\sim}{x})) = \frac{(\theta-\delta(\underset{\sim}{x}))^2}{\theta(1-\theta)} ,$$

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

$$\delta^*(\underset{\sim}{x}) = \frac{\int_0^1 \theta^{\alpha+\sum_i x_i-1}(1-\theta)^{\beta+n-\sum_i x_i-2}\,d\theta}{\int_0^1 \theta^{\alpha+\sum_i x_i-2}(1-\theta)^{\beta+n-\sum_i x_i-2}\,d\theta} =$$

$$= \frac{B(\alpha+\sum_i x_i,\ \beta+n-\sum_i x_i-1)}{B(\alpha+\sum_i x_i-1,\ \beta+n-\sum_i x_i-1)} =$$

$$= \frac{\alpha+\sum_i x_i-1}{\alpha+\beta+n-2} ,$$

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

Je-li $0<\beta\le 1$ a $\sum_i x_i = 0$, pak

$$\delta^*(\underset{\sim}{x}) = 0.$$

Je-li $0<\beta\le 1$ a $\sum_i x_i = n$, pak $\delta^*$ není definováno vztahem (4.6), ale můžeme ho dodefinovat pomocí limity:

$$\delta^*(\underset{\sim}{x}) = \lim_{a\to 1_-} \frac{\int_0^a \theta^{\alpha+n-1}(1-\theta)^{\beta-2}\,d\theta}{\int_0^a \theta^{\alpha+n-2}(1-\theta)^{\beta-2}\,d\theta} =$$

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

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

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

$$= \sum_{x=0}^{n} \binom{n}{x} \frac{B(\alpha+x,\beta+n-x)}{B(\alpha-1,\beta-1)(\alpha+\beta+n-2)} = \frac{(\alpha-1)(\beta-1)}{(\alpha+\beta-1)(\alpha+\beta-2)(\alpha+n+\beta-2)},$$

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

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

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

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

<u>Příklad 4.2.</u> 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 $\theta$ má inverzní gama rozdělení s parametry $(\beta,\alpha)$, $\alpha > 0$, $\beta > 0$ dané hustotou

$$q(\theta;\alpha,\beta) = (\Gamma(\alpha))^{-1}\beta^{-\alpha}\,\theta^{-\alpha-1}\,e^{-(\theta\beta)^{-1}} \qquad \theta > 0$$

$$= 0 \qquad \theta \leq 0.$$

Aposteriorní hustota parametru $\theta$ je

$$\pi(\theta|\underset{\sim}{x}) = (\Gamma(\alpha+n))^{-1}(1/\beta+\sum_{i=1}^{n} x_i)^{\alpha+n}\theta^{-\alpha-n-1}\exp\left\{-(\tfrac{1}{\beta}+\sum_{i=1}^{n} x_i)/\theta\right\} \qquad \theta > 0$$

$$= 0 \qquad \theta \leq 0,$$

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

$$\delta^*(x) = \int_{-\infty}^{+\infty} \theta\pi(\theta|\underset{\sim}{x})d\theta = \frac{\frac{1}{\beta} + \sum_i x_i}{\alpha+n-1}$$

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

$$\varrho^*(q) = E\{var(\theta|\underset{\sim}{X})\} = E(\frac{1}{\beta} + \sum_i x_i)^2 (\alpha + n - 1)^{-2} (\alpha + n - 2)^{-1} =$$

$$= \frac{\beta^2}{(\alpha-1)(\alpha-2)(\alpha+n-1)} ,$$

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

$$\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)\beta^{\alpha}(1/\beta+\sum_i x_i)^{\alpha+n}} \qquad \text{pro } x_i \geq 0,\ i = 1,\dots,n$$

a rovna nule jinak.

Příklad 1.2 (pokračování ze str. 8 ). Předpokládejme, že dva fyzikové vyjádří svou představu o sledované fyzikální konstantě $\theta$ následovně. Zkušenější z nich říká, že možné hodnoty $\theta$ mají rozdělení $N(900,400)$. Zatímco druhý z nich (méně zkušený) říká, že možné hodnoty $\theta$ 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

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

a při použití apriorní informace fyzika méně zkušeného je

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

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

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

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

a méně zkušeného fyzika

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

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

Vedle dvou uvedených typů odhadů se používá ještě <u>bayesovský odhad $\hat{\theta}_M$ maximálně věrohodného typu</u>, který je definován následovně:

$$\max_{\theta \in \Theta} r(\underset{\sim}{x} \mid \theta) q(\theta) = r(\underset{\sim}{x} \mid \hat{\theta}_M) q(\hat{\theta}_M),$$

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

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

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

pro $\theta > 0$, výpočtem dostaneme

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

tj.

$$\hat{\theta}_M = \delta^*_{2,1}(\underset{\sim}{x}).$$

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. Nechť $X_1,\ldots,X_n$ jsou nezávislé náhodné veličiny, $X_i$ má rozdělení $N(\theta,\sigma_0^2)$, kde $\sigma_0^2>0$ je známé. Nechť $Y_1,\ldots,Y_N$ jsou nezávislé náhodné veličiny, které představují výsledky z minulosti, $Y_i$ mají rozdělení $N(\theta_q,\sigma_q^2)$, kde $\theta_q$ i $\sigma_q^2>0$ jsou neznámé.

Uvažujme úlohu odhadu parametru $\theta$ při ztrátové funkci $L_2$ dané (4.1).

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

$$\hat{\theta}_q = \bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\begin{aligned}\hat{\sigma}_q^2 &= \hat{\sigma}^2 - \sigma_0^2 && \text{je-li } \hat{\sigma}^2 > \sigma_0^2\\ &= 0 && \text{jinak,}\end{aligned}$$

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

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

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

Je-li navíc $n = 1$, pak

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

Příklad 4.4. Nechť $X$ má Poissonovo rozdělení s parametrem $\lambda$. Nechť $Y_1,\ldots,Y_N$ představují výsledky z minulosti a jsou to nezávislé náhodné veličiny, $Y_i$ má Poissonovo rozdělení s parametrem $\lambda_i$; $\lambda_1,\ldots\lambda_N$ jsou nezávislé náhodné veličiny, $\lambda_i$ má rozdělení gama $(\alpha,\beta)$, $\alpha>0$, $\beta>0$. Pro nepodmíněnou střední hodnotu a nepodmíněný rozptyl náhodné veličiny $Y_i$ platí

$$EY_i = E\lambda_i = \frac{\beta}{\alpha}$$

$$\operatorname{var} Y_i = E\lambda_i + E(\lambda_i - \tfrac{\beta}{\alpha})^2 = \frac{\beta}{\alpha} + \frac{\beta}{\alpha^2} = \frac{\beta}{\alpha}\left(\frac{\alpha+1}{\alpha}\right)$$

$$\hat{\alpha} = \frac{\bar{Y}}{\hat{\sigma}^2 - \bar{Y}} \qquad \hat{\beta} = \frac{\bar{Y}^2}{\hat{\sigma}^2 - \bar{Y}} \qquad \text{je-li } \hat{\sigma}^2 - \bar{Y} > 0$$

