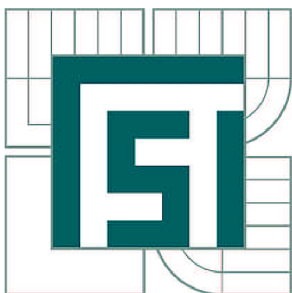


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY

FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

KVAZINORMY DISKRÉTNÍCH ROZDĚLENÍ PRAVDĚPODOBNOTI A JEJICH APLIKACE

QUASINORMS OF DISCRETE PROBABILITY DISTRIBUTIONS AND THEIR APPLICATIONS

DIZERTAČNÍ PRÁCE

DOCTORAL THESIS

AUTOR PRÁCE

AUTHOR

Ing. JAKUB ŠÁCHA

VEDOUCÍ PRÁCE

SUPERVISOR

doc. RNDr. ZDENĚK KARPÍŠEK, CSc.

BRNO 2012

Abstract

Dizertační práce je zaměřena na řešení statistického problému nalezení rozdělení pravděpodobnosti diskrétní náhodné veličiny na základě pozorovaných dat. Tyto odhady jsou získány minimalizací kvazinorem za předem daných omezujících podmínek kladených na hledaná rozdělení. Práce se dále zabývá odvozením intervalů spolehlivosti pro odhadované pravděpodobnosti. Obsahuje také praktické aplikace těchto metod.

Summary

Doctoral thesis is focused on solution of the statistical problem to find a probability distribution of a discrete random variable on the basis of the observed data. These estimates are obtained by minimizing quasi-norms with given constraints. The thesis further focuses on deriving confidence intervals for estimated probabilities. It also contains practical application of these methods.

Klíčová slova

kvazinorma, f-divergence, odhad diskrétního rozdělení pravděpodobnosti, obecné podmínky, interval spolehlivosti

Keywords

quasinorm, f-divergence, discrete probability distribution estimation, general conditions, confidence interval

ŠÁCHA, J. *Kvazinormy diskrétních rozdělení pravděpodobnosti a jejich aplikace*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2012. 85 s. Vedoucí dizertační práce doc. RNDr. Zdeněk Karpíšek, CSc..

Prohlášení

Prohlašuji, že jsem dizertační práci zpracoval samostatně, podle pokynů vedoucího práce a s použitím uvedené literatury.

V Brně dne 29. 8. 2012

Ing. Jakub Šácha

Děkuji doc. RNDr. Zdeňku Karpíškovi, CSc. za cenné rady a podněty k této dizertační práci. Rovněž bych rád poděkoval své ženě za její podporu.

Obsah

Úvod	5
1 Konvexní funkce	7
2 f -divergence diskrétních rozdělení pravděpodobnosti	9
3 Duální f -divergence	12
4 Kvazinorma diskrétního rozdělení pravděpodobnosti	14
5 Odhad diskrétního rozdělení pravděpodobnosti	17
5.1 Hellingerova kvazinorma	17
5.2 Shannonova kvazinorma	19
5.3 Pearsonova kvazinorma	20
6 Podmínky pro generující funkci f	22
7 Příklady	24
7.1 Příklad 1	24
7.2 Příklad 2	28
8 Aproximace základních rozdělení pravděpodobnosti pomocí kvazinorem	30
8.1 Binomické rozdělení	30
8.2 Poissonovo rozdělení	34
8.3 Normální rozdělení	38
8.4 Weibullovo rozdělení	44
9 Odhad rozdělení s obecnými lineárními podmínkami	51
9.1 Hellingerova kvazinorma	52
9.2 Shannonova kvazinorma	53
9.3 Pearsonova kvazinorma	54
9.4 Příklad	55
10 Intervalové odhady pravděpodobností p_j pro Shannonovu kvazinormu	58
10.1 Odvození rozdělení pro vektor Lagrangeových multiplikátorů	58
10.2 Odvození rozdělení pro vektor pravděpodobností	59
10.3 Příklad	60
11 Intervalové odhady pravděpodobností p_j pro Shannonovu kvazinormu a obecné podmínky	69
11.1 Odvození rozdělení pro vektor Lagrangeových multiplikátorů	69
11.2 Odvození rozdělení pro vektor pravděpodobností	70
11.3 Příklad	71

Závěr	77
Literatura	79
Autorovy publikace	83
Autorovy ostatní publikace	84
Seznam použitých zkratk a symbolů	85

Úvod

Jednou nejdůležitějších úloh a aplikací matematické statistiky je nalezení tvaru rozdělení pravděpodobnosti pozorované náhodné veličiny nebo náhodného vektoru. Tato práce se zabývá odhadem diskrétního rozdělení pravděpodobnosti z pozorovaných hodnot, který respektuje předem dané požadavky na vlastnosti tohoto rozdělení.

Stěžejním pojmem pro tyto odhady je pojem f -divergence dvou rozdělení pravděpodobnosti, která vyjadřuje jejich vzdálenost. Z f -divergence pak vychází v práci používaný pojem tzv. kvazinormy rozdělení, kterou je f -divergence diskrétního rozdělení pravděpodobnosti náhodné veličiny nabývající konečně mnoha hodnot a určitého pevného rozdělení na témže pravděpodobnostním prostoru. Jde vlastně o analogii zavedení indukované normy na lineárním prostoru s metrikou pomocí neutrálního prvku. Za tento "neutrální prvek" volíme rozdělení diskrétní se stejnými hodnotami pravděpodobnosti, protože se jedná o nejvíce neurčitý prvek na daném pravděpodobnostním prostoru, např. ve smyslu Shannonovy entropie. Kvazinormu tedy můžeme chápat jako jistou míru neurčitosti rozdělení. V této práci se nejvíce zabýváme třemi typy kvazinorem a sice Hellingerovou, Shannonovou a Pearsonovou. Pojem kvazinormy je pak rozšířen o tzv. duální kvazinormu.

Odhad rozdělení pravděpodobnosti pozorované veličiny je založen na myšlence minimalizovat kvazinormu tohoto odhadu při současném zachování rovnosti vybraných číselných charakteristik pozorovaného a odhadovaného rozdělení. Za tyto číselné charakteristiky volíme obecné momenty. Jinými slovy hledáme nejvíce neurčité rozdělení (ve smyslu kvazinormy), které splňuje zadané vedlejší podmínky (zachovává vybrané momenty). Proto hovoříme o tzv. pesimistických odhadech. V práci pak je popsáno rozšíření těchto momentových podmínek na obecné lineární podmínky.

Stanovení odhadu pozorovaného diskrétního rozdělení pravděpodobnosti je řešením optimalizační úlohy, kterou je minimalizace zvolené kvazinormy při omezeních daných vedlejšími podmínkami. Při absenci vedlejších podmínek je odhad totožný s maximálně entropickým rozdělením, ale přidáváním omezení se odhad blíží pozorovanému empirickému rozdělení. Přitom testujeme hypotézu o shodě pozorovaného a odhadovaného rozdělení a vedlejší podmínky přidáváme, dokud hypotéza o shodě není nezamítnuta.

Po popisu teoretických aspektů této problematiky je práce zaměřena také aplikačně. Jsou v ní konkrétní příklady na stanovení těchto odhadů, které jsou doplněny grafickou interpretací výsledků. Podrobně jsou v práci popsány aproximace významných diskrétních rozdělení (binomické a Poissonovo) i diskretizovaných spojitých rozdělení (normální a Weibullovo). Jsou použity tři nejvýznamnější kvazinormy a v závěrech je provedeno srovnání, která z nich je nejvhodnější pro dané rozdělení.

V závěrečných dvou kapitolách je práce zaměřena na nalezení aproximací asymptotických intervalových (konfidenčních) odhadů hodnot pravděpodobnostní funkce pozorované diskrétní náhodné veličiny pro Shannonovu kvazinormu. Pro konstrukci těchto odhadů jsou použity jak momentové, tak i obecné lineární podmínky. Řešení této úlohy je ilustrováno konkrétními příklady výpočtů 95% konfidenčních intervalů pro odhadnuté pravděpodobnosti.

Výsledky prezentované v práci jsou součástí řešení třídy úloh fitování pozorovaných

diskrétních rozdělení pravděpodobnosti v projektech: výzkumné centrum MŠMT České republiky čís. 1M06047 "Centrum pro jakost a spolehlivost výroby CQR", grantového projektu GAČR reg. č. P403/11/2085 "Konstrukce metod pro vícefaktorové měření komplexní podnikové výkonnosti ve vybraném odvětví" a výzkumného úkolu Akademie Sting v Brně "Podpora řízení malých a středních firem s využitím matematických metod".

1 Konvexní funkce

V dalším textu značí \mathbb{R} množinu reálných čísel a \mathbb{R}^* množinu reálných čísel rozšířenou o nevlastní prvky ∞ a $-\infty$. Množina $I(a, b) \subset \mathbb{R}^*$ je otevřený, polootevřený anebo uzavřený interval, který může být ohraničený anebo neohraničený a také rozšířený o nevlastní prvky $-\infty$ a $+\infty$. Definice, věty a důsledky uvedené v tomto oddílu jsou převzaty z [1], kde jsou také důkazy uvedených tvrzení.

Definice 1.1 Funkce $f: I(a, b) \rightarrow \mathbb{R}^*$ je **konvexní v bodě** $u_0 \in (a, b)$, jestliže je spojitá v intervalu $(a, b) \subset I(a, b)$, spojitá zprava v bodě a a spojitá zprava v bodě b , pokud tyto body patří do intervalu $I(a, b)$, a jestliže existuje $\lambda(u_0) \in \mathbb{R}$ tak, že $f(u) \geq f(u_0) + \lambda(u_0)(u - u_0)$ pro všechna $u \in I(a, b)$, $u \neq u_0$. Konvexní funkce f se nazývá **striktně konvexní v bodě** u_0 , jestliže nahradíme výše uvedenou neostrou nerovnost ostrou nerovností. Funkce f je **konvexní**, resp. **striktně konvexní**, v $I(a, b)$, je-li f konvexní, resp. striktně konvexní, v každém bodě $u \in I(a, b)$.

Jestliže $f: E \rightarrow \mathbb{R}^*$, kde $E \subset (\mathbb{R}^*)^N$ je konvexní množina, tak konvexnost funkce f definujeme podmínkou spojitosti na vnějšku množiny E , resp. na hranici (když ta patří do E) a podmínkou (1.1), kde ale $\lambda(u_0) \in \mathbb{R}$ a součin $\lambda(u_0)(u - u_0)$ chápeme jako skalární součin v $(\mathbb{R}^*)^N$.

Věta 1.1 Jestliže $f: I(a, b) \rightarrow \mathbb{R}^*$ je konvexní a $u' < u < u''$ jsou libovolné tři body z $I(a, b)$, pak

$$\frac{f(u) - f(u')}{u - u'} \leq \frac{f(u'') - f(u')}{u'' - u'} \leq \frac{f(u'') - f(u)}{u'' - u},$$

přičemž nerovnosti jsou ostré, když je f striktně konvexní v bodě u .

Důsledek 1.1 Jestliže je funkce f spojitá na $I(a, b)$ a splňuje Jensenovu nerovnost pro $m \geq 2$, pak je konvexní.

Z monotónnosti derivací vyplývají i určité lokální monotónnosti konvexních funkcí f a to, že existuje $u_0 \in (a, b)$ takové, že f je monotónní na intervalech (a, u_0) a (u_0, b) . Z toho plyne následující důsledek.

Důsledek 1.2 Jestliže je $f: I(a, b) \rightarrow \mathbb{R}$ konvexní, pak existují v \mathbb{R}^* limity

$$f(a) = \lim_{u \rightarrow a^+} f(u), f(b) = \lim_{u \rightarrow b^-} f(u),$$

přičemž takto (jednoznačně) rozšířená funkce $\langle a, b \rangle \rightarrow \mathbb{R}^*$ je konvexní. Rozšířená funkce je striktně konvexní v bodě $u_0 \in (a, b)$, právě když f je striktně konvexní v bodě u_0 .

Důsledek 1.3 Jestliže $f: I(a, b) \rightarrow \mathbb{R}^*$ je konvexní, tj. $f(a) > -\infty$, resp. $f(b) > -\infty$, pak a , resp. b jsou konečné, kde $f(a)$, resp. $f(b)$ definujeme pro $a \notin I(a, b)$ a $b \notin I(a, b)$ stejně jako v důsledku 1.2.

Věta 1.2 *Jestliže $f: I(a, b) \rightarrow \mathbb{R}^*$ je dvakrát diferencovatelná v $I(a, b)$, pak je konvexní, právě když $f''(u) \geq 0$ pro všechna $u \in (a, b)$, a striktně konvexní v bodě $u_0 \in (a, b)$, právě když $f''(u_0) > 0$.*

Věta 1.2 platí i pro funkci f více proměnných, když nezápornost druhé derivace nahradíme nezápornou definitností matice druhých parciálních derivací funkce f apod.

2 f -divergence diskretních rozdění pravděpodobnosti

Předpokládáme dále, že je dán diskretní pravděpodobnostní prostor (Ω, Σ, P) , ale až na výjimky vystačíme s diskretním pravděpodobnostním modelem (Ω, \mathbf{p}) , kde \mathbf{p} je hustota pravděpodobnosti P . Omezujeme se na konečný (případně spočetný) základní prostor Ω a σ -algebra Σ nehraje v našich úvahách podstatnou roli. Definice, věty a důsledky uvedené v tomto oddílu jsou převzaty z [1], kde jsou také důkazy uvedených tvrzení.

V matematické statistice mají velký význam číselné míry podobnosti (vzdálenosti) dvojic hustot \mathbf{p}, \mathbf{q} definované pomocí určité konvexní funkce $f: I(a, b) \rightarrow \mathbb{R}$. Pro tyto míry se v literatuře vžil název f -divergence. Jde přitom o f -divergenci pravděpodobnostních modelů $(\Omega, \mathbf{p}), (\Omega, \mathbf{q})$, resp. f -divergenci hustot \mathbf{p}, \mathbf{q} nebo jim příslušných pravděpodobností P, Q .