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

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

$L_2 = (\lambda - \delta(x))^2$

V případě $\hat{\sigma}^2 - \bar{Y} < 0$ není tento odhad vhodný. Můžeme však postupovat jiným způsobem. Při ztrátové funkci $L_2$ je obécné 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

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

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

## 4.3 BODOVÝ ODHAD; VÍCEROZMĚRNÝ PŘÍPAD

Ve vícerozměrném případě nejčastěji používáme ztrátovou funkci

$$L_A(\underset{\sim}{\theta}, \underset{\sim}{\delta}(\underset{\sim}{x})) = (\underset{\sim}{\theta} - \underset{\sim}{\delta})' \underset{\sim}{A} (\underset{\sim}{\theta} - \underset{\sim}{\delta}),$$

kde $\underset{\sim}{A}$ je symetrická pozitivně semidefinitní matice typu $k \times k$ a $\underset{\sim}{\delta}(\underset{\sim}{x}) = (\delta_1(\underset{\sim}{x}), \ldots, \delta_k(\underset{\sim}{x}))'$ je odhad parametru $\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k)'$. Za předpokladu, že prvky varianční matice $\mathrm{var}\{\underset{\sim}{X}|\underset{\sim}{\theta}\}$ jsou konečné, splňuje bayesovský odhad $\underset{\sim}{\delta}_A^*(\underset{\sim}{x})$ parametru $\theta$ následující vztah

$$(\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-E(\underset{\sim}{\theta}|\underset{\sim}{X}=\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-E(\underset{\sim}{\theta}|\underset{\sim}{X}=\underset{\sim}{x})) = 0 \qquad (4.13)$$

a bayesovské riziko je rovno

$$\varrho_A^*(q) = E\{\mathrm{tr}\{\underset{\sim}{A}\ \mathrm{var}(\underset{\sim}{\theta}|\underset{\sim}{X})\} \cdot \int_{\Theta} r(\underset{\sim}{X}|\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta})\} . \qquad (4.14)$$

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

$$\int_{\Theta}(\underset{\sim}{\theta}-\underset{\sim}{\delta}(\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\theta}-\underset{\sim}{\delta}(\underset{\sim}{x}))r(\underset{\sim}{x}|\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}) =$$

$$= \int_{\Theta}\{(\underset{\sim}{\theta}-\underset{\sim}{\delta}_A^*(\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\theta}-\underset{\sim}{\delta}_A^*(\underset{\sim}{x}))+2(\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-\underset{\sim}{\delta}(\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\theta}-\underset{\sim}{\delta}_A^*(\underset{\sim}{x})) +$$

$$+ (\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-\underset{\sim}{\delta}(\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-\underset{\sim}{\delta}(\underset{\sim}{x}))\}r(\underset{\sim}{x}|\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}) =$$

$$= \mathrm{tr}\{\underset{\sim}{A}\ \mathrm{var}\{\underset{\sim}{\theta}|\underset{\sim}{X}=\underset{\sim}{x}\}\} + \int_{\Theta} r(\underset{\sim}{x}|\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}) \ .$$

$$\cdot\ (\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-\underset{\sim}{\delta}(\underset{\sim}{x}))'\underset{\sim}{A}(\underset{\sim}{\delta}_A^*(\underset{\sim}{x})-\underset{\sim}{\delta}(\underset{\sim}{x})) \geq \mathrm{tr}\{\underset{\sim}{A}\ \mathrm{var}\{\underset{\sim}{\theta}|\underset{\sim}{X}=\underset{\sim}{x}\}\} .$$

Je-li matice $\underset{\sim}{A}$ regulární, pak platí

$$\underset{\sim}{\delta}_{\underset{\sim}{A}}^*(\underset{\sim}{x}) = E(\underset{\sim}{\theta}|\underset{\sim}{X}=\underset{\sim}{x}). \qquad (4.15)$$

Je-li matice $\underset{\sim}{A}$ singulární, pak existuje řešení více.

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

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

$$L(\mu,\sigma^{-2};\delta_1(\underset{\sim}{x}),\delta_2(\underset{\sim}{x})) = (\mu-\delta_1(\underset{\sim}{x}))^2+(\sigma^{-2}-\delta_2(\underset{\sim}{x}))^2,$$

tj. v (4.12) klademe $\underset{\sim}{A} = \underset{\sim}{I}_2$ a podle (4.15) a (2.15) je bayesovský odhad $\mu$ roven

$$\frac{ra + \sum_{i=1}^{n} X_i}{r + n} \tag{4.16}$$

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

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

Tedy klademe-li $r = 0$, $c = 0$, $d = -1/2$, dostáváme běžný odhad $\mu$ a $\sigma^2$ užívaný v klasické statistice. Zobecněný maximálně věrohodný odhad je shodný s (4.16) a zobecněný maximálně věrohodný odhad $\sigma^{-2}$ je

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

<u>Příklad 4.5.</u> Nechť $\underset{\sim}{X} = (X_1,\dots,X_k)'$ má multinomické rozdělení s parametry $(m,\underset{\sim}{\theta})$, kde $m$ je známé a $\underset{\sim}{\theta} = (\theta_1,\dots,\theta_k)'$ má Dirichletovo rozdělení s parametry $\underset{\sim}{\alpha} = (\alpha_1,\dots,\alpha_k)'$. Označme $\alpha_0 = \sum_{i=1}^{k}\alpha_i$. Úkolem je najít bayesovský odhad parametru $\underset{\sim}{\theta}$ a bayesovské riziko $\varrho^*(q)$ při ztrátové funkci

$$L(\underset{\sim}{\theta},\underset{\sim}{\delta}(\underset{\sim}{x})) = \sum_{i=1}^{k}(\theta_i-\delta_i(\underset{\sim}{x}))^2,$$

kde $\underset{\sim}{\delta}(\underset{\sim}{x}) = (\delta_1(\underset{\sim}{x}),\dots,\delta_k(\underset{\sim}{x}))'$. Aposteriorní rozdělení parametru $\underset{\sim}{\theta}$ je Dirichletovo $(\alpha_1+x_1,\dots,\alpha_k+x_k)$. Dále platí

$$\varrho(q,\delta) = \sum_{i=1}^{k}\int_{R_n}\Big(\int_{\Theta}(\theta_i-\delta_i(\underset{\sim}{x}))^2\tau(\underset{\sim}{\theta}|\underset{\sim}{x})d\theta\int_{\Theta}\prod_{i=1}^{n}r(x_i|z)q(z)dz\Big)\,dx_1\dots dx_n \tag{4.17}$$

$$= \sum_{i=1}^{k}E\{E((\theta_i-\delta_i(\underset{\sim}{X}))^2|\underset{\sim}{X}\},$$

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

Vzhledem k (4.16) bude minima dosaženo pro

$$\delta^*(\underset{\sim}{x}) = (\delta_1^*(\underset{\sim}{x}),\dots,\delta_k^*(\underset{\sim}{x}))'$$

$$\delta_i^*(\underset{\sim}{x}) = E(\theta_i|\underset{\sim}{x}) = \frac{\alpha_i + x_i}{\alpha_0 + m},$$

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

$$E((\theta_i-\delta_i^*(X))^2|\underset{\sim}{X}=\underset{\sim}{x}) = E((\theta_i-E(\theta_i|\underset{\sim}{X}=\underset{\sim}{x}))^2|\underset{\sim}{X}=\underset{\sim}{x}) = \tag{4.18}$$

$$= \frac{(\alpha_i+x_i)(\alpha_0+m-\alpha_i-x_i)}{(\alpha_0+m)^2(\alpha_0+m+1)}.$$

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

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

$$= \frac{1}{\binom{m+\alpha_0}{m}} \prod_{i=1}^{k} \binom{\alpha_i+x_i-1}{\alpha_i-1}$$

a tudíž marginální rozdělení $X_i$ je

$$\frac{1}{\binom{m+\alpha_0}{m}} \binom{\alpha_i+x_i-1}{\alpha_1-1} \binom{m+\alpha_0-\alpha_i-x_i-1}{\alpha_0-\alpha_1-1} \quad \text{pro } x_i = 0,\dots,n,\quad \sum_{i=1}^{k} x_i = m$$

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

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

## 4.4 VĚROHODNOSTNÍ MNOŽINY

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

$$P(\underset{\sim}{\theta} \in D_\alpha(\underset{\sim}{X}) \mid \underset{\sim}{\theta}) = 1 - \alpha \quad \text{pro vš. } \underset{\sim}{\theta} \in \Theta \tag{4.19}$$

Při bayesovském přístupu konfidenčním množinám odpovídají tzv. <u>100(1-$\alpha$)% věrohodnostní množiny</u> (credible region v angličtině) <u>parametru $\underset{\sim}{\theta}$</u>. 100(1-$\alpha$)% věrohodnostní množina parametru $\underset{\sim}{\theta}$ je definovaná jako libovolná množina $C_\alpha(\underset{\sim}{X}) \subset \Theta$ taková, že

$$P(\underset{\sim}{\theta} \in C_\alpha(\underset{\sim}{X}) \mid \underset{\sim}{X}) = \int_{C_\alpha(\underset{\sim}{X})} \pi(\underset{\sim}{\theta} \mid \underset{\sim}{X}) \, d\lambda(\underset{\sim}{\theta}) = 1 - \alpha \ ; \tag{4.20}$$

$1-\alpha$ nazýváme <u>věrohodnost</u>. Někdy nahrazujeme poslední rovnost nerovností $\geq$. Protože $\pi(\underset{\sim}{\theta} \mid \underset{\sim}{x})$ je hustota na $\Theta$, můžeme mluvit o pravděpodobnosti, že $\underset{\sim}{\theta}$ náleží do $C_\alpha(\underset{\sim}{X})$. Na rozdíl od klasického přístupu, kdy konfidenční množinu $D_\alpha(\underset{\sim}{X})$ lze interpretovat jen v termínech pravděpodobnosti pokrytí. Jak uvidíme na příkladech v řadě případů jsou množiny $C_\alpha(\underset{\sim}{X})$ a $D_\alpha(\underset{\sim}{X})$ totožné.

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

$$C^*_\alpha(\underset{\sim}{X}) = \{\underset{\sim}{\theta} \in \Theta;\ \pi(\underset{\sim}{\theta}|\underset{\sim}{X}) \geq k_\alpha\}, \tag{4.21}$$

kde $k_\alpha$ je největší konstanta taková, že

$$\int_{C^*_\alpha(\underset{\sim}{X})} \pi(\underset{\sim}{\theta}|\underset{\sim}{X})\,d\lambda(\underset{\sim}{\theta}) = 1 - \alpha \quad (\text{popř.} \geq 1 - \alpha). \tag{4.22}$$

Najít $C^*_\alpha(\underset{\sim}{X})$ je mnohdy značně obtížné. Poznamenejme, že $C^*_\alpha(\underset{\sim}{X})$ je věrohodnostní množina splňující (4.21) a navíc

$$\int_{C^*_\alpha(\underset{\sim}{X})} d\lambda(\underset{\sim}{\theta}) \leq \int_{C_\alpha(\underset{\sim}{X})} d\lambda(\underset{\sim}{\theta}) \text{ pro vš. } C_\alpha(\underset{\sim}{X}). \tag{4.23}$$

Je-li $\theta$ jednorozměrný parametr, hledáme obvykle věrohodnou množinu ve tvaru intervalu a mluvíme o <u>$100(1-\alpha)\%$ věrohodnostním intervalu</u>.

Nechť $X_1,\dots,X_n$ je náhodný výběr z $N(\mu,\sigma^2)$, kde $\mu$ 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 $\mu^*$, $c^*$, $d^*$ jsou dány, (2.15), (2.16) resp. (2.17), je t-rozdělení o $2d^*$ stupních volnosti. Z vlastností t-rozdělení plyne, že interval

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

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

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

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

ale není symetrické kolem tohoto bodu. Tudíž zkonstruovat věrohodnostní interval je numericky velmi obtížné a tak většinou bereme následující interval za $100(1-\alpha)\%$ věrohodnostní pro $\sigma^{-2}$.

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

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

Není-li o rozdělení $(\mu,\sigma^{-2})$ 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 $\mu$ a $\sigma^{-2}$, 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 $\mu$ 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 $\gamma^2$ má tvar

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

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

str. 8, 45

Příklad 1.1 (pokračování). Chceme-li najít 95% věrohodný interval pro kvocient inteligence $\Theta$, 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

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

Klasický 95% interval spolehlivosti má tvar

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

Tedy bayesovský interval je kratší.

Příklad 4.4. Při kontrole jakosti výrobků bylo náhodně vybráno a zkontrolováno n výrobků. Předpokládáme, že podíl vadných $\theta$ v základním souboru má beta rozdělení $(a,b)$, $a > 0$, $b > 0$. Chceme nalézt $100(1-\alpha)\%$ věrohodnostní interval. Aposteriorní rozdělení $\theta$ 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 $\sigma^{-2}$. 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 $(1-(1-\alpha/2)^{(b+n)^{-1}},\ 1-(\alpha/2)^{(b+n)^{-1}})$. 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 $\theta$. O hodnotě $\theta$ nejsou dostupné žádné informace a tak za apriorní hustotu zvolíme

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

$$= 0 \qquad \theta \leq 0.$$

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 zkonstruovat $100(1-\alpha)\%$ interval věrohodnosti pro parametr $\theta$. Za $100(1-\alpha)\%$ věrohodnostní interval obvykle bereme

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

kde $H^{-1}(x;c,d)$ je kvantilová funkce gama rozdělení s parametry $(c,d)$. Je-li $\sum_i x_i = 1$, 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. TESTOVÁNÍ HYPOTÉZ

## 5.1 ÚVOD

Nechť náhodný vektor $\underset{\sim}{X} = (X_1, \ldots, X_n)'$ má hustotu $r(\underset{\sim}{x}|\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\nu_n$, kde $\underset{\sim}{\theta}$ je parametr náležející do neprázdné borelovské množiny $\Theta \subset R_k$. Parametr $\underset{\sim}{\theta}$ má apriorní hustotu $q(\underset{\sim}{\theta})$ vzhledem k $\sigma$-konečné míře $\lambda$.