Jak vyplývá z důsledku 1.2, existuje právě jedno spojitě rozšíření $f(0), f(\infty)$ tak, že rozšířená funkce je konvexní na $\langle 0, \infty \rangle$ a $f(0) > \infty$. Bez újmy na obecnosti proto předpokládáme, že funkce f je definovaná na $\langle 0, \infty \rangle$, striktně konvexní v $u = 1$ a $f(0) < \infty$.

Věta 2.1 *Existuje limita*

$$f(*) = \lim_{u \rightarrow \infty} \frac{f(u)}{u} \in \mathbb{R}^*,$$

přičemž platí

$$-\infty < f(1) < f(0) + f(*) .$$

Dále platí

$$\lim_{\substack{u \rightarrow 0+ \\ v \rightarrow v_0}} u f\left(\frac{v}{u}\right) = v_0 f(*), \quad \lim_{\substack{u \rightarrow 0+ \\ v \rightarrow v_0}} v f\left(\frac{v}{u}\right) = v_0 f(0)$$

pro každé $v_0 \in (0, \infty)$.

Důsledek 2.1 *Jestliže je $f(u)$ konvexní na $(0, \infty)$, pak funkce*

$$\tilde{f}(u) = f(u) - f(1)$$

je také konvexní na $(0, \infty)$, přičemž $\tilde{f}(u) = 0$. Jestliže je $f(u)$ striktně konvexní v $u = 1$, je také $\tilde{f}(u)$ striktně konvexní v $u = 1$, přičemž $\tilde{f}(u) + \tilde{f}() > 0$.*

Tento důsledek nám zajišťuje, že můžeme požadovat $f(1) = 0$ bez omezení obecnosti v následující definici f -divergence.

Definice 2.1 *Nechť funkce $f(u)$ je konvexní $(0, \infty)$, striktně konvexní v $u = 1$, a $f(1) = 0$. f -divergencí pravděpodobnostních modelů $(\Omega, \mathbf{p}), (\Omega, \mathbf{q})$, resp. hustot \mathbf{p} a \mathbf{q} na (Ω, Σ, P) rozumíme funkcionál*

$$D_f(\mathbf{p}, \mathbf{q}) = \sum_{x \in \Omega} q(x) f\left(\frac{p(x)}{q(x)}\right),$$

kde klademe $0f(\frac{0}{0}) = 0$ a $0f(\frac{p}{0}) = pf(*)$ pro všechna $p \in (0, 1)$. Říkáme, že funkce f generuje danou f -divergenci.

Definice 2.2 Pravděpodobnostní modely (Ω, \mathbf{p}) , (Ω, \mathbf{q}) , resp. jejich hustoty \mathbf{p} a \mathbf{q} , jsou **ortogonální** a píšeme $\mathbf{p} \perp \mathbf{q}$, jestliže existují takové disjunktní množiny $E, F \subset \Omega$, že $\sum_{x \in E} p(x) = 1$ a $\sum_{x \in F} q(x) = 1$.

Další věta vyjadřuje, jak f -divergence měří podobnost pravděpodobnostních modelů.

Věta 2.2 Pro libovolnou f -divergenci platí nerovnost

$$0 \leq D_f(\mathbf{p}, \mathbf{q}) \leq f(0) + f(*),$$

přičemž obě dvě rovnosti nemohou nastat současně. Levá rovnost platí, právě když $\mathbf{p} = \mathbf{q}$, a pravá rovnost platí, právě když \mathbf{p} a \mathbf{q} jsou ortogonální a současně je $f(0) + f(*) < \infty$.

Z věty 2.2 plyne, že pravděpodobnostní modely (Ω, \mathbf{p}) a (Ω, \mathbf{q}) jsou si podobné, jestliže jejich f -divergence $D_f(\mathbf{p}, \mathbf{q})$ je blízká 0. Maximální podobnost je vlastně shoda pravděpodobností P a Q na téže σ -algebře podmnožin množiny Ω , což nastává právě tehdy, když \mathbf{p} a \mathbf{q} jsou totožné na Ω . Naopak modely jsou tím více nepodobné, čím více se jejich f -divergence blíží maximální hodnotě $f(0) + f(*)$. Maximálně divergentní jsou modely ortogonální.

Přehled nejpoužívanějších f -divergencí je v tabulce 1 a jejich vztahy popisuje následující věta. Místo $\sum_{x \in \Omega}$ dále píšeme \sum_x .

Věta 2.3 Pro všechny hustoty pravděpodobnosti \mathbf{p} , \mathbf{q} na Ω platí:

$$D_1(\mathbf{p}, \mathbf{q}) = \chi^1(\mathbf{p}, \mathbf{q}) = V(\mathbf{p}, \mathbf{q}),$$

$$D_{1/2}(\mathbf{p}, \mathbf{q}) = 2D^{1/2}(\mathbf{p}, \mathbf{q}),$$

$$\chi^2(\mathbf{p}, \mathbf{q}) = D^2(\mathbf{p}, \mathbf{q}),$$

$$V(\mathbf{p}, \mathbf{q}) < (\chi^\alpha(\mathbf{p}, \mathbf{q}))^{1/\alpha} \text{ pro } \alpha \in [1, \infty),$$

$$\ln \left(\frac{V(\mathbf{p}, \mathbf{q})^2}{4} + \frac{V(\mathbf{p}, \mathbf{q})^4}{36} \right) \leq I(\mathbf{p}, \mathbf{q}) \leq \ln \chi^2(\mathbf{p}, \mathbf{q}),$$

$$1 - \left(1 + \frac{V(\mathbf{p}, \mathbf{q})}{2} \right)^{\max(\alpha, 1-\alpha)} \left(1 - \frac{V(\mathbf{p}, \mathbf{q})}{2} \right)^{\min(\alpha, 1-\alpha)} \leq D^\alpha(\mathbf{p}, \mathbf{q}) \leq \frac{V(\mathbf{p}, \mathbf{q})}{2} \text{ pro } \alpha \in (0, 1).$$

Dále platí, že $D^{1/2}(\mathbf{p}, \mathbf{q})$, $\chi^1(\mathbf{p}, \mathbf{q})$, $V(\mathbf{p}, \mathbf{q})$ a $D_\beta(\mathbf{p}, \mathbf{q})^\beta$ jsou pro libovolné β metriky na množině všech hustot pravděpodobnosti na Ω .

V této práci se budeme nejvíce zaměřovat f -divergence z tabulky 2.

Tabulka 1: Nejpoužívanější třídy f -divergencí

$f(u)$	parametr	označení $D_f(\mathbf{p}, \mathbf{q})$ název	tvar $D_f(\mathbf{p}, \mathbf{q})$
$u \ln u$	-	$I(\mathbf{p}, \mathbf{q})$ <i>I-divergence</i>	$\sum_x p(x) \ln \frac{p(x)}{q(x)}$
$ u^\beta - 1 ^{1/\beta}$	$\beta \in (0, 1)$	$D_\beta(\mathbf{p}, \mathbf{q})$ <i>β-divergence</i>	$\sum_x \left p(x)^\beta - q(x)^\beta \right ^{1/\beta}$
	$\beta = 1/2$	$D_{1/2}(\mathbf{p}, \mathbf{q})$ <i>Hellingerova vzdálenost</i>	$2 \left(1 - \sum_x (p(x) q(x))^{1/2} \right)$
$ u - 1 ^\alpha$		$\chi^1(\mathbf{p}, \mathbf{q}), V(\mathbf{p}, \mathbf{q})$ <i>Totální variace</i>	$\sum_x p(x) - q(x) $
	$\alpha \in (1, \infty)$	$\chi^\alpha(\mathbf{p}, \mathbf{q})$ <i>χ^α-divergence</i>	$\sum_x \frac{ p(x) - q(x) ^\alpha}{q(x)^{\alpha-1}}$
	$\alpha = 2$	$\chi^2(\mathbf{p}, \mathbf{q})$ <i>χ^2-divergence</i>	$\sum_x \frac{(p(x) - q(x))^2}{q(x)}$
$\text{sign}(\alpha - 1)(u^\alpha - 1)$	$\alpha \in (0, 1)$	$D^\alpha(\mathbf{p}, \mathbf{q})$ <i>α-divergence</i>	$1 - \sum_x p(x)^\alpha q(x)^{1-\alpha}$
	$\alpha \in (1, \infty)$	$D^\alpha(\mathbf{p}, \mathbf{q})$ <i>α-divergence</i>	$\sum_x \frac{p(x)^\alpha}{q(x)^{1-\alpha}} - 1$

 Tabulka 2: Nejpoužívanější f -divergence

$f(u)$	$f(0) + f(*)$	název	$D_f(\mathbf{p}, \mathbf{q})$
$u \ln u$	∞	$I(\mathbf{p}, \mathbf{q})$ <i>I-divergence</i>	$\sum_x p(x) \ln \frac{p(x)}{q(x)}$
$(u^{1/2} - 1)^2$	2	$D_{1/2}(\mathbf{p}, \mathbf{q})$ <i>Hellingerova vzdálenost</i>	$\left(1 - \sum_x (p(x) q(x))^{1/2} \right)$
$(u - 1)^2$	∞	$\chi^2(\mathbf{p}, \mathbf{q})$ <i>χ^2-divergence</i>	$\sum_x \frac{(p(x) - q(x))^2}{q(x)}$

3 Duální f -divergence

f -divergence $D_f(\mathbf{p}, \mathbf{q})$ vyjadřuje vzdálenost mezi pravděpodobnostními modely \mathbf{p} a \mathbf{q} . Vzdálenost bývá obvykle symetrická funkce, definice f -divergence však tuto symetrii nezaručuje, neboť f -divergence nemusí být metrika. Obecně tedy platí $D_f(\mathbf{p}, \mathbf{q}) \neq D_f(\mathbf{q}, \mathbf{p})$, tj. vzdálenost \mathbf{p} od \mathbf{q} není stejná jako vzdálenost \mathbf{q} od \mathbf{p} ve smyslu téže f -divergence. Zamysleme se však, zda vzdálenost \mathbf{p} od \mathbf{q} ve smyslu f -divergence není rovna vzdálenosti \mathbf{q} od \mathbf{p} ve smyslu jiné F -divergence. Pokusme se najít vztah mezi funkcemi f a F a zjistit, jestli F splňuje požadavky na funkci generující f -divergenci. Vyjdeme z požadavku

$$D_f(\mathbf{p}, \mathbf{q}) = D_F(\mathbf{q}, \mathbf{p}).$$

Po dosazení z definice f -divergence dostáváme

$$\sum_x q(x) f\left(\frac{p(x)}{q(x)}\right) = \sum_x p(x) F\left(\frac{q(x)}{p(x)}\right).$$

Tuto rovnost nám zaručí rovnost uvnitř sum, ze které zavedením substituce $u = \frac{q(x)}{p(x)}$ dostáváme vztah mezi funkcemi f a F

$$F(u) = u f\left(\frac{1}{u}\right).$$

Záměnnou u a $\frac{1}{u}$ lze snadno ukázat, že vztah je symetrický mezi f a F , tj.

$$f(u) = u F\left(\frac{1}{u}\right).$$

Zbývá ukázat, že funkce F také splňuje požadavky na funkci generující f -divergenci, tj. konvexnost na $(0, \infty)$, striktní konvexnost v bodě 1 a $F(1) = 0$. Předpokládejme, že f je dvakrát diferencovatelná na $(0, \infty)$. Pak požadavky na funkci f můžeme psát jako $f''(u) \geq 0$ na $(0, \infty)$, $f''(1) > 0$ a $f(1) = 0$. Podle věty o derivaci složené funkce a derivaci součinu ze vztahu $F(u) = u f\left(\frac{1}{u}\right)$ plyne

$$F''(u) = \frac{1}{u^3} f''\left(\frac{1}{u}\right)$$

a tedy $F''(u) \geq 0$ na $(0, \infty)$, $F''(1) > 0$. $F(1) = 0$ platí zřejmě. Funkce F tedy také splňuje podmínky na funkci generující F -divergenci.

Předcházející úvahy můžeme shrnout do následující věty.

Věta 3.1 *Nechť D_f je f -divergence a funkce F je dána vztahem $F(u) = u f\left(\frac{1}{u}\right)$. Pak D_F je opět F -divergence a platí $D_f(\mathbf{p}, \mathbf{q}) = D_F(\mathbf{q}, \mathbf{p})$.*

Definice 3.1 *F -divergenci, která je generována funkcí F danou vztahem z věty 3.1 budeme nazývat **duální** k f -divergenci generované funkcí f .*

Relace duality je symetrická, tj. duální f -divergence k duální f -divergenci je původní f -divergence.

Zkusme nyní najít generující funkce duálních F -divergencí k používaným třídám divergencí z tabulky 1.

$$f(u) = u \ln u \rightarrow F(u) = u \cdot \frac{1}{u} \ln \frac{1}{u} = -\ln u,$$

$$f(u) = |u^\beta - 1|^{1/\beta} \rightarrow F(u) = u \cdot \left| \frac{1}{u} - 1 \right|^{1/\beta} = \left| u^\beta \frac{1}{u} - u^\beta \right|^{1/\beta} = |u^{1-\beta}|^{1/\beta} = f(u),$$

$$f(u) = |u - 1|^\alpha \rightarrow F(u) = u \cdot \left| \frac{1}{u} - 1 \right|^\alpha = u \cdot \frac{|1 - u|^\alpha}{u^\alpha} = \frac{|u - 1|^\alpha}{u^{\alpha-1}}.$$

Třída generovaná funkcí $f(u) = |u^\beta - 1|^{1/\beta}$ je symetrická a patří do ní používaná Hellingerova vzdálenost.

Rozšířme nyní tabulku 2 nejpoužívanějších f -divergencí o jejich duální protějšky. V tabulce 3 jsou uvedeny vzájemně duální f -divergence. Konkrétně I -divergence $I(\mathbf{p}, \mathbf{q})$ je duální s logaritmickou divergencí $I_D(\mathbf{p}, \mathbf{q})$, χ^2 -divergence $\chi^2(\mathbf{p}, \mathbf{q})$ je duální s Pearsonovou divergencí $\chi_D^2(\mathbf{p}, \mathbf{q})$ a Hellingerova vzdálenost $D_{1/2}(\mathbf{p}, \mathbf{q})$ je duální sama se sebou.

Tabulka 3: Nejpoužívanější f -divergence a jejich dualní protějšky

$f(u)$	označení název	$D_f(\mathbf{p}, \mathbf{q})$
$u \ln u$	$I(\mathbf{p}, \mathbf{q})$ <i>I-divergence</i>	$\sum_x p(x) \ln \frac{p(x)}{q(x)}$
$(u^{1/2} - 1)^2$	$D_{1/2}(\mathbf{p}, \mathbf{q})$ <i>Hellingerova vzdálenost</i>	$\left(1 - \sum_x (p(x)q(x))^{1/2} \right)$
$(u - 1)^2$	$\chi^2(\mathbf{p}, \mathbf{q})$ <i>χ^2-divergence</i>	$\sum_x \frac{(p(x) - q(x))^2}{q(x)} = \sum_x \frac{p^2(x)}{q(x)} - 1$
$-\ln u$	$I_D(\mathbf{p}, \mathbf{q})$ <i>Logaritmická divergence</i>	$\sum_x q(x) \ln \frac{q(x)}{p(x)}$
$\frac{(u-1)^2}{u}$	$\chi_D^2(\mathbf{p}, \mathbf{q})$ <i>Pearsonova divergence</i>	$\sum_x \frac{(q(x) - p(x))^2}{p(x)} = \sum_x \frac{q^2(x)}{p(x)} - 1$

4 Kvazinorma diskrétního rozdělení pravděpodobnosti

Významnou roli v aplikacích metod matematické statistiky při intervalových odhadech parametrů a parametrických testech statistických hypotéz hraje nalezení tvaru rozdělení pravděpodobnosti pozorované náhodné veličiny nebo náhodného vektoru. Z výše uvedeného pojmu f -divergence lze vyvodit postupy, které umožňují takové rozdělení odhadnout [7], [8], [9]. Tyto postupy však musí obvykle respektovat další podmínky kladené na toto rozdělení. Jde nejčastěji o podmínky dané apriorním stanovením hodnot vybraných číselných charakteristik, např. střední hodnoty, rozptylu apod. Základní ideou je najít takové rozdělení, které je v jistém smyslu blízké nějakému vhodně zvolenému pevnému rozdělení. Přesněji jde o nalezení rozdělení, které je s takovým pevným rozdělením totožné při absenci vedlejších podmínek, ale s přidáváním podmínek se od tohoto pevného rozdělení postupně vzdaluje, ale s minimalizací zvolené f -divergence hledaného a daného pevného rozdělení. Poznamenejme ještě, že místo pojmů pravděpodobnostní model a hustota užívaných v [1], budeme pro jednoduchost a vzhledem ke směřování výsledků hovořit ve stejném smyslu o rozdělení pravděpodobnosti.

Věta 4.1 *Nechť (Ω, Σ, P) je konečný pravděpodobnostní prostor, kde P je libovolná pravděpodobnostní míra na Ω a $D_f(\mathbf{p}, \mathbf{q})$ je f -divergence rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ a $\mathbf{q} = (q_1, \dots, q_m)$, $m > 1$ z tohoto prostoru. Označme*

$$V(\mathbf{q}) = \int_S D_f(\mathbf{p}, \mathbf{q}) dS,$$

kde $S = \{\mathbf{p} \in \mathbb{R}^m : \forall p_j \geq 0, \sum_{j=1}^m p_j = 1\}$, integrál f -divergencí všech rozdělení \mathbf{p} od nějakého pevně zvoleného rozdělení \mathbf{q} . Jestliže existuje funkce $V(\mathbf{q})$, $\mathbf{q} \in S$ a funkce

$$G(q_j) = \frac{\partial V(\mathbf{q})}{\partial q_j} = \int_S \frac{\partial D_f(\mathbf{p}, \mathbf{q})}{\partial q_j} dS,$$

(tj. oba uvedené integrály konvergují), dále jestliže existuje derivace $G'(q_j)$ v $\langle 0, 1 \rangle$ a funkce f má spojitou druhou derivaci v $(0, \infty)$, pak $V(\mathbf{q})$ nabývá absolutního minima na S v rozdělení pravděpodobnosti

$$\mathbf{q} = \mathbf{p}_0 = \left(\frac{1}{m}, \dots, \frac{1}{m} \right).$$

S je vlastně množina všech rozdělení z (Ω, Σ, P) a oba integrály jsou integrály 1. druhu po nadploše S dimenze $m - 1$ v \mathbb{R}^m .

Důkaz: Označme

$$\Lambda(\mathbf{q}, \lambda) = V(\mathbf{q}) + \lambda \left(\sum_{j=1}^m q_j - 1 \right).$$

Lagrangeovu funkci pro vázaný extrém funkce $V(\mathbf{q})$ za podmínky $\sum_{j=1}^m q_j = 1$, tj. extrém na nadploše S . Funkce f je konvexní (jde o f -divergenci), takže z předpokladu existence její druhé derivace je $f'' \geq 0$ v $(0, \infty)$. Protože

$$\frac{\partial D_f(\mathbf{p}, \mathbf{q})}{\partial q_j} = f\left(\frac{p_j}{q_j}\right) - \frac{p_j}{q_j} f'\left(\frac{p_j}{q_j}\right),$$

pak

$$G'(q_j) = \int_S \frac{d}{dq_j} \left(f\left(\frac{p_j}{q_j}\right) - \frac{p_j}{q_j} f'\left(\frac{p_j}{q_j}\right) \right) dS = \int_S \frac{p_j^2}{q_j^3} f''\left(\frac{p_j}{q_j}\right) dS > 0$$

a odtud funkce $G(q_j)$ je rostoucí na $[0, 1]$. Protože

$$\frac{\partial \Lambda(\mathbf{q}, \lambda)}{\partial q_j} = G(q_j) + \lambda,$$

existuje na $[0, 1]$ jediný kořen $q_j = G^{-1}(-\lambda)$ rovnice $G(q_j) + \lambda = 0$. Z podmínky $\sum_{j=1}^m q_j = 1$ pak dostaneme $q_j = \frac{1}{m}$, $j = 1, \dots, m$. Dále je

$$\frac{\partial^2 V(\mathbf{q})}{\partial q_j \partial q_k} = \frac{\partial^2 \Lambda(\mathbf{q}, \lambda)}{\partial q_j \partial q_k} = \begin{cases} G'(q_j) > 0 & \text{pro } j = k, \\ 0 & \text{pro } j \neq k, \end{cases}$$

$j = 1, \dots, m$ a $k = 1, \dots, m$. Z toho plyne, že Hessova matice funkce $V(\mathbf{q})$ je diagonální a pozitivně definitní. Jacobiho matice pro jedinou podmínku $\sum_{j=1}^m q_j - 1 = 0$ má hodnotu 1, takže $V(\mathbf{q})$ má za dané podmínky v $\mathbf{q} = \mathbf{p}_0 = \left(\frac{1}{m}, \dots, \frac{1}{m}\right)$ absolutní minimum. \square

Podle věty 4.1 se pro pozorovanou diskrétní náhodnou veličinu X nabývající konečně mnoha hodnot nabízí zvolit za pevné diskrétní rozdělení, od něhož budeme vzdálenost pomocí f -divergence měřit, rozdělení $\mathbf{p}_0 = \left(\frac{1}{m}, \dots, \frac{1}{m}\right)$. Toto rozdělení je všem rozdělením z (Ω, Σ, P) nejbližší ve smyslu minima integrálu $V(\mathbf{q})$ všech f -divergencí a navíc má také maximální možnou neurčitost vyjádřenou pomocí Shannonovy entropie. To nás opravňuje k zavedení následujícího pojmu [7]. Jde přitom o jistou analogii zavedení indukované normy na lineárním prostoru s metrikou pomocí neutrálního prvku.

Definice 4.1 *Nechť $\mathbf{p} = (p_1, \dots, p_m)$ a $\mathbf{p}_0 = \left(\frac{1}{m}, \dots, \frac{1}{m}\right)$, $m > 1$ jsou diskrétní rozdělení z pravděpodobnostního prostoru (Ω, Σ, P) a D_f je f -divergence definovaná na daném prostoru. **Kvazinormou** rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ na (Ω, Σ, P) rozumíme f -divergenci $D_f(\mathbf{p}, \mathbf{p}_0)$.*

Věta 4.2 *Pro každou kvazinormu $D_f(\mathbf{p}, \mathbf{p}_0)$ platí:*

$$a) D_f(\mathbf{p}, \mathbf{p}_0) = \frac{1}{m} \sum_{j=1}^m f(mp_j),$$

b) $D_f(\mathbf{p}, \mathbf{p}_0)$ je nezáporná symetrická funkce proměnných p_j , $j = 1, \dots, m$.

Výsledky z kapitoly o duálních f -divergencích nás opravňují k rozšíření tohoto pojmu na kvazinormu.

Definice 4.2 *Nechť $D_f(\mathbf{p}, \mathbf{p}_0)$ kvazinorma rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ na (Ω, Σ, P) a funkce F je dána vztahem $F(u) = uf\left(\frac{1}{u}\right)$, pak **duální kvazinormou** k $D_f(\mathbf{p}, \mathbf{p}_0)$ budeme nazývat kvazinormu $D_F(\mathbf{p}, \mathbf{p}_0)$.*

V souladu s předcházejícími výsledky můžeme shrnout vlastnosti duální kvazinormy.

Věta 4.3 *Nechť $D_f(\mathbf{p}, \mathbf{p}_0)$ je kvazinorma a $D_F(\mathbf{p}, \mathbf{p}_0)$ je kvazinorma k ní duální. Pak platí:*

- a) $D_F(\mathbf{p}, \mathbf{p}_0) = D_f(\mathbf{p}_0, \mathbf{p})$, $D_f(\mathbf{p}, \mathbf{p}_0) = D_F(\mathbf{p}_0, \mathbf{p})$,
b) $D_F(\mathbf{p}, \mathbf{p}_0) = \frac{1}{m} \sum_{j=1}^m F(mp_j) = \sum_{j=1}^m p_j f\left(\frac{1}{mp_j}\right)$.

Tabulka 4 shrnuje nejpoužívanější kvazinormy včetně jejich duálních protějšků. Značení uvedených kvazinorem odpovídá dalším kapitolám v práci.

Tabulka 4: Nejpoužívanější kvazinormy

$f(u)$	označení název	$D_f(\mathbf{p}, \mathbf{p}_0)$
$u \ln u$	$S(\mathbf{p}, \mathbf{p}_0)$ Shannonova kvazinorma	$\frac{1}{m} \sum_{j=1}^m mp_j \ln(mp_j) = \sum_{j=1}^m p_j \ln p_j + \ln m$
$(u^{1/2} - 1)^2$	$H(\mathbf{p}, \mathbf{p}_0)$ Hellingerova kvazinorma	$\sum_{j=1}^m \left(\sqrt{p_j} - \sqrt{\frac{1}{m}} \right)^2 = 2 - \frac{2}{\sqrt{m}} \sum_{j=1}^m \sqrt{p_j}$
$(u - 1)^2$	$\chi^2(\mathbf{p}, \mathbf{p}_0)$ Kvadratická kvazinorma	$\frac{1}{m} \sum_{j=1}^m (mp_j - 1)^2 = m \sum_{j=1}^m p_j^2 - 1$
$-\ln u$	$L(\mathbf{p}, \mathbf{p}_0)$ Logaritmická kvazinorma	$-\frac{1}{m} \sum_{j=1}^m \ln(mp_j) = -\sum_{j=1}^m \ln p_j - \ln m$
$\frac{(u-1)^2}{u}$	$P(\mathbf{p}, \mathbf{q}_0)$ Pearsonova kvazinorma	$\frac{1}{m} \sum_{j=1}^m \frac{(mp_j - 1)^2}{mp_j} = \frac{1}{m^2} \sum_{j=1}^m \frac{1}{p_j} - 1$

5 Odhad diskrétního rozdělení pravděpodobnosti

Předpokládáme, že pozorovaná diskrétní náhodná veličina X na (Ω, Σ, P) , jejíž rozdělení pravděpodobnosti chceme odhadnout (fitovat), nabývá nejvýše konečně mnoha různých hodnot $x_j^* \in \mathbb{R}$, (tj. $\Omega = \{x_1^*, \dots, x_m^*\} \subset \mathbb{R}$) s neznámými pravděpodobnostmi

$$p_j = P(X = x_j^*), \quad j = 1, \dots, m, \quad m > 1.$$

Pozorováním náhodné veličiny X získáme statistický soubor (x_1, \dots, x_n) a jeho roztríděním dostaneme roztríděný statistický soubor

$$\left(\left(x_1^*, \frac{f_1}{n} \right), \dots, \left(x_m^*, \frac{f_m}{n} \right) \right),$$

kde f_j je absolutní četnost pozorované hodnoty x_j^* . Dále předpokládáme, že $n > m$ a $f_j > 0$ pro všechna $j = 1, \dots, m$ (jestliže $f_j = 0$, pak j -tou třídu vynecháme). Pro odhad rozdělení \mathbf{p} požadujeme, aby toto rozdělení navíc splňovalo nějaké zadané podmínky, jejichž počet je $K \geq 1$. Mezi tyto podmínky nezařazujeme zřejmou podmínku $\sum_{j=1}^m p_j = 1$, ale další např. momentové podmínky

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 1, \dots, K.$$

Hledáme pak takové rozdělení \mathbf{p} , které má minimální kvazinormu $D_f(\mathbf{p}, \mathbf{p}_0)$.