Úlohu testovat hypotézu $H_0$: $\underset{\sim}{\theta} \in \Theta_0$ proti $H_1$: $\underset{\sim}{\theta} \in \Theta - \Theta_0 = \Theta_1$, $\Theta_i$ jsou neprázdné borelovské podmnožiny $\Theta$, lze formulovat jako statistický rozhodovací problém $(\Theta, \Delta, R)$, kde $\Delta$ 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 $H_i$: $\underset{\sim}{\theta} \in \Theta_i$, $i = 1, \ldots, q$, $\Theta_1, \ldots, \Theta_q$ jsou neprázdné disjunktní borelovské podmnožiny $\Theta$, $\bigcup_{i=1}^{q} \Theta_i = \Theta$, můžeme formulovat jako statistický rozhodovací problém $(\Theta, \Delta, R)$, kde $\Delta$ je množina rozhodovacích funkcí, které nabývají pouze hodnot $d_1, \ldots, d_k$, kde $d_i$ označuje rozhodnutí, že platí hypotéza $H_i$, $i = 1, \ldots, k$.

## 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 obdobné). Typ velice často souvisí s tvarem hypotéz.

Obecně volíme ztrátovou funkci L s vlastnostmi

$$L(\underset{\sim}{\theta}, d_i) = 0 \qquad \theta \in \Theta_i, \ i = 0,1$$

$$L(\underset{\sim}{\theta}, d_i) \geqslant 0 \qquad \theta \notin \Theta_i \quad .$$

Jsou-li obě hypotézy jednoduché, volíme ztrátovou funkci

$$L(\underset{\sim}{\theta}_i, d_i) = 0 \qquad i = 0,1$$

$$L(\underset{\sim}{\theta}_0, d_1) = a_1$$

$$L(\underset{\sim}{\theta}_1, d_0) = a_0$$

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

$$a_1 \xi\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_0) > a_0(1-\xi)\, r(\underset{\sim}{x}|\underset{\sim}{\theta}_1), \tag{5.1}$$

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:

$$\begin{aligned} L_*(\underset{\sim}{\theta}, \underset{\sim}{d}_i) &= 0 \qquad \theta \in \Theta_i \\ &= a_i \qquad \theta \notin \Theta_i \qquad i = 0,1, \end{aligned} \tag{5.2}$$

$$\begin{aligned} L_{**}(\underset{\sim}{\theta}, \underset{\sim}{d}_i) &= 0 \qquad \theta \in \Theta_i \\ &= K_i\, d(\underset{\sim}{\theta}, \Theta_i) \qquad \theta \notin \Theta_i \qquad i = 0,1 \end{aligned} \tag{5.3}$$

kde $d(\underset{\sim}{\theta}, \Theta_i)$ je vzdálenost (obvykle Eukleidova) $\underset{\sim}{\theta}$ od množiny $\Theta_i$, $a_i > 0$, $K_i > 0$, $i = 0,1$. Zatímco ztrátová funkce $L_*$ závisí pouze na tom, zda $\underset{\sim}{\theta}$ náleží do $\Theta_0$ nebo $\Theta_1$, ztrátová funkce $L_{**}$ nabývá tím větší hodnoty, čím je skutečná hodnota $\underset{\sim}{\theta}$ vzdálenější od hypotézy, pro kterou jsme se rozhodli.

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

$$\begin{aligned} R_*(\underset{\sim}{\theta}, \delta) &= a_0\, P(\delta(\underset{\sim}{X}) = d_0 / \underset{\sim}{\theta}) \qquad \underset{\sim}{\theta} \in \Theta_1 \\ &= a_1\, P(\delta(\underset{\sim}{X}) = d_1 / \underset{\sim}{\theta}) \qquad \Theta \in \Theta_0 \end{aligned}$$

a odtud dále obdržíme pro bayesovskou rizikovou funkci

$$\varrho_*(q,\delta) = a_0 \int_{\Theta_1} P(\delta(\underset{\sim}{X})=d_0/\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}) + \qquad (5.4)$$

$$+ a_1 \int_{\Theta_0} P(\delta(\underset{\sim}{X})=d_1/\underset{\sim}{\theta})q(\underset{\sim}{\theta})d\lambda(\underset{\sim}{\theta}).$$

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

Pro libovolnou rozhodovací funkci $\delta$ platí

$$P(\delta(\underset{\sim}{X})=d_0/\underset{\sim}{\theta}) + P(\delta(\underset{\sim}{X})=d_1/\underset{\sim}{\theta}) = 1 \text{ pro vš. } \underset{\sim}{\theta}\in\Theta \qquad (5.5)$$

$$\int_{\Theta_1} L_*(\underset{\sim}{\theta},d_0)r(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta}) = a_0P(\underset{\sim}{\theta}\in\Theta_1/\underset{\sim}{X}=\underset{\sim}{x}) \qquad (5.7)$$

$$\int_{\Theta_0} L_*(\underset{\sim}{\theta},d_1)r(\underset{\sim}{\theta}|\underset{\sim}{x})d\lambda(\underset{\sim}{\theta}) = a_1P(\underset{\sim}{\theta}\in\Theta_0/\underset{\sim}{X}=\underset{\sim}{x}) \qquad (5.8)$$

Vzhledem k (3.10) se rozhodneme pro $d_0$, jestliže

$$a_0P(\underset{\sim}{\theta}\in\Theta_1/\underset{\sim}{X}=\underset{\sim}{x}) < a_1P(\underset{\sim}{\theta}\in\Theta_0/\underset{\sim}{X}=\underset{\sim}{x}). \qquad (5.9)$$

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

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

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

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

Čísly $a_0,a_1$ 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 určitou 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 tratích plyne, že $\theta$ můžeme považovat za náhodnou veličinu s Paretovým rozdělením $(a, x_0)$.

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

$$P(\theta \leq d | \underset{\sim}{X} = \underset{\sim}{x}) = \int_0^d (1+a)(\max_{0 \leq i \leq n} x_i)^{a+1} \theta^{-a-2} I\{\theta > \max_{0 \leq i \leq n} x_i\} d\theta = \qquad (5.11)$$

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

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

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

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

kde $\underset{\sim}{X} = (X_1, \dots, X_n)'$.

Při ztrátové funkci

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

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

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

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

kde $d_i$ označuje rozhodnutí, že platí hypotéza $H_i$, $i = 0,1$; $K > 0$, je bayesovská rozhodovací funkce $\delta^*$ následující (viz (5.9)):

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

$$= d_0 \qquad \text{pro } P(\theta \leq d | \underset{\sim}{X} = \underset{\sim}{x}) > P(\theta > d | \underset{\sim}{X} = \underset{\sim}{x})$$

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

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

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

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

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

Rozhodneme se tedy pro $d_1$.