Pro odhady diskrétních rozdělení za momentových podmínek pomocí minimálních kvazinorem volíme Hellingerovu, Pearsonovu a Shannonovu kvazinormu.

5.1 Hellingerova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Hellingerovu kvazinormu za K počátečních momentových podmínek**

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 0, \dots, K,$$

jestliže jeho Hellingerova kvazinorma

$$H(\mathbf{p}, \mathbf{p}_0) = 2 - \frac{2}{\sqrt{m}} \sum_{j=1}^m \sqrt{p_j}$$

je minimální pro

$$M_k = \sum_{j=1}^m p_j x_j^{*k}, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme [8]

$$p_j(\boldsymbol{\lambda}) = \frac{1}{m \left(\sum_{k=0}^K \lambda_k x_j^{*k} \right)^2}, \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$, jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = H(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j x_j^{*k} - M_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce, anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $H_K = \min H(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Hellingerovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$H_K = 2 - 2 \sum_{j=1}^m \left| \sum_{k=0}^K \lambda_k x_j^{*k} \right|.$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $D_0 = 0$. Speciálně pro $K = m - 1$ jde o interpolaci $p_j = \frac{f_j}{n}$, $j = 1, \dots, m$ a $H_{m-1} = 2 - \frac{2}{\sqrt{mn}} \sum_{j=1}^m \sqrt{f_j}$. Platí, že $H_0 \leq \dots \leq H_{m-1}$.

Má-li pozorovaná náhodná veličina X empirické rozdělení $\mathbf{f} = (\frac{f_1}{n}, \dots, \frac{f_m}{n})$, pak statistika

$$\chi^2(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) = \sum_{j=1}^m \frac{(f_j - np_j(\boldsymbol{\lambda}))^2}{np_j(\boldsymbol{\lambda})} = \frac{1}{n} \sum_{j=1}^m \frac{f_j^2}{p_j(\boldsymbol{\lambda})} - n$$

má pro $n \rightarrow \infty$ asymptoticky rozdělení chí-kvadrát s $m - K - 1$ stupni volnosti. Asymptotickou vlastnost můžeme použít k testování vhodnosti nalezeného rozdělení pravděpodobnosti $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$. Pro praktické použití požadujeme [2], aby bylo $np_j(\boldsymbol{\lambda})$ pro všechna $j = 1, \dots, m$. Pro test vhodnosti odhadnutého rozdělení pravděpodobnosti $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ můžeme také využít přímo Hellingerovu vzdálenost. Jde o nepříliš známý tzv. **Pitmanův - Hellingerův** test shody [12], [14], který spočívá ve skutečnosti, že statistika

$$4nH(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) = 4n \sum_{j=1}^m \sqrt{p_j(\boldsymbol{\lambda})} - \sqrt{\frac{f_j}{n}}$$

má pro $n \rightarrow \infty$ asymptoticky rozdělení chí-kvadrát s $m - K - 1$ stupni volnosti. Postupným přidáváním momentových podmínek a opakovaným odhadem rozdělení pravděpodobnosti pomocí minimální Hellingerovy kvazinormy lze určit minimální potřebný počet K těchto podmínek tak, aby platilo $\chi^2(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, resp. $4nH(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, kde χ_{n-1}^2 je $(1 - \alpha)$ -kvantil rozdělení chí-kvadrát s daným počtem stupňů volnosti pro hladinu významnosti α .

5.2 Shannonova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Shannonovu kvazinormu za K počátečních momentových podmínek**

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 0, \dots, K,$$

jestliže jeho Shannonova kvazinorma

$$S(\mathbf{p}, \mathbf{p}_0) = \sum_{j=1}^m p_j \ln p_j + \ln m$$

je minimální pro

$$M_k = \sum_{j=1}^m p_j x_j^{*k}, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme [5], [6]

$$p_j(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \lambda_k x_j^{*k}\right), \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$, jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = S(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j x_j^{*k} - M_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce

$$\sum_{j=1}^m \exp\left(-1 + \sum_{k=0}^K \lambda_k x_j^{*k}\right) x_j^{*\nu} = M_\nu, \quad \nu = 1, \dots, K,$$

anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $S_K = \min S(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Hellingerovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$S_K = \ln m - \sum_{j=1}^m \left(1 + \sum_{k=0}^K \lambda_k x_j^{*k}\right) \exp\left(-1 - \sum_{k=0}^K \lambda_k x_j^{*k}\right).$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $S_0 = 0$. Specialně pro $K = m - 1$ jde o interpolaci $p_j = \frac{f_j}{n}$, $j = 1, \dots, m$ a $S_{m-1} = \frac{1}{n} \sum_{j=1}^m f_j \ln f_j$. Platí, že $S_0 \leq \dots \leq S_{m-1}$.

Dále pak můžeme aplikovat Pearsonův, resp. Pitmanův-Hellingerův, test shody rozdělení $\mathbf{p}(\boldsymbol{\lambda})$ stejně jako v případě Hellingerovy kvazinormy. Postupným přidáváním momentových podmínek a opakovaným odhadem rozdělení pravděpodobnosti pomocí minimální Shannonovy kvazinormy lze určit minimální potřebný počet K těchto podmínek tak, aby platilo, $\chi^2(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, resp. $4nD(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, kde χ_{n-1}^2 je $(1 - \alpha)$ -kvantil rozdělení chí-kvadrát s daným počtem stupňů volnosti pro hladinu významnosti α .

5.3 Pearsonova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Pearsonovu kvazinormu za K počátečních momentových podmínek**

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 0, \dots, K,$$

jestliže jeho Pearsonova kvazinorma

$$P(\mathbf{p}, \mathbf{p}_0) = \frac{1}{m^2} \sum_{j=1}^m \frac{1}{p_j} - 1$$

je minimální pro

$$M_k = \sum_{j=1}^m p_j x_j^{*k}, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme [16], [19]

$$p_j(\boldsymbol{\lambda}) = \frac{1}{m \sqrt{\sum_{k=0}^K \lambda_k x_j^{*k}}}, \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$, jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = P(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j x_j^{*k} - M_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce, anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $P_K = \min P(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Shannonovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$P_K = \frac{1}{m} \sum_{j=1}^m \sqrt{\sum_{k=0}^K \lambda_k x_j^{*k}} - 1.$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $P_0 = 0$. Specialně pro $K = m - 1$ jde o interpolaci $p_j = \frac{f_j}{n}$, $j = 1, \dots, m$ a $P_{m-1} = \frac{n}{m^2} \sum_{j=1}^m \frac{1}{f_j}$. Platí, že $P_0 \leq \dots \leq P_{m-1}$.

Dále pak můžeme aplikovat Pearsonův, resp. Pitmanův-Hellingerův, test shody rozdělení $\mathbf{p}(\boldsymbol{\lambda})$ stejně jako v případě Hellingerovy nebo Shannonovy kvazinormy. Postupným přidáváním momentových podmínek a opakovaným odhadem rozdělení pravděpodobnosti pomocí minimální Shannonovy kvazinormy lze určit minimální potřebný počet K těchto podmínek tak, aby platilo, $\chi^2(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, resp. $4nD(\mathbf{f}, \mathbf{p}(\boldsymbol{\lambda})) \leq \chi_{n-1}^2$, kde χ_{n-1}^2 je $(1 - \alpha)$ -kvantil rozdělení chí-kvadrát s daným počtem stupňů volnosti pro hladinu významnosti α .

6 Podmínky pro generující funkci f

Odhad rozdělení spočívá v řešení minimalizační úlohy s účelovou funkcí

$$\frac{1}{m} \sum_{j=1}^m f(mp_j)$$

v proměnných p_j a s omezeními ve tvaru lineárních rovnic

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 1, \dots, K.$$

Existenci a jednoznačnost řešení zaručuje konvexnost funkce f [4]. Proměnné p_j v modelu mají význam pravděpodobností a je tedy potřeba zajistit jejich nezápornost. Tu přímo nezaručují momentové podmínky a obecně je tedy nutné přidat do modelu podmínky nezápornosti $p_j \geq 0$. Tyto podmínky nezápornosti by však bylo možné nahradit nějakou vlastností funkce f vytvářející příslušnou kvazinormu. Jinak řečeno, ptejme se, co musí splňovat funkce f , aby byla zaručena nezápornost p_j .

Zapišme Lagrangeovu funkci pro naši úlohu

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = \frac{1}{m} \sum_{j=1}^m f(mp_j) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j x_j^{*k} - M_k \right).$$

Podmínkou minima je nulovost gradientu Lagrangeovy funkce, tj.

$$\frac{\partial \Lambda}{\partial p_j} = f'(mp_j) + \sum_{k=0}^K \lambda_k x_j^{*k} = 0,$$

$$f'(mp_j) = - \sum_{k=0}^K \lambda_k x_j^{*k} = 0.$$

Po úpravě dostáváme

$$p_j = \frac{1}{m} (f')^{-1} \left(- \sum_{k=0}^K \lambda_k x_j^{*k} \right) \geq 0,$$

kde $(f')^{-1}$ je inverzní funkce k derivaci funkce f .

Předcházející úvahy můžeme shrnout do následující věty.

Věta 6.1 *Jestliže $(f')^{-1}(v) \geq 0$ pro všechna $v \in \mathbb{R}$, pak pro odhad vzniklý minimalizací kvazinormy příslušné funkce f s momentovými omezeními platí $p_j \geq 0$, $j = 1 \dots m$.*

Nyní zjistíme, které z používaných kvazinorem splňují předpoklady této věty.

Pro Helingerovu kvazinormu $f(u) = (\sqrt{u} - 1)^2$ je

$$f'(u) = 1 - \frac{1}{\sqrt{u}},$$

takže

$$(f')^{-1}(v) = \frac{1}{(v-1)^2} \geq 0$$

a podmínka věty 6.1 je splněna.

Pro kvadratickou kvazinormu $f(u) = (u-1)^2$ je

$$f'(u) = 2(u-1),$$

takže

$$(f')^{-1}(v) = \frac{v}{2} + 1$$

a podmínka věty 6.1 není splněna.

Pro Pearsonovu kvazinormu $f(u) = \frac{(u-1)^2}{u}$ je

$$f'(u) = 1 - \frac{1}{u^2},$$

takže

$$(f')^{-1}(v) = \frac{1}{\sqrt{1-v}} \geq 0$$

a podmínka věty 6.1 je splněna.

Pro Shannonovu kvazinormu $f(u) = u \ln u$ je

$$f'(u) = 1 + \ln u,$$

takže

$$(f')^{-1}(v) = e^{v-1} \geq 0$$

a podmínka věty 6.1 je splněna.

Pro logaritmickou kvazinormu $f(u) = -\ln u$ je

$$f'(u) = -\frac{1}{u} \text{ pro } u \in (0, \infty),$$

takže

$$(f')^{-1}(v) = -\frac{1}{v} \geq 0 \text{ pro } v \in (-\infty, 0),$$

a podmínka věty 6.1 je splněna.

7 Příklady

7.1 Příklad 1

Pozorováním diskrétní náhodné veličiny X jsme získali statistický soubor o rozsahu $n = 180$. Po jeho roztřídění jsme obdrželi diskrétní empirické rozdělení náhodné veličiny X dané následující tabulkou 5, kde jsou x_j^* středy tříd a f_j pozorované absolutní četnosti [15].

Tabulka 5: Pozorování

x_j^*	1	2	3	4	5	6	7
f_j	15	36	21	15	27	42	24

Hledáme minimum Hellingerovy, Shannonovy a Personovy kvazinormy za vedlejších podmínek daných prvními pěti obecnými momenty

$$M_0 = \frac{1}{n} \sum_{j=1}^m f_j = 1,$$

$$M_1 = \frac{1}{n} \sum_{j=1}^m f_j x_j^* = 12,75,$$

$$M_2 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*2} = 65,85,$$

$$M_3 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*3} = 375,15,$$

$$M_4 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*4} = 2251,05.$$

Výpočty byly provedeny pomocí speciálně vytvořeného softwaru s optimalizačním jádrem z programu GAMS. Výsledky za postupného přidávání momentových podmínek pro všechny kvazinormy jsou ilustrovány v tabulce 6.

V tabulce 7 jsou vypočtené odhady absolutních četností, kde horní index odhadnuté zaokrouhlené četnosti f_j^H , resp. f_j^S , resp. $f_j^P = np_j(\boldsymbol{\lambda})$ odpovídá prvnímu písmenu názvu dané kvazinormy.

Výsledky chi-kvadrat testu a Pitman-Hellingerova testu dobré shody pozorovaného a fitovaného rozdělení jsou v tabulce 8, kde "Ano" znamená, že hypotézu o shodě rozdělení nezamítáme, "Ne" znamená zamítnutí. Z výsledků je vidět, že oba testy mají podobnou citlivost a při $K = 4$ je dosaženo již dosaženo dobré shody. Grafická ilustrace je na obrázku 1.