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

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

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

pak bayesovská rozhodovací funkce je rovna $d_0$, jestliže

$$E(\theta | \underset{\sim}{X} = \underset{\sim}{x}) \le \theta_0 \tag{5.14}$$

a je $d_1$, jestliže platí nerovnost $>$.

Přesvědčíme se, že tato rozhodovací funkce je bayesovská. Podle (3.10) stačí spočítat pro libovolnou rozhodovací funkci $\delta$ podmíněnou střední hodnotu ztrátové funkce:

$$\begin{aligned} E(L_{**}(\theta, \delta(\underset{\sim}{X})) | \underset{\sim}{X} = \underset{\sim}{x}) &= \int_{\theta_0}^{+\infty} (\theta - \theta_0) \pi(\theta | \underset{\sim}{x}) \, d\lambda(\theta) \qquad \text{pro } \delta(\underset{\sim}{x}) = d_0 \\ &= \int_{-\infty}^{\theta_0} (\theta_0 - \theta) \pi(\theta | \underset{\sim}{x}) \, d\lambda(\theta) \qquad \text{pro } \delta(\underset{\sim}{x}) = d_1 . \end{aligned}$$

Podle (3.10) se rozhodneme pro $d_0$, jestliže

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

což je ekvivalentní

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

Pro bayesovské riziko postupně dostaneme:

$$\varrho^*(q) = \int_{R_n}\int_{-\infty}^{+\infty} L_{**}(\theta,\delta(\underset{\sim}{x}))r(\underset{\sim}{x}|\theta)q(\underset{\sim}{\theta})d\lambda(\theta) =$$

$$= \int_{\{x;E(\theta|\underset{\sim}{X}=\underset{\sim}{x})\leq\theta_0\}} \left(\int_{\theta_0}^{+\infty}(\theta-\theta_0)r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta)\right)d\nu_n(\underset{\sim}{x})$$

$$- \int_{\{x;E(\theta|\underset{\sim}{X}=\underset{\sim}{x})>\theta_0\}} \left(\int_{-\infty}^{\theta_0}(\theta-\theta_0)r(\underset{\sim}{x}|\theta)q(\theta)d\lambda(\theta)\right)d\nu_n(\underset{\sim}{x}) =$$

$$= \int_{\{x;E(\theta|\underset{\sim}{X}=\underset{\sim}{x})\leq\theta_0\}} (E(\theta|\underset{\sim}{X}=\underset{\sim}{x})-\theta_0)r(\underset{\sim}{x})d\nu_n(\underset{\sim}{x}) - \int_{-\infty}^{\theta_0} (\theta-\theta_0)q(\theta)d\lambda(\theta).$$

<u>Příklad 5.1</u> (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

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

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 $\lambda$) $\theta \in \Theta_i$ kde $i = 0$ nebo 1, pak bychom se nikdy nemohli rozhodnout pro hypotézu $H_i$, neboť nulovost apriorní hustoty implikuje nulovost aposteriorní hustoty. Takováto situace nastane např. je-li $\Theta_1 \in \mathcal{B}_r$, $\Theta \in \mathcal{B}_k$, $r < k$, $\lambda$ Lebesguova míra a $\lambda(\Theta) > 0$ (u všech apriorních hustot uvažovaných ve 2. kapitole byla $\lambda$ 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 $H_0^*$: $\theta \in \Theta_0^*$, $H_1^*$: $\theta \in \Theta - \Theta_0^* = \Theta_1^*$, 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. $\Theta_0$ jednobodová, bereme za $\Theta_0^*$ vhodné okolí tohoto bodu.

str. 8, 45
Příklad 1.1 (pokračování). 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 $\theta_0$ přiřadíme apriorní pravdě podobnost $q > 0$ 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 $\theta$ je

$$P(\theta = \theta_0 | X = x) = \frac{q\, r(x|\theta_0)}{q\, r(x|\theta_0) + (1-q) \int_\Theta r(x|\theta) q(\theta) d\lambda(\theta)} \tag{5.15}$$

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

Marginální hustota X je rovna

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

Na $\Theta - \{\theta_0\}$ existují apriorní i aposteriorní hustoty $q^*$ a $\pi^*$ vzhledem k $\lambda$;

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

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

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 testujme 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 $\theta$ 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

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

Dále z (5.15) plyne, že aposteriorní rozdělení je

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

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

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

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

Dosadíme-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

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

přičemž výraz na pravé straně je větší než 5. Docházíme tedy k paradoxnímu závěru. Pro hypotézu $H_0$ se totiž rozhodneme i v případě

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

zatímco při klasickém přístupu bychom $H_0$ zamítli i na hladině $5{,}1.10^{-7}$.

Z toho lze soudit, že buď ztrátová funkce nebo apriorní rozdělení nebyly vhodně zvoleny (např. hodnoty blízké $\theta_0$ 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 aposteriorních hustot. Pro test hypotézy $H_0$: $\underset{\sim}{\theta} \in \Theta_0$ proti hypotéze $H_1$: $\theta \notin \Theta_0$ sestavíme podíl

$$\frac{\sup\limits_{\underset{\sim}{\theta} \in \Theta_0} \pi(\underset{\sim}{\theta} | \underset{\sim}{x})}{\sup\limits_{\underset{\sim}{\theta} \in \Theta} \pi(\underset{\sim}{\theta} | \underset{\sim}{x})} \tag{5.22}$$

Tento podíl je vždy $\leq 1$. Hodnoty podílu blízké jedné indikují platnost hypotézy $H_0$, zatímco malé hodnoty indikují platnost hypotézy $H_1$.

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

Je analogií testu podílem věrohodností v klasické statistice a používá se též jako v klasické statistice pro test lineárních hypotéz v lineárním modelu. Dosazením (2.30) do (5.22) obdržíme po delším 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 $X_1, \ldots, X_n$ je náhodný výběr z normálního rozdělení $N(\mu, \sigma_0^2)$, $\sigma_0^2 > 0$ známé a $\mu$ je neznámý parametr.

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

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

$a \in R_1$, $b^2 > 0$. Pak platí

$$\frac{P(\mu=\mu_0|\underset{\sim}{X}=\underset{\sim}{x})}{P(\mu \neq \mu_0|\underset{\sim}{X}=\underset{\sim}{x})} = \frac{q(2\pi\sigma_0^2)^{-n/2} \exp\{-\sum_i (x_i-\mu_0)^2(2\sigma_0^2)^{-1}\}}{(1-q).r(\underset{\sim}{x})} \tag{5.23}$$

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

$$\begin{aligned} L_*(\mu,d_0) &= 0 & \mu &= \mu_0 \\ &= a_0 & \mu &\neq \mu_0 \\ L_*(\mu,d_1) &= 0 & \mu &\neq \mu_0 \\ &= a_1 & \mu &= \mu_0, \end{aligned} \tag{5.24}$$

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

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

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

Při ztrátové funkci

$$\begin{aligned} L_{**}(\mu,d_0) &= k_1(\mu-\mu_0)^2 & \mu &\in R_1 \\ L_{**}(\mu,d_1) &= k_2 & \mu &= \mu_0 \\ &= 0 & \mu &\neq \mu_0, \end{aligned} \tag{5.25}$$

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

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

tj.

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

kde $\mu_1$ a $\sigma_1^2$ jsou dány (2.9).

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

$$
\begin{aligned}
L_0(\mu, d_0) &= 0 & \mu \le \mu_0 \\
&= \mu - \mu_0 & \mu > \mu_0 \\
L_0(\mu, d_1) &= 0 & \mu > \mu_0 \\
&= \mu_0 - \mu & \mu \le \mu_0
\end{aligned}
\tag{5.26}
$$

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

$$\frac{\sum_i X_i b^2 + a\,\sigma_0^2}{nb^2 + \sigma_0^2} \le \mu_0 .$$

B. Nechť $X_1, \ldots, X_n$ je náhodný výběr z $N(\mu, \sigma^2)$, kde $\mu$ 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 $(\mu, \sigma^{-2})$ je následující: $P(\mu = \mu_0) = q$, podmíněná apriorní hustota $\sigma^{-2}$ při daném $\mu = \mu_0$ je gama rozdělení $(c,d)$, podmíněná apriorní hustota $(\mu, \sigma^{-2})$ při podmínce $\mu \neq \mu_0$ je normální-gama s parametry $(a,r,c,d)$ (hustota je dána (2.18)). Pak pro podmíněnou hustotu $r(\underset{\sim}{x} | \mu = \mu_0)$ náhodného vektoru $\underset{\sim}{X}$ při podmínce $\mu = \mu_0$ platí

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

a tedy podle (5.15)

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

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

$$\left(\frac{a_1 q}{a_0(1-q)}\sqrt{\frac{r+n}{r}}\right)^{\frac{2}{n+2d}} < \frac{c+\frac{1}{2}\sum_i (x_i-\mu_0)^2}{c+\frac{1}{2}\sum_i (x_i-\bar{x})^2+\frac{rn}{2(r+n)}(\bar{x}-a)^2} \cdot$$

Nyní uvažujme ztrátovou funkci

$$\begin{aligned} L_{**}(\mu,\sigma^{-2};d_0) &= k_0\sigma^{-2}(\mu-\mu_0)^2 & \mu \in R_1,\ \sigma^{-2}>0 \\ L_{**}(\mu,\sigma^{-2};d_1) &= k_1 & \mu = \mu_0,\ \sigma^{-2}>0 \\ &= 0 & \mu \neq \mu_0,\ \sigma^{-2}>0, \end{aligned}$$

kde $k_0 > 0$, $k_1 > 0$. Pro libovolnou rozhodovací funkci $\delta(\underset{\sim}{x})$ platí

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

Využitím (2.15), (2.16) obdržíme

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

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

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

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 $X_1,\ldots,X_n$ z $N(\mu_1,\sigma^2)$ a $Y_1,\ldots$ $\ldots,Y_{n_2}$ z $N(\mu_2,\sigma^2)$, $\mu_1$, $\mu_2$, $\sigma^2 > 0$ neznámé a úlohu testovat hypotézu $H_0$: $\mu_1 \leq \mu_2$ proti $H_1$: $\mu_1 > \mu_2$. Předpokládejme, že apriorní rozdělení $(\mu_1, \mu_2, \sigma^{-2})$ lze popsat následovně. Podmíněné apriorní rozdělení $(\mu_1, \mu_2)$ při daném $\sigma^{-2}$ je $N((a_1, a_2), \sigma^2 \begin{pmatrix} r_1^{-1} & 0 \\ 0 & r_2^{-1} \end{pmatrix})$, marginální apriorní rozdělení $\sigma^{-2}$ je gama s parametry $(c,d)$. Pak aposteriorní rozdělení $\pi(\mu_1, \mu_2, \sigma^{-2} | \underset{\sim}{x}, \underset{\sim}{y})$ parametrů $(\mu_1, \mu_2, \sigma^{-2})$ je následující. Podmíněné aposteriorní rozdělení $(\mu_1, \mu_2)$ při daném $\sigma^{-2}$ je $N(a_1^*, a_2^*)'$, $\sigma^2 \begin{pmatrix} (r_1+n_1)^{-1} & 0 \\ 0 & (r_2+n_2)^{-1} \end{pmatrix})$, kde

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

Marginální aposteriorní rozdělení $\sigma^{-2}$ 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

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

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

$$E(L(\mu_1, \mu_2, \sigma^{-2}; \delta(\underset{\sim}{X}, \underset{\sim}{Y})) | \underset{\sim}{X} = \underset{\sim}{x}, \underset{\sim}{Y} = \underset{\sim}{y}) =$$

$$= \int\int_{\mu_1 > \mu_2} \int_0^{+\infty} (\mu_1 - \mu_2)\ (\mu_1, \mu_2, \sigma^{-2} | \underset{\sim}{x}, \underset{\sim}{y}) d\mu_1 d\mu_2\, d\sigma^{-2} \text{ pro } \delta(\underset{\sim}{x}, \underset{\sim}{y}) = d_0,$$

$$= \int\int_{\mu_1 \leq \mu_2} \int_0^{+\infty} (\mu_2 - \mu_1)\ (\mu_1, \mu_2, \sigma^{-2}) | \underset{\sim}{x}, \underset{\sim}{y})\ d\mu_1 d\mu_2 d\sigma^{-2} \text{ pro } \delta(\underset{\sim}{x}, \underset{\sim}{y}) = d_1.$$

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

$$E(\mu_1|\underset{\sim}{X}=\underset{\sim}{x}) \leq E(\mu_2|\underset{\sim}{Y}=\underset{\sim}{y})$$

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 nerovnost $>$.

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 $H_i$ platí):

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

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

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

Tedy ztráta závisí na vzdálenosti od hypotetické množiny, jestliže jde o $\theta$ ze "sousední" množiny a na jejím dvojnásobku, jestliže $\theta$ nenáleží ani do hypotetické ani do "sousední" množiny. Přímým výpočtem dostaneme pro lib. rozhodovací funkci

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

$$E(L(\theta,\delta(X))|\underset{\sim}{X}=\underset{\sim}{x}) = \int_{-\infty}^{90} (90-\theta)\pi(\theta|x)d\theta + \int_{110}^{+\infty} (\theta-110)\pi(\theta|x)d\theta \quad \text{jestliže } \delta(x) = d_2,$$

$$E(L(\theta,\delta(X))|\underset{\sim}{X}=\underset{\sim}{x}) = \int_{-\infty}^{90} 2(110-\theta)\pi(\theta|x)d\theta + \int_{90}^{110} (110-\theta)\pi(\theta|x)d\theta \quad \text{jestliže } \delta(x) = d_3.$$

Specielně při $x = 115$ platí

$$E(L(\theta,\delta(X))|X=115) = 34{,}32 \quad \text{při } \delta(115) = d_1$$
$$= 3{,}55 \quad = d_2$$
$$= 3{,}24 \quad = d_3.$$

Tedy bayesovské rozhodnutí je $d_3$.