Tabulka 6: Hodnoty kvazinorem

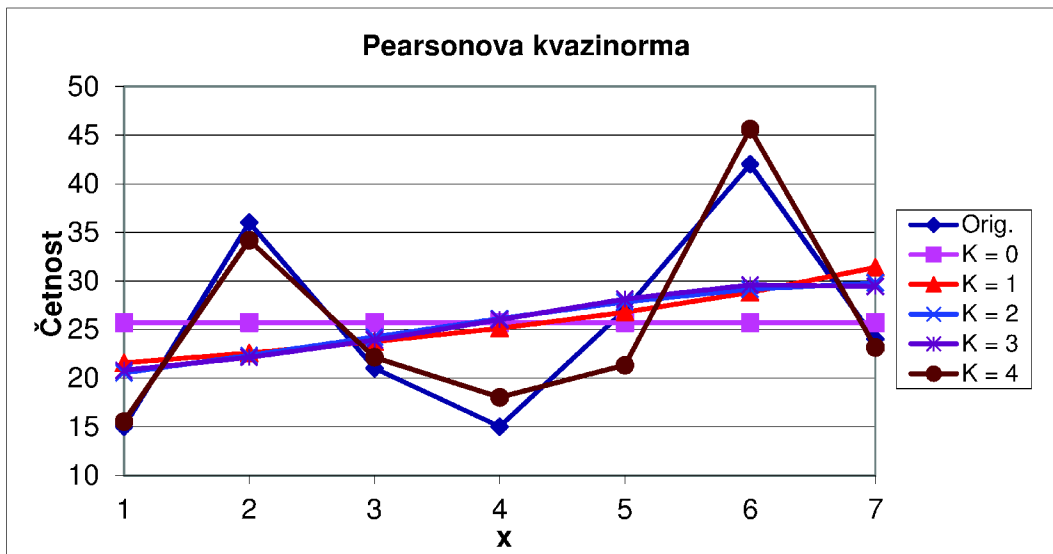
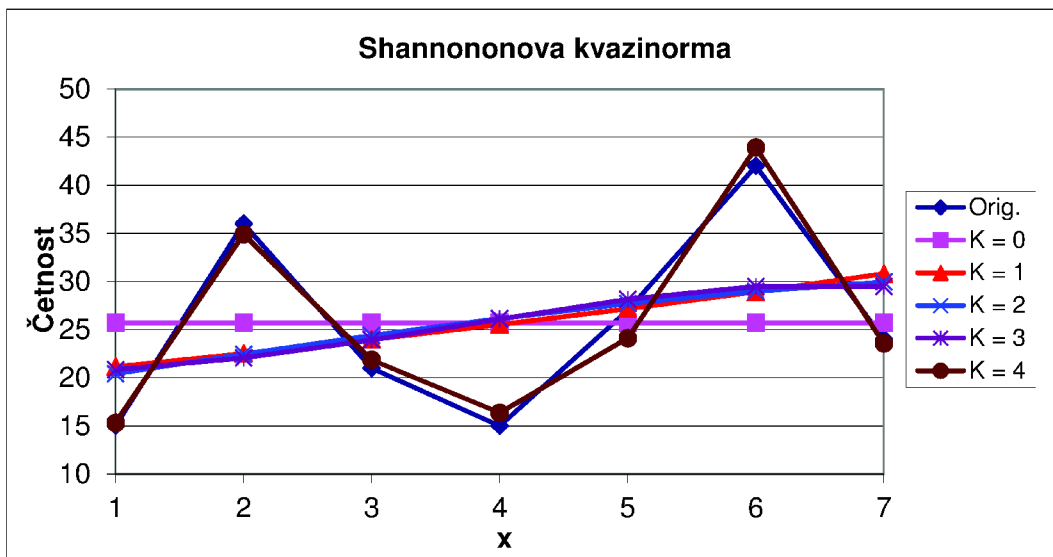
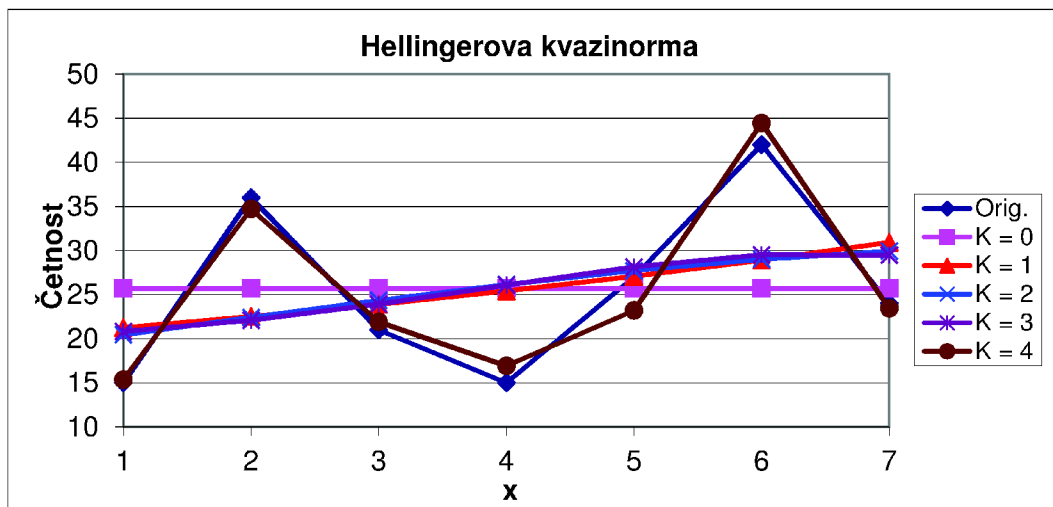
Kvazinorma \ K	0	1	2	3	4
$D(\mathbf{p}, \mathbf{p}_0)$	0	0,0039	0,0041	0,0041	0,0318
$S(\mathbf{p}, \mathbf{p}_0)$	0	0,0078	0,0080	0,0081	0,0647
$P(\mathbf{p}, \mathbf{p}_0)$	0	0,0156	0,0168	0,0169	0,1250

Tabulka 7: Výsledky odhadů

x_j^*		1	2	3	4	5	6	7
f_j		15	36	21	15	27	42	24
f_j^H	$K = 0$	25,7	25,7	25,7	25,7	25,7	25,7	25,7
	$K = 1$	21,3	22,5	23,9	25,4	27,1	28,9	30,9
	$K = 2$	20,4	22,4	24,3	26,1	27,7	29,0	29,9
	$K = 3$	20,8	22,1	23,9	26,1	28,1	29,5	29,4
	$K = 4$	15,4	34,7	21,9	16,9	23,2	44,5	23,4
f_j^S	$K = 0$	25,7	25,7	25,7	25,7	25,7	25,7	25,7
	$K = 1$	21,1	22,5	24,0	25,5	27,2	28,9	30,8
	$K = 2$	20,4	22,4	24,4	26,1	27,7	29,0	29,9
	$K = 3$	20,8	22,1	23,9	26,1	28,1	29,5	29,5
	$K = 4$	15,3	34,9	21,9	16,4	24,1	43,9	23,6
f_j^P	$K = 0$	25,7	25,7	25,7	25,7	25,7	25,7	25,7
	$K = 1$	21,6	22,6	23,8	25,1	26,8	28,8	31,4
	$K = 2$	20,5	22,3	24,2	26,1	27,8	29,1	29,8
	$K = 3$	20,8	22,1	23,9	26,0	28,1	29,6	29,4
	$K = 4$	15,5	34,2	22,1	18,0	21,3	45,6	23,2

Tabulka 8: Výsledky testů

Kvazinorma	K	Chi-kvadrát test			Pitman-Hellingerův test		
		Krit.	P -hod.	Hyp.	Krit.	P -hod.	Hyp.
$D(\mathbf{p}, \mathbf{p}_0)$	0	24,4017	0,00044	Ne	24,0245	0,00052	Ne
	1	22,0124	0,00052	Ne	20,9359	0,00083	Ne
	2	21,8920	0,00021	Ne	20,8431	0,00034	Ne
	3	21,8031	0,00007	Ne	20,6977	0,00012	Ne
	4	1,0803	0,58266	Ano	1,0516	0,59108	Ano
$S(\mathbf{p}, \mathbf{p}_0)$	0	24,4000	0,00044	Ne	24,0185	0,00052	Ne
	1	21,9841	0,00053	Ne	20,9087	0,00084	Ne
	2	21,8884	0,00021	Ne	20,8450	0,00034	Ne
	3	21,8959	0,00007	Ne	20,7791	0,00012	Ne
	4	0,6264	0,73111	Ano	0,6138	0,73571	Ano
$P(\mathbf{p}, \mathbf{p}_0)$	0	24,4000	0,00044	Ne	24,0185	0,00052	Ne
	1	22,1566	0,00049	Ne	21,0824	0,00078	Ne
	2	21,8447	0,00022	Ne	20,7839	0,00035	Ne
	3	21,5744	0,00008	Ne	20,4961	0,00013	Ne
	4	2,5037	0,28598	Ano	2,3902	0,30268	Ano



Obrázek 1: Grafické znázornění

7.2 Příklad 2

Počítačovou simulací diskrétní náhodné veličiny X s Poissonovým rozdělením pravděpodobnosti s parametrem $\lambda = 1,5$ jsme získali statistický soubor pozorovaných hodnot x_i , $i = 1, \dots, 100$ [35]. Po jeho roztrídění a sloučení původních tří tříd s malými četnostmi pro $x_j^* = 4, 5, 6$ dostaneme roztríděný statistický soubor, který je uveden v tabulce 9.

Tabulka 9: Zadání

j	1	2	3	4	5
x_j^*	0	1	2	3	5
f_j	21	36	27	9	7

Počet tříd $m = 5$ a rozsah $n = 100$, takže

$$M_0 = 1,$$

$$M_1 = \bar{x} = \frac{1}{100} \sum_{j=1}^5 f_j x_j^* = 1,52,$$

$$M_2 = \bar{x} = \frac{1}{100} \sum_{j=1}^5 f_j x_j^{*2} = 4.$$

Pomocí optimalizačního nástroje Řešitel z Excelu pro určení minima Pearsonovy kvazinormy a následujícím chí-kvadrát testem jsme získali výsledky v tabulkách 10 a 11.

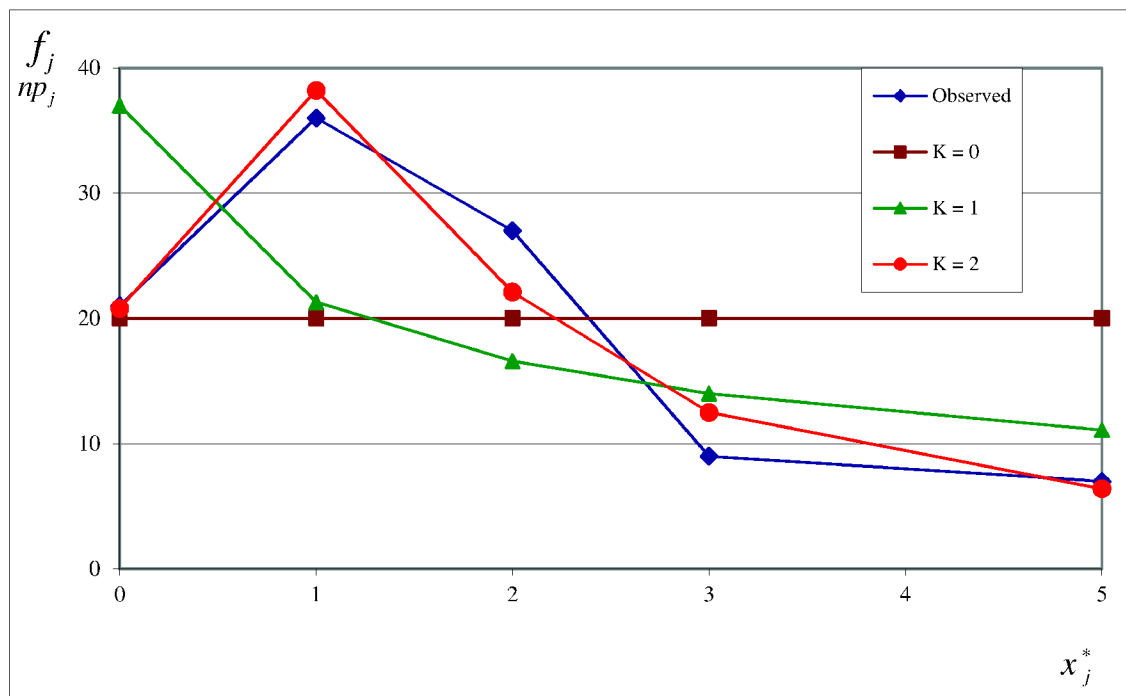
Tabulka 10: Výsledky - Lagrangeovy multiplikátory

K	λ_k	p_j
0	$\lambda_0 = 1$	0,2
1	$\lambda_0 = 0,2918$ $\lambda_1 = 0,5862$	$p_j = \frac{1}{5(\sqrt{\lambda_0 + \lambda_1 x_j^*})}$
2	$\lambda_0 = 0,9217$ $\lambda_1 = -1,244$ $\lambda_2 = 0,5971$	$p_j = \frac{1}{5(\sqrt{\lambda_0 + \lambda_1 x_j^* + \lambda_2 x_j^{*2}})}$

Tabulka 11: Výsledky odhadů

K	\hat{f}_j	21	36	27	9	7	P_K	χ^2	$\chi_{0,95}^2$	Hyp.
0	np_j	20	20	20	20	20	0	29,80	9,488	Ne
1	np_j	37,0	21,3	16,6	14,0	11,1	0,18	26,94	7,815	Ne
2	np_j	20,8	38,2	22,1	12,5	6,4	0,42	29,80	5,991	Ano

Rozdělení pravděpodobnosti získaná pro $K = 0$ a $K = 1$ a vedlejších momentových podmínek na hladině významnosti 0,05 zamítáme. Dokládají to také odhady četností np_j v porovnání s pozorovanými četnostmi v tabulce 11 i znázornění výsledků na obrázku 2.