Dále platí, že

$$P(\theta \leq 90|x=115) = 0{,}005$$
$$P(90 < \theta < 110|x=115) = 0{,}475$$
$$P(110 \geq \theta|x=115) = 0{,}520.$$

Odtud plyne, že při ztrátové funkci

$$L^*(\theta,d_i) = 0 \quad \theta \in \Theta_i$$
$$= 1 \quad \theta \notin \Theta_i \quad i = 1,2,3$$

přičemž $\Theta_1 = (-\infty,90\rangle$, $\Theta_2 = (90,110)$, $\Theta_3 = \langle 110,+\infty)$, je bayesovské rozhodnutí opět $d_3$.

## APENDIX; PŘEHLED POUŽITÝCH ROZDĚLENÍ

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

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

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

Platí

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

při $m = 1$ mluvíme o alternativním rozdělení.

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

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

Platí

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

Nega;ivně binomické rozdělení s parametry (a,p) ($a > 0,\ p \in \langle 0,1\rangle$) má hustotu

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

Platí

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

Při $a = 1$ mluvíme o geometrickém rozdělení.

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

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

Platí

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

Gama rozdělení s parametry (a,p) $(a > 0), \ p > 0)$ má hustotu

$$\begin{aligned} r(x|a,p) &= a^p(\Gamma(p))^{-1}\, x^{p-1}\, e^{-ax} && x \geq 0, \\ &= 0 && x < 0, \end{aligned} \tag{A.9}$$

kde $\Gamma(p) = \int\limits_0^{+\infty} t^{p-1}\, e^{-t}\, dt.$

Platí

$$EX = p\, a^{-1}, \qquad \mathrm{var}\, X = p\, a^{-2}. \tag{A.10}$$

Při $p = 1$ mluvíme o exponenciálním rozdělení. Při $p = n/2$ a $a = 1/2$ mluvíme o $\chi^2$-rozdělení (centrálním) o n stupních volnosti.

Beta rozdělení s parametry (a,b) $(a > 0, \ b > 0)$ má hustotu

$$\begin{aligned} r(x|a,b) &= (B(a,b))^{-1}\, x^{a-1}\, (1-x)^{b-1} && x \in (0,1), \\ &= 0 && x \notin (0,1), \end{aligned} \tag{A.11}$$

kde $B(a,b) = \int\limits_0^1 x^{a-1}\, (1-x)^{b-1}\, dx.$

Platí

$$EX = a(a+b)^{-1}, \qquad \mathrm{var}\, X = ab\, \{(a+b)^2(a+b+1)\}^{-1}. \tag{A.12}$$

Rovnoměrné rozdělení s parametry (a,b) $(a < b)$ má hustotu

$$\begin{aligned} r(x|a,b) &= (b-a)^{-1} && x \in (a,b) \\ &= 0 && x \notin (a,b) \end{aligned} \tag{A.13}$$

Platí

$$EX = (a+b)/2, \qquad \mathrm{var}\, X = (b-a)^2/12. \tag{A.14}$$

Paretovo rozdělení s parametry (a,b) ($a>0$, $b>0$) má hustotu

$$r(x|a,b) = (a/b)\,(b/x)^{a+1} \qquad x \geq b \qquad \text{(A.15)}$$
$$= 0 \qquad x < b.$$

Pro $a > 2$ platí

$$EX = ab(a-1)^{-1}, \qquad \operatorname{var} X = ab^2\{(a-1)^2(a-2)\} \qquad \text{(A.16)}$$

Studentovo (t-) rozdělení o n stupních volnosti s parametrem $\mu$ ($\mu \in R_1$) má hustotu

$$r(x|n,\mu) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\ \Gamma(\frac{n}{2})} \left(1 + \frac{(x-\mu)^2}{n}\right)^{-\frac{n+1}{2}} \qquad x \in R_1. \qquad \text{(A.17)}$$

Má-li náhodná veličina X rozdělení $N(\mu,1)$, Y rozdělení $\chi^2$ 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 $\mu$.
Platí

$$EX = \mu \qquad \text{(A.18)}$$

a pro $n > 2$ platí

$$\operatorname{var} X = \frac{n}{n-2}.$$

F-rozdělení s $n_1$ a $n_2$ stupni volnosti má hustotu

$$r(x|n_1,n_2) = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2})\ \Gamma(\frac{n_2}{2})} \left(\frac{n_1}{n_2}\right)^{n_1/2} \left(1 + \frac{n_1}{n_2}x\right)^{-\frac{n_1+n_2}{2}} x^{n_1/2-1}$$
$$x \in R_1. \qquad \text{(A.19)}$$

Jsou-li náhodné veličiny X a Y nezávislé a mají-li $\chi^2$-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

$$EF = \frac{n_2}{n_2-2} \tag{A.20}$$

a pro $n_2 > 4$

$$\operatorname{var} F = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-4)(n_2-2)^2} . \tag{A.21}$$

Multinomické rozdělení s parametry $(n, p)$ ($n = 1,2,\ldots$; $\underset{\sim}{p} = (p_1,\ldots,p_k)'$, $p_i \in \langle 0,1\rangle$, $i = 1,\ldots,k$; $\sum_{i=1}^{k} p_i = 1$, $k \geq 2$) má hustotu

$$r(\underset{\sim}{x}|n,p) = \frac{n!}{x_1!\ldots x_k!} p_1^{x_1}\ldots p_k^{x_k} \quad \underset{\sim}{x} = (x_1,\ldots,x_k)', \tag{A.22}$$

$$x_i = 0,\ldots,n,\ i = 1,\ldots,k,\ \sum_{i=1}^{k} x_i = n$$

$$= 0 \qquad \text{jinak.}$$

Platí

$$EX_i = np_i, \qquad \operatorname{var} X_i = np_i(1-p_i) \qquad i = 1,\ldots,k, \tag{A.23}$$

$$\operatorname{cov}(X_i,X_j) = -np_ip_j \qquad i,j = 1,\ldots,k;\ i \neq j, \tag{A.24}$$

$$J(p_1,\ldots,p_{k-1}) = \frac{n}{p_k} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix} + n \begin{pmatrix} 1/p_1 & & 0 \\ & \ddots & \\ 0 & & 1/p_{k-1} \end{pmatrix}. \tag{A.25}$$

Dirichletovo rozdělení (mnohorozměrné beta rozdělení) s parametry $\underset{\sim}{\alpha} = (\alpha_1,\ldots,\alpha_k)'$ ($\alpha_i > 0$, $i = 1,\ldots,k$) má hustotu

$$r(\underset{\sim}{x}|\underset{\sim}{\alpha}) = \frac{\Gamma(\alpha_1+\ldots+\alpha_k)}{\Gamma(\alpha_1)\ldots\Gamma(\alpha_k)} x_1^{\alpha_1-1}\ldots x_k^{\alpha_k-1} \quad \underset{\sim}{x} = (x_1,\ldots,x_k)' \tag{A.26}$$

$$x_i > 0,\ i = 1,\ldots,k,\ \sum_{i=1}^{k} x_i = 1$$

$$= 0 \qquad \text{jinak.}$$

Platí

$$EX_i = \frac{\alpha_i}{\alpha_0}, \qquad \operatorname{var} X_i = \frac{\alpha_i(\alpha_0-\alpha_i)}{\alpha_0^2(\alpha_0+1)}, \quad i = 1,\ldots,k, \tag{A.27}$$

$$\operatorname{cov}(X_i,X_j) = -\frac{\alpha_i\alpha_j}{\alpha_0^2(\alpha_0+1)}\,, \quad j,i = 1,\dots,k;\ i \neq j, \tag{A.28}$$

kde $\alpha_0 = \sum_{i=1}^{k} \alpha_i$.