Obrázek 2: Grafické znázornění

8 Aproximace základních rozdělení pravděpodobnosti pomocí kvazinorem

V této kapitole aplikujeme popsany způsob odhadu rozdělení pomocí minimalizace kvaziny normy na významných rozdělení pravděpodobnosti, jak diskretních, tak diskretizovaných spojitých [45]. Konkrétně se bude jednat o binomické, Poissonovo, normální a Weibullovo. Cílem kapitoly bude posoudit, jak dobře konkrétní kvazinorma s přibývajícím počtem momentových omezení aproximuje vybrané rozdělení. Budeme srovnávat tři nej-používanější kvaziny normy - Hellingerovu, Pearsonovu a Shannonovu.

Pro každý typ rozdělení bude vybráno několik zástupců s konkrétními hodnotami parametrů. Pro tyto zástupce pak bude uměle vytvořen datový soubor o rozsahu $n = 100$. Nebude zde pracováno se simulovanými datovými soubory s využitím generátoru náhodných čísel, ale četnosti jednotlivých tříd budou stanoveny výpočtem podle vztahu $f_i = nP_i$, kde P_i jsou vypočtené pravděpodobnosti jednotlivých tříd pomocí pravděpodobnostní či distribuční funkce. Můžeme hovořit tedy v jistém smyslu o teoretických četnostech. Pro srovnatelnost jednotlivých datových souborů budeme volit rozsah souboru vždy $n = 100$ a počet tříd $m = 9$. Tyto teoretické četnosti budeme pak pokládat za naše pozorování a budeme provádět odhad tohoto rozdělení minimalizací tří kvazinorem. Budeme sledovat, jak se tyto odhady blíží tomuto "teoretickému pozorování" s přibývajícím počtem momentových omezení $K = 0 \dots 8$. Příklad $K = 8$ nebude až tak zajímavý, protože dojde k úplné shodě pro všechny kvaziny normy. Míru shody pozorování a odhadu budeme posuzovat testovým kritériem chí-kvadrát $\chi^2 = \frac{(f_i - np_i)^2}{np_i}$, kde p_i jsou odhady pravděpodobností jednotlivých tříd a budeme také testovat shodu na hladině významnosti 0,95. Ke stanovení hodnot pravděpodobnostní funkce odhadovaného rozdělení, minimalizací příslušné kvaziny normy byl použit speciální software Estimator 1.4.1 [52] na bázi optimalizačního řešitele GAMS, který byl vytvořen pro tyto účely. Pro každé pozorování na závěr vytvoříme graf závislosti chí-kvadrát kritéria na počtu momentových podmínek, ze kterého bude vidět, pro kterou kvazinormu dochází k nejlepší shodě pozorování a odhadu.

8.1 Binomické rozdělení

Jako zástupce binomického rozdělení jsme zvolili rozdělení $Bi(8; 0,5)$ a $Bi(8; 0,25)$. První parametr je volen záměrně, aby vzniklo 9 přirozených tříd pro $x_i = 0, \dots, 8$. Volbou druhého parametru docílíme v prvním případě symetrie, ve druhém levostranné šikmosti. Hodnoty teoretických četností f_i jsou vypočteny přímo z pravděpodobnostní funkce P_i binomického rozdělení a odhadované pravděpodobnosti jsou vypočteny Estimatorem pro počty momentových podmínek $K = 0 \dots 8$. Podrobné výsledky výpočtů jsou v tabulkách, uvedených z důvodu převedení z Excelu jako obrázky 3 a 4. V posledních dvou řádcích tabulky je testová statistika chí-kvadrát testu a příslušná kritická hodnota. Testové kritérium je vyznačeno barevně v případě, že chí-kvadrát test zamítá shodu pozorování a odhadu.

V grafech na obrázcích 5 a 6 je znázorněna závislost testového kritéria χ^2 na počtu

Bi(8;0,5)		Hellinger									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	0,39	0,0039	0,1111	0,1111	0,0125	0,0125	0,0051	0,0051	0,0040	0,0040	0,0039
1	3,12	0,0312	0,1111	0,1111	0,0299	0,0299	0,0267	0,0267	0,0307	0,0307	0,0312
2	10,94	0,1094	0,1111	0,1111	0,0801	0,0801	0,1127	0,1127	0,1111	0,1111	0,1094
3	21,88	0,2188	0,1111	0,1111	0,2104	0,2104	0,2270	0,2270	0,2154	0,2154	0,2188
4	27,34	0,2734	0,1111	0,1111	0,3343	0,3343	0,2569	0,2569	0,2777	0,2777	0,2734
5	21,88	0,2188	0,1111	0,1111	0,2104	0,2104	0,2270	0,2270	0,2154	0,2154	0,2188
6	10,94	0,1094	0,1111	0,1111	0,0801	0,0801	0,1127	0,1127	0,1111	0,1111	0,1094
7	3,12	0,0312	0,1111	0,1111	0,0299	0,0299	0,0267	0,0267	0,0307	0,0307	0,0312
8	0,39	0,0039	0,1111	0,1111	0,0125	0,0125	0,0051	0,0051	0,0040	0,0040	0,0039
		Krit.:	76,7676	76,7676	4,5144	4,5144	0,39058	0,39058	0,02433	0,02433	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Bi(8;0,5)		Pearson									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	0,39	0,0039	0,1111	0,1111	0,0250	0,0250	0,0090	0,0090	0,0057	0,0057	0,0039
1	3,12	0,0312	0,1111	0,1111	0,0334	0,0334	0,0184	0,0184	0,0169	0,0169	0,0312
2	10,94	0,1094	0,1111	0,1111	0,0499	0,0499	0,0835	0,0835	0,1596	0,1596	0,1094
3	21,88	0,2188	0,1111	0,1111	0,0988	0,0988	0,3554	0,3554	0,1184	0,1184	0,2188
4	27,34	0,2734	0,1111	0,1111	0,5856	0,5856	0,0674	0,0674	0,3989	0,3989	0,2734
5	21,88	0,2188	0,1111	0,1111	0,0988	0,0988	0,3554	0,3554	0,1184	0,1184	0,2188
6	10,94	0,1094	0,1111	0,1111	0,0499	0,0499	0,0835	0,0835	0,1596	0,1596	0,1094
7	3,12	0,0312	0,1111	0,1111	0,0334	0,0334	0,0184	0,0184	0,0169	0,0169	0,0312
8	0,39	0,0039	0,1111	0,1111	0,0250	0,0250	0,0090	0,0090	0,0057	0,0057	0,0039
		Krit.:	76,7676	76,7676	63,5283	63,5283	77,4377	77,4377	26,7123	26,7123	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Bi(8;0,5)		Shannon									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	0,39	0,0039	0,1111	0,1111	0,0054	0,0054	0,0040	0,0040	0,0039	0,0039	0,0039
1	3,12	0,0312	0,1111	0,1111	0,0305	0,0305	0,0308	0,0308	0,0312	0,0312	0,0312
2	10,94	0,1094	0,1111	0,1111	0,1047	0,1047	0,1099	0,1099	0,1094	0,1094	0,1094
3	21,88	0,2188	0,1111	0,1111	0,2192	0,2192	0,2191	0,2191	0,2187	0,2187	0,2188
4	27,34	0,2734	0,1111	0,1111	0,2804	0,2804	0,2725	0,2725	0,2735	0,2735	0,2734
5	21,88	0,2188	0,1111	0,1111	0,2192	0,2192	0,2191	0,2191	0,2187	0,2187	0,2188
6	10,94	0,1094	0,1111	0,1111	0,1047	0,1047	0,1099	0,1099	0,1094	0,1094	0,1094
7	3,12	0,0312	0,1111	0,1111	0,0305	0,0305	0,0308	0,0308	0,0312	0,0312	0,0312
8	0,39	0,0039	0,1111	0,1111	0,0054	0,0054	0,0040	0,0040	0,0039	0,0039	0,0039
		Krit.:	76,7676	76,7676	0,15089	0,15089	0,00259	0,00259	1,9E-05	1,9E-05	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 3: Výsledky pro binomické rozdělení Bi(8; 0,5)

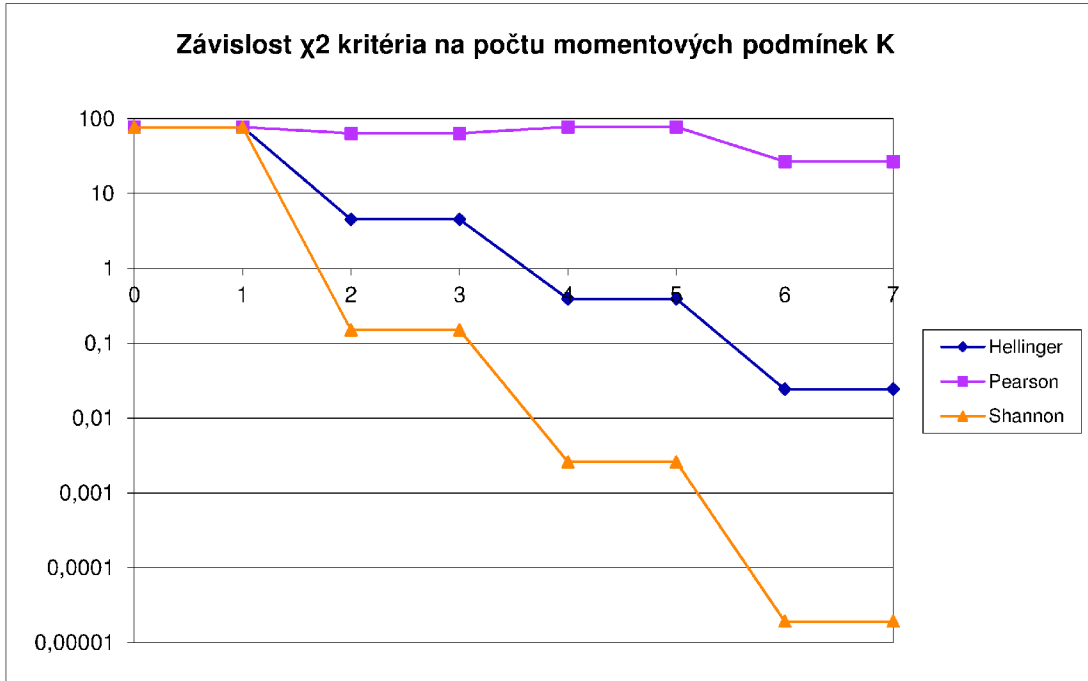
Bi(8;0,25)		Hellinger									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	10,01	0,1001	0,1111	0,3479	0,0853	0,1208	0,0965	0,1010	0,1000	0,1001	0,1001
1	26,70	0,2670	0,1111	0,2018	0,2554	0,2108	0,2816	0,2618	0,2679	0,2669	0,2670
2	31,15	0,3115	0,1111	0,1316	0,3829	0,3428	0,2914	0,3230	0,3090	0,3119	0,3115
3	20,76	0,2076	0,1111	0,0926	0,1834	0,2339	0,2150	0,1950	0,2114	0,2068	0,2076
4	8,65	0,0865	0,1111	0,0686	0,0584	0,0672	0,0922	0,0927	0,0831	0,0875	0,0865
5	2,31	0,0231	0,1111	0,0529	0,0203	0,0170	0,0192	0,0228	0,0247	0,0223	0,0231
6	0,38	0,0038	0,1111	0,0420	0,0083	0,0050	0,0033	0,0031	0,0036	0,0042	0,0038
7	0,04	0,0004	0,1111	0,0342	0,0039	0,0018	0,0007	0,0004	0,0003	0,0003	0,0004
8	0,00	0,0000	0,1111	0,0283	0,0020	0,0007	0,0002	0,0001	0,0000	0,0000	0,0000
		Krit.:	106,523	70,4299	4,10619	3,4135	0,40228	0,19821	0,04102	0,01415	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Bi(8;0,25)		Pearson									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	10,01	0,1001	0,1111	0,4670	0,0407	0,1820	0,0599	0,1097	0,0992	0,1001	0,1001
1	26,70	0,2670	0,1111	0,1197	0,2249	0,0348	0,4444	0,2125	0,2729	0,2667	0,2670
2	31,15	0,3115	0,1111	0,0860	0,6168	0,4699	0,0201	0,4372	0,2945	0,3125	0,3115
3	20,76	0,2076	0,1111	0,0706	0,0425	0,2736	0,4089	0,0599	0,2346	0,2055	0,2076
4	8,65	0,0865	0,1111	0,0614	0,0244	0,0174	0,0546	0,1745	0,0614	0,0891	0,0865
5	2,31	0,0231	0,1111	0,0550	0,0173	0,0090	0,0061	0,0035	0,0362	0,0210	0,0231
6	0,38	0,0038	0,1111	0,0502	0,0133	0,0058	0,0029	0,0014	0,0007	0,0048	0,0038
7	0,04	0,0004	0,1111	0,0466	0,0109	0,0042	0,0018	0,0008	0,0003	0,0001	0,0004
8	0,00	0,0000	0,1111	0,0436	0,0092	0,0032	0,0012	0,0005	0,0001	0,0000	0,0000
		Krit.:	106,523	148,709	107,289	196,015	447,781	57,2585	3,20937	0,13982	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Bi(8;0,25)		Shannon									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	10,01	0,1001	0,1111	0,3062	0,1118	0,1034	0,1004	0,1002	0,1001	0,1001	0,1001
1	26,70	0,2670	0,1111	0,2166	0,2468	0,2580	0,2657	0,2666	0,2671	0,2670	0,2670
2	31,15	0,3115	0,1111	0,1532	0,3094	0,3170	0,3134	0,3123	0,3113	0,3116	0,3115
3	20,76	0,2076	0,1111	0,1084	0,2200	0,2111	0,2066	0,2069	0,2079	0,2074	0,2076
4	8,65	0,0865	0,1111	0,0767	0,0888	0,0839	0,0863	0,0867	0,0864	0,0868	0,0865
5	2,31	0,0231	0,1111	0,0542	0,0203	0,0219	0,0234	0,0232	0,0231	0,0229	0,0231
6	0,38	0,0038	0,1111	0,0384	0,0026	0,0041	0,0038	0,0038	0,0039	0,0039	0,0038
7	0,04	0,0004	0,1111	0,0271	0,0002	0,0006	0,0003	0,0004	0,0003	0,0004	0,0004
8	0,00	0,0000	0,1111	0,0192	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000
		Krit.:	106,523	50,0595	0,47448	0,0906	0,00604	0,00299	0,00198	0,0015	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 4: Výsledky pro binomické rozdělení Bi(8; 0,25)

momentových podmínek K . Z obrázků je názorně vidět srovnání našich tří kvazinorem z hlediska kvality aproximace.

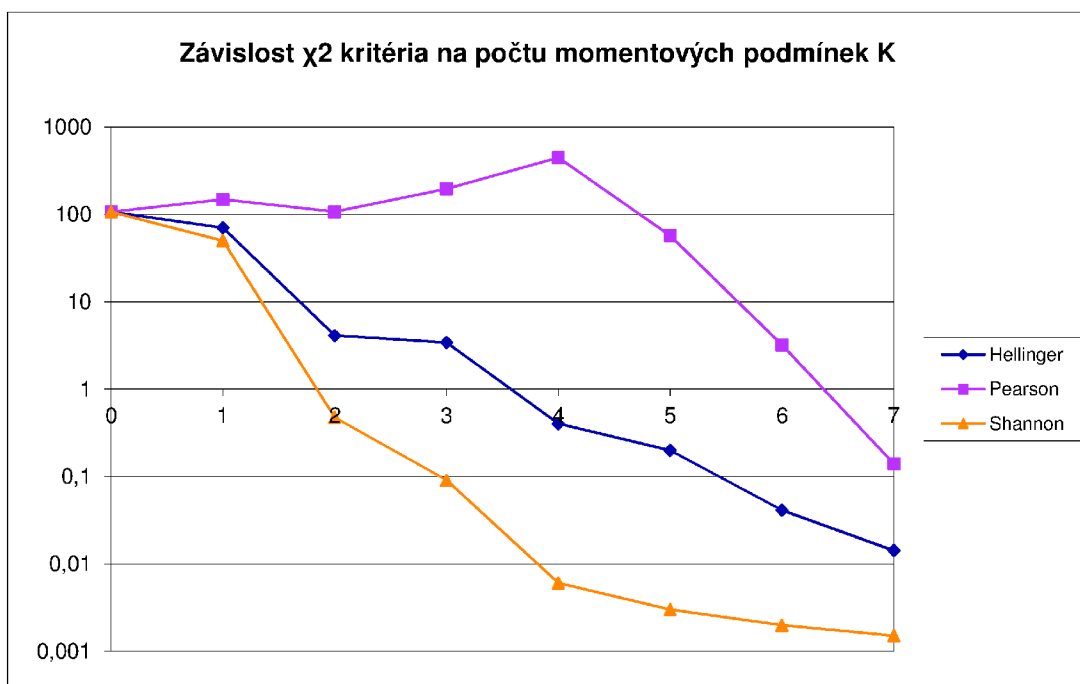


Obrázek 5: Graf závislosti χ^2 na K pro rozdělení $\text{Bi}(8; 0,5)$

Z výsledků je vidět, že pro aproximování binomického rozdělení se nejvíce hodí Shannonova kvazinorma, o něco hůře Hellingerova a daleko nejhůře Pearsonova. U prvních dvou je dosaženo dobré shody na hladině významnosti 0,95 už při $K = 2$. Pro Pearsonovu kvazinormu je dosaženo dobré shody při vyšším počtu momentových omezení. Tato kvazinorma se chová "zvláště" tím, že přidáním další omezující podmínky je někdy dosaženo horší shody s pozorováním. Znamená to, že přidáním další informace o pozorovaném rozdělení v podobě předepsání další momentové podmínky pro odhad, se tento odhad vzdálí od pozorování ve smyslu chí-kvadrát kritéria. Tento fakt se může zdát překvapivý, nicméně není v rozporu s naším přístupem k odhadu rozdělení.

Hellingerova a Shannonova kvazinorma se takto nechovají, tam se zvýšením K nedochází k poklesu χ^2 . U rozdělení $\text{Bi}(8; 0,5)$ pouze přidání liché momentové podmínky nezpůsobí změnu dohadu a tudíž ani χ^2 . Tato skutečnost je způsobena symetrií rozdělení. U nesymetrického $\text{Bi}(8; 0,25)$ dochází k poklesu χ^2 při každém vzrůstu K .

Co se týče srovnání odhadů našich dvou zástupců binomického rozdělení, dá se říct, že pro všechny kvazinormy jsou aproximace vesměs o něco lepší pro rozdělení $\text{Bi}(8; 0,5)$. Tato vlastnost je zřejmě způsobena menšími odlišnostmi v teoretických četnostech jednotlivých tříd pro rozdělení $\text{Bi}(8; 0,5)$, tudíž vyšší neurčitostí pozorování, tj. nižší odlišností samotného pozorování od maximálně entropického rozdělení \mathbf{p}_0 . Tato relativně malá odlišnost v odhadech těchto dvou rozdělení se ještě více zmenšuje s rostoucím K , což svědčí o dobré flexibilitě odhadu rozdělení s využitím minimalizace kvazinorem.



Obrázek 6: Graf závislosti χ^2 na K pro rozdělení $\text{Bi}(8; 0,25)$

8.2 Poissonovo rozdělení

Jako zástupce Poissonova rozdělení jsme zvolili dvě rozdělení $\text{Po}(2)$ a $\text{Po}(3)$. Hodnoty teoretických četností f_i jsou vypočteny přímo z pravděpodobnostní funkce P_i Poissonova rozdělení. Abychom dostali devět tříd, stejně jako u binomického rozdělení, zvolíme osm přirozených tříd $x = 0, \dots, 7$ a poslední třída, kterou budeme reprezentovat hodnotou $x = 8$ vznikne sloučením tříd pro $x \geq 8$. Odhadované pravděpodobnosti jsou vypočteny Estimátorem pro počty momentových podmínek $K = 0 \dots 8$. Podrobné výsledky výpočtů jsou v tabulkách, uvedených z důvodu převedení z Excelu jako obrázky 7 a 8. V posledních dvou řádcích tabulky je testová statistika chí-kvadrát testu a příslušná kritická hodnota. Testové kritérium je vyznačeno barevně v případě, že chí-kvadrát test zamítá shodu pozorování a odhadu.

V grafech na obrázcích 9 a 10 je znázorněna závislost testového kritéria χ^2 na počtu momentových podmínek K . Z obrázků je názorně vidět srovnání našich tří kvazinorem z hlediska kvality aproximace.

Výsledky jsou velmi podobné jako u binomického rozdělení. Hellingerova a Shannonova kvazinorma dopadli dobře. K dobré shodě na hladině významnosti 0,95 dochází již při $k = 2$. Pearsonova kvazinorma dopadá hůře, obzvláště pro rozdělení $\text{Po}(2)$, kde se podobně jako u binomického rozdělení někdy objeví pokles χ^2 s rostoucím K . Důvody byly nastíněny v podkapitole o binomickém rozdělení a souvisí zřejmě s velkou nevyrovnaností teoretických četností.

Po(2)		Hellinger									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	13,53	0,1353	0,1111	0,3480	0,1242	0,1495	0,1349	0,1354	0,1353	0,1353	0,1353
1	27,07	0,2707	0,1111	0,2018	0,2655	0,2343	0,2719	0,2701	0,2704	0,2707	0,2707
2	27,07	0,2707	0,1111	0,1316	0,3113	0,2875	0,2698	0,2721	0,2716	0,2706	0,2707
3	18,05	0,1805	0,1111	0,0925	0,1747	0,1983	0,1795	0,1789	0,1793	0,1808	0,1805
4	9,02	0,0902	0,1111	0,0686	0,0714	0,0830	0,0918	0,0911	0,0910	0,0899	0,0902
5	3,61	0,0361	0,1111	0,0529	0,0293	0,0297	0,0358	0,0361	0,0360	0,0364	0,0361
6	1,20	0,0120	0,1111	0,0420	0,0133	0,0111	0,0115	0,0117	0,0118	0,0119	0,0120
7	0,34	0,0034	0,1111	0,0342	0,0066	0,0045	0,0035	0,0035	0,0035	0,0034	0,0034
8	0,11	0,0011	0,1111	0,0283	0,0036	0,0021	0,0011	0,0011	0,0011	0,0011	0,0011
		Krit.:	86,3354	47,1513	1,65319	1,24012	0,00697	0,00431	0,00284	0,00052	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Po(2)		Pearson									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	13,53	0,1353	0,1111	0,4671	0,0544	0,2134	0,1004	0,1455	0,1338	0,1353	0,1353
1	27,07	0,2707	0,1111	0,1196	0,2984	0,0606	0,4174	0,2142	0,2808	0,2703	0,2707
2	27,07	0,2707	0,1111	0,0860	0,4912	0,3892	0,0508	0,3972	0,2417	0,2720	0,2707
3	18,05	0,1805	0,1111	0,0706	0,0562	0,2674	0,3000	0,0395	0,2262	0,1779	0,1805
4	9,02	0,0902	0,1111	0,0613	0,0324	0,0304	0,1016	0,1651	0,0486	0,0935	0,0902
5	3,61	0,0361	0,1111	0,0550	0,0229	0,0158	0,0152	0,0266	0,0570	0,0335	0,0361
6	1,20	0,0120	0,1111	0,0502	0,0177	0,0102	0,0073	0,0069	0,0078	0,0133	0,0120
7	0,34	0,0034	0,1111	0,0465	0,0145	0,0073	0,0044	0,0033	0,0028	0,0030	0,0034
8	0,11	0,0011	0,1111	0,0436	0,0122	0,0056	0,0029	0,0019	0,0014	0,0011	0,0011
		Krit.:	86,3354	112,438	62,8061	97,0787	109,694	60,1205	5,8759	0,05465	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Po(2)		Shannon									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	13,53	0,1353	0,1111	0,3062	0,1564	0,1396	0,1362	0,1353	0,1353	0,1353	0,1353
1	27,07	0,2707	0,1111	0,2166	0,2412	0,2599	0,2674	0,2705	0,2706	0,2707	0,2707
2	27,07	0,2707	0,1111	0,1532	0,2591	0,2754	0,2742	0,2711	0,2710	0,2707	0,2707
3	18,05	0,1805	0,1111	0,1084	0,1940	0,1856	0,1805	0,1800	0,1801	0,1805	0,1805
4	9,02	0,0902	0,1111	0,0767	0,1012	0,0889	0,0887	0,0904	0,0905	0,0902	0,0902
5	3,61	0,0361	0,1111	0,0542	0,0368	0,0339	0,0359	0,0361	0,0361	0,0361	0,0361
6	1,20	0,0120	0,1111	0,0384	0,0093	0,0114	0,0125	0,0119	0,0119	0,0120	0,0120
7	0,34	0,0034	0,1111	0,0271	0,0016	0,0038	0,0037	0,0035	0,0035	0,0034	0,0034
8	0,11	0,0011	0,1111	0,0192	0,0002	0,0014	0,0009	0,0011	0,0011	0,0011	0,0011
		Krit.:	86,3354	31,1364	1,57514	0,11216	0,02205	0,00055	0,00054	1E-05	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

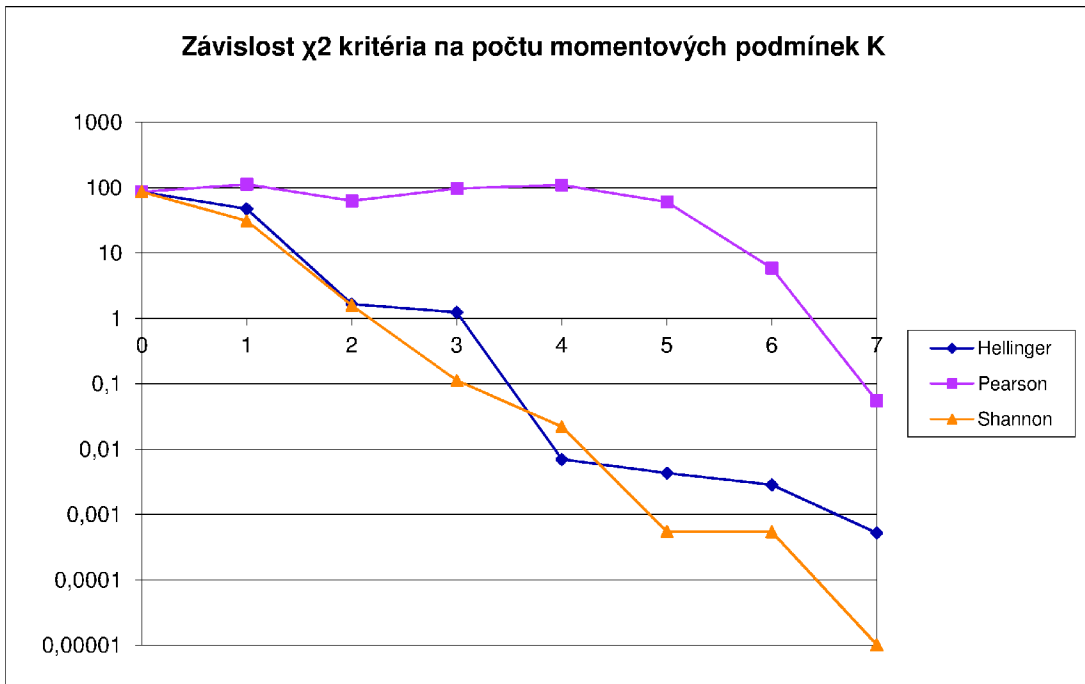
Obrázek 7: Výsledky pro Poissonovo rozdělení Po(2)