Marginální rozdělení $X_i$ je beta rozdělení s parametry $\alpha_i$, $\alpha_0 - \alpha_i$.

k-rozměrné normální rozdělení s parametry $\underset{\sim}{\mu}$ a $\underset{\sim}{\Sigma}$ (ozn. $N_k(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$) ($\underset{\sim}{\mu} \in R_k$, $\underset{\sim}{\Sigma}$ -symetrická pozitivně definitní matice typu $k \times k$) má hustotu

$$r(\underset{\sim}{x}|\underset{\sim}{\mu},\underset{\sim}{\Sigma}) = (2\pi)^{-n/2}(\det \underset{\sim}{\Sigma})^{-1/2} \exp\left\{-\tfrac{1}{2}(\underset{\sim}{x}-\underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\right\}, \quad \underset{\sim}{x} \in R_k. \tag{A.29}$$

Platí

$$E\underset{\sim}{X} = \underset{\sim}{\mu}\,, \qquad \operatorname{var} \underset{\sim}{X} = \underset{\sim}{\Sigma}\,. \tag{A.30}$$

k-rozměrné Wishartovo rozdělení (centrální) s n stupni volnosti s parametrickou maticí $\underset{\sim}{\Sigma}$ ($\underset{\sim}{\Sigma}$ - symetrická pozitivně definitní matice typu $k \times k$) má hustotu

$$r(\underset{\sim}{x}|n,\underset{\sim}{\Sigma}) = c_{k,n}(\det \underset{\sim}{\Sigma})^{-n/2} (\det \underset{\sim}{x})^{(n-k-1)/2}\ . \tag{A.31}$$
$$.\ \exp\left\{-\frac{1}{2}\operatorname{tr}(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{x})\right\} \quad \text{pro vš. } \underset{\sim}{x} = \begin{pmatrix} x_{11}, & \dots, & x_{1k} \\ \vdots & \ddots & \vdots \\ x_{k1}, & \dots, & x_{kk} \end{pmatrix},$$

kde

$$c_{k,n}^{-1} = 2^{nk/2}\,\pi^{k(k-1)/4} \prod_{j=1}^{k} \Gamma\left(\frac{n+1-j}{2}\right). \tag{A.32}$$

Je-li $\underset{\sim}{X}_1,\dots,\underset{\sim}{X}_n$ náhodný výběr z $N_k(\underset{\sim}{0},\underset{\sim}{\Sigma})$, pak náhodná matice

$$\underset{\sim}{S} = \sum_{i=1}^{n} \underset{\sim}{X}_i\,\underset{\sim}{X}_i'$$

má k-rozměrné Wishartovo rozdělení s n stupni volnosti s parametrickou matici $\underset{\sim}{\Sigma}$. Náhodná veličina

$$\underset{\sim}{a}'\underset{\sim}{S}\,\underset{\sim}{a}\,(\underset{\sim}{a}'\underset{\sim}{\Sigma}\,\underset{\sim}{a})^{-1},$$

kde $\underset{\sim}{a} = (a_1, \dots, a_k)$ je vektor konstant, má $\chi^2$-rozdělení s n stupni volnosti.

<u>k-rozměrné t-rozdělení (centrální) s n stupni volnosti s parametry $\mu$ a $\Sigma$</u> ($\mu \in R_k$, $\Sigma$ - symetrická pozitivně definitní matice typu $k \times k$) má hustotu

$$r(\underset{\sim}{x}|\underset{\sim}{\mu},\underset{\sim}{\Sigma}) = d_{kn}\left(1 + \frac{1}{n}(\underset{\sim}{x}-\underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{x}-\underset{\sim}{\mu})\right)^{-\frac{k+n}{2}}, \quad \underset{\sim}{x}=(x_1,\dots,x_k)' \in R_k \tag{A.33}$$

$$d_{kn}^{-1} = \frac{\Gamma(\frac{n+k}{2})(\det \underset{\sim}{\Sigma})^{-1/2}}{\Gamma(\frac{n}{2})\ (n\pi)^{k/2}} . \tag{A.34}$$

Platí

$$E\underset{\sim}{T} = \underset{\sim}{\mu} \tag{A.35}$$

a pro $n > 2$

$$\operatorname{var} \underset{\sim}{T} = \frac{n}{n-2}\underset{\sim}{\Sigma}. \tag{A.36}$$

Nechť má náhodný vektor $\underset{\sim}{Y} = (Y_1, \dots, Y_k)'$ rozdělení $N_k(\underset{\sim}{0}, \underset{\sim}{\Sigma})$, kde $\underset{\sim}{\Sigma}$ je regulární, nechť má náhodná veličina Z $\chi^2$-rozdělení s n stupni volnosti a $\underset{\sim}{Y}$ a Z jsou nezávislé. Definujme náhodný vektor $\underset{\sim}{X} = (X_1, \dots, X_k)'$ předpisem

$$X_i = \frac{Y_i}{\sqrt{Z}}\sqrt{n} + \mu_i \qquad i = 1, \dots, k. \tag{A.37}$$

Pak náhodný vektor $\underset{\sim}{X}$ má k-rozměrné t-rozdělení s n stupni volnosti a parametry $\underset{\sim}{\mu} = (\mu_1, \dots, \mu_k)$ a $\underset{\sim}{\Sigma}$.

<u>Dvojrozměrné Paretovo rozdělení s parametry $(r_1, r_2, a)$</u> ($r_1 < r_2$, $a > 0$) má hustotu

$$r(x_1, x_2 | r_1, r_2, a) = \frac{a(a+1)(r_2-r_1)^a}{(x_2-x_1)^{a+2}} \quad (x_1, x_2) \in R_2,\ x_1 < r_1,\ r_2 < x_2. \tag{A.34}$$

$$= 0 \quad \text{jinak.}$$

Pro $a > 1$ platí

$$EX_1 = \frac{ar_1 - r_2}{a-1}\,, \qquad EX_2 = \frac{ar_2 - r_1}{a-1} \tag{A.35}$$

a pro $a > 2$

$$\operatorname{var} X_1 = \operatorname{var} X_2 = \frac{a(r_2 - r_1)^2}{(a-1)^2(a-2)}\,. \tag{A.36}$$

## 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 Gřshick, 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, Massachusets, 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.