Po(3)		Hellinger									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	4,98	0,0498	0,1111	0,1995	0,0636	0,0608	0,0530	0,0499	0,0500	0,0498	0,0498
1	14,94	0,1494	0,1111	0,1634	0,1270	0,1274	0,1390	0,1488	0,1483	0,1495	0,1494
2	22,40	0,2240	0,1111	0,1363	0,2147	0,2202	0,2319	0,2260	0,2267	0,2237	0,2240
3	22,40	0,2240	0,1111	0,1154	0,2451	0,2467	0,2290	0,2210	0,2207	0,2246	0,2240
4	16,81	0,1681	0,1111	0,0990	0,1748	0,1710	0,1626	0,1700	0,1696	0,1674	0,1681
5	10,08	0,1008	0,1111	0,0858	0,0933	0,0909	0,0977	0,1013	0,1017	0,1014	0,1008
6	5,04	0,0504	0,1111	0,0751	0,0459	0,0457	0,0518	0,0488	0,0489	0,0501	0,0504
7	2,16	0,0216	0,1111	0,0663	0,0232	0,0239	0,0244	0,0225	0,0223	0,0217	0,0216
8	0,81	0,0119	0,1111	0,0590	0,0124	0,0133	0,0105	0,0117	0,0118	0,0119	0,0119
		Krit.:	50,0471	39,8803	1,06008	0,9927	0,21803	0,01702	0,01858	0,00108	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

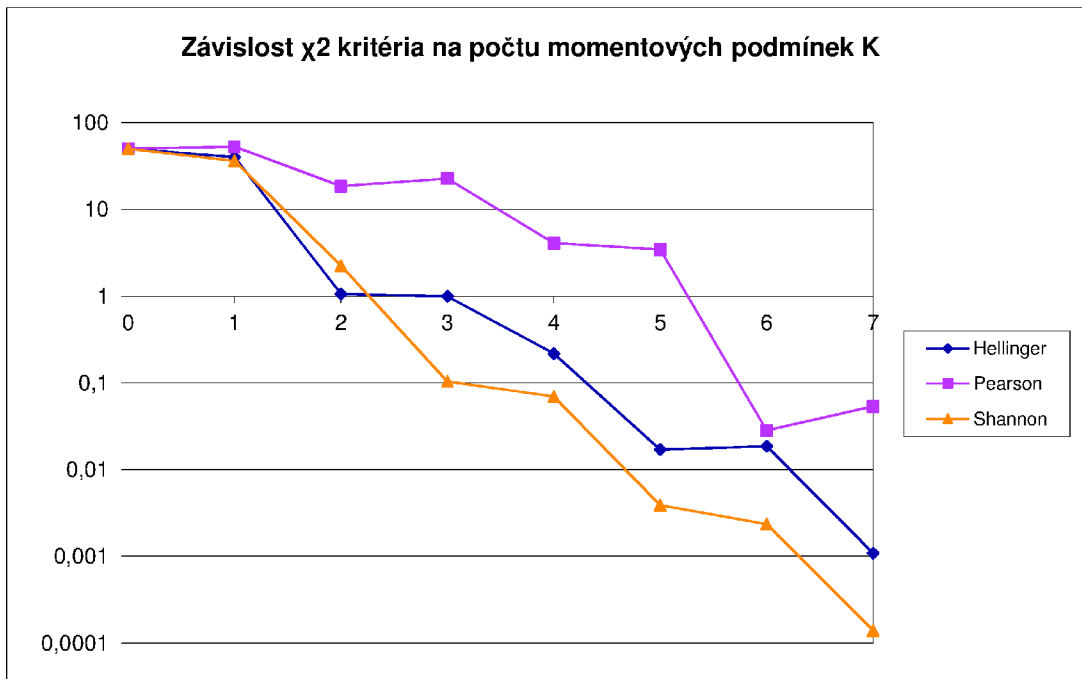
Po(3)		Pearson									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	4,98	0,0498	0,1111	0,2324	0,0543	0,0900	0,0474	0,0549	0,0500	0,0497	0,0498
1	14,94	0,1494	0,1111	0,1516	0,0891	0,0912	0,1490	0,1227	0,1481	0,1500	0,1494
2	22,40	0,2240	0,1111	0,1208	0,2537	0,1339	0,2545	0,2784	0,2272	0,2219	0,2240
3	22,40	0,2240	0,1111	0,1034	0,3496	0,3963	0,1604	0,1743	0,2202	0,2282	0,2240
4	16,81	0,1681	0,1111	0,0918	0,0981	0,1436	0,2119	0,1799	0,1696	0,1629	0,1681
5	10,08	0,1008	0,1111	0,0835	0,0575	0,0622	0,1017	0,1128	0,1022	0,1050	0,1008
6	5,04	0,0504	0,1111	0,0770	0,0407	0,0375	0,0402	0,0431	0,0484	0,0483	0,0504
7	2,16	0,0216	0,1111	0,0719	0,0315	0,0259	0,0215	0,0214	0,0226	0,0222	0,0216
8	0,81	0,0119	0,1111	0,0676	0,0257	0,0194	0,0134	0,0126	0,0117	0,0118	0,0119
		Krit.:	50,0471	52,9496	18,5187	22,68	4,07957	3,43818	0,02819	0,05352	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Po(3)		Shannon									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
0	4,98	0,0498	0,1111	0,1913	0,0701	0,0531	0,0513	0,0497	0,0499	0,0498	0,0498
1	14,94	0,1494	0,1111	0,1637	0,1361	0,1422	0,1448	0,1496	0,1491	0,1494	0,1494
2	22,40	0,2240	0,1111	0,1401	0,1990	0,2244	0,2264	0,2239	0,2248	0,2239	0,2240
3	22,40	0,2240	0,1111	0,1199	0,2195	0,2301	0,2274	0,2234	0,2231	0,2242	0,2240
4	16,81	0,1681	0,1111	0,1026	0,1825	0,1687	0,1664	0,1689	0,1683	0,1678	0,1681
5	10,08	0,1008	0,1111	0,0878	0,1144	0,0975	0,0982	0,1010	0,1014	0,1010	0,1008
6	5,04	0,0504	0,1111	0,0752	0,0540	0,0488	0,0505	0,0495	0,0497	0,0503	0,0504
7	2,16	0,0216	0,1111	0,0643	0,0192	0,0234	0,0240	0,0222	0,0219	0,0216	0,0216
8	0,81	0,0119	0,1111	0,0550	0,0052	0,0118	0,0109	0,0118	0,0119	0,0119	0,0119
		Krit.:	50,0471	36,0561	2,2462	0,10398	0,06949	0,00387	0,00235	0,00014	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 8: Výsledky pro Poissonovo rozdělení Po(3)



Obrázek 9: Graf závislosti χ^2 na K pro rozdělení Po(2)



Obrázek 10: Graf závislosti χ^2 na K pro rozdělení Po(3)

8.3 Normální rozdělení

Pro odhad normálního rozdělení jsme zvolili tři zástupce se střední hodnotou 0 a směrodatnými odchylkami 3, 2 a 1,5. Pro všechna tři rozdělení byla zvolena stejná diskretizace s šířkou třídy rovnou 1. Ve výsledkových tabulkách jsou třídy reprezentovány celými čísly od -4 do 4, které mají význam středů tříd, s výjimkou krajních dvou reprezentujících rozšířené třídy $(-\infty; -3,5)$ a $(3,5; \infty)$. Teoretické četnosti byly vypočteny z distribuční funkce normálního rozdělení. Z hlediska zásad třídění se toto třídění nejlépe hodí pro rozdělení $N(0; 4)$. Pokud bychom však zvolili třídění úměrné směrodatné odchylce rozdělení, vyšly by teoretické četnosti tříd stejně a naše srovnání by postrádalo smysl.

Odhadované pravděpodobnosti jsou vypočteny Estimátorem pro počty momentových podmínek $K = 0 \dots 8$, stejně jako předchozích podkapitolách. Podrobné výsledky výpočtů jsou v tabulkách, uvedených z důvodu převedení z Excelu jako obrázky 11, 12 a 13. V posledních dvou řádcích tabulky je testová statistika chí-kvadrát testu a příslušná kritická hodnota. Testové kritérium je vyznačeno barevně v případě, že chí-kvadrát test zamítá shodu pozorování a odhadu.

V grafech na obrázcích 15, 14 a 16 je znázorněna závislost testového kritéria χ^2 na počtu momentových podmínek K . Z obrázků je názorně vidět srovnání našich tří kvazinorem z hlediska kvality aproximace.

Pro aproximaci normálního rozdělení je dosaženo velmi uspokojivých výsledků. Pro rozdělení s největším rozptylem $N(0; 9)$ dosaženo shody s pozorováním již při samotném rozdělení \mathbf{p}_0 ($K = 0$). Je to však způsobeno malými rozdíly mezi četnostmi jednotlivých tříd. Použité roztrídění datového souboru z tohoto rozdělení je však nevhodné, proto se jím už více nebudeme zabývat.

Pro další dvě rozdělení, a to $N(0; 4)$ a $N(0; 2,25)$, je dosaženo dobré shody již pro $K = 2$ pro všechny kvazinormy s výjimkou Pearsonovy u rozdělení $N(0; 2,25)$. Toto rozdělení s nejmenším rozptylem dosahuje největších rozdílů v četnostech tříd a obdobně jako případech v minulých podkapitolách je Pearsonova kvazinorma nevhodná.

Ostatní dvě kvazinormy dosahují podobných velmi dobrých výsledků a obzvláště pro nejlépe tříděný případ $N(0; 4)$ nelze jednoznačně říct, která je lepší. Přidání liché momentové podmínky nevede ke změně odhadu, což je podobně jako u $Bi(8; 0,5)$ způsobeno symetrií normálního rozdělení.

Další zajímavostí je rychlost poklesu χ^2 při růstu K v závislosti na variabilitě normálního rozdělení. U rozdělení $N(0; 9)$ s největším rozptylem dochází k dobré shodě již pro $K = 0$, avšak hodnota testového kritéria chí-kvadrát testu s rostoucím K klesá nejpomaleji. Pro rozdělení s větším rozptylem, kde jsou větší rozdíly v teoretických četnostech, hodnota χ^2 začíná na logicky na vyšší hodnotě, avšak s rostoucím K je pokles χ^2 razantnější a od jisté hodnoty K je dosaženo lepší shody než pro více variabilní soubory. To opět svědčí o dobré flexibilitě našich odhadů. Tato závislost je nejmarkantnější pro Shannonovu kvazinormu (viz obrázek 17).

N(0;9)		Hellinger									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	12,17	0,1217	0,1111	0,1111	0,1045	0,1045	0,1203	0,1203	0,1217	0,1217	0,1217
-3	8,07	0,0806	0,1111	0,1111	0,1093	0,1093	0,0867	0,0867	0,0808	0,0808	0,0806
-2	10,62	0,1062	0,1111	0,1111	0,1130	0,1130	0,0984	0,0984	0,1057	0,1057	0,1062
-1	12,53	0,1253	0,1111	0,1111	0,1152	0,1152	0,1247	0,1247	0,1264	0,1264	0,1253
0	13,24	0,1324	0,1111	0,1111	0,1160	0,1160	0,1398	0,1398	0,1311	0,1311	0,1324
1	12,53	0,1253	0,1111	0,1111	0,1152	0,1152	0,1247	0,1247	0,1264	0,1264	0,1253
2	10,62	0,1062	0,1111	0,1111	0,1130	0,1130	0,0984	0,0984	0,1057	0,1057	0,1062
3	8,07	0,0806	0,1111	0,1111	0,1093	0,1093	0,0867	0,0867	0,0808	0,0808	0,0806
4	12,17	0,1217	0,1111	0,1111	0,1045	0,1045	0,1203	0,1203	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,56312	2,56312	0,25103	0,25103	0,00378	0,00378	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

N(0;9)		Pearson									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	12,17	0,1217	0,1111	0,1111	0,1046	0,1046	0,1199	0,1199	0,1217	0,1217	0,1217
-3	8,07	0,0806	0,1111	0,1111	0,1092	0,1092	0,0881	0,0881	0,0809	0,0809	0,0806
-2	10,62	0,1062	0,1111	0,1111	0,1129	0,1129	0,0971	0,0971	0,1052	0,1052	0,1062
-1	12,53	0,1253	0,1111	0,1111	0,1153	0,1153	0,1230	0,1230	0,1274	0,1274	0,1253
0	13,24	0,1324	0,1111	0,1111	0,1161	0,1161	0,1437	0,1437	0,1298	0,1298	0,1324
1	12,53	0,1253	0,1111	0,1111	0,1153	0,1153	0,1230	0,1230	0,1274	0,1274	0,1253
2	10,62	0,1062	0,1111	0,1111	0,1129	0,1129	0,0971	0,0971	0,1052	0,1052	0,1062
3	8,07	0,0806	0,1111	0,1111	0,1092	0,1092	0,0881	0,0881	0,0809	0,0809	0,0806
4	12,17	0,1217	0,1111	0,1111	0,1046	0,1046	0,1199	0,1199	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,53889	2,53889	0,39914	0,39914	0,01393	0,01393	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

N(0;9)		Shannon									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	12,17	0,1217	0,1111	0,1111	0,1045	0,1045	0,1204	0,1204	0,1217	0,1217	0,1217
-3	8,07	0,0806	0,1111	0,1111	0,1094	0,1094	0,0862	0,0862	0,0807	0,0807	0,0806
-2	10,62	0,1062	0,1111	0,1111	0,1130	0,1130	0,0990	0,0990	0,1058	0,1058	0,1062
-1	12,53	0,1253	0,1111	0,1111	0,1152	0,1152	0,1251	0,1251	0,1262	0,1262	0,1253
0	13,24	0,1324	0,1111	0,1111	0,1160	0,1160	0,1387	0,1387	0,1313	0,1313	0,1324
1	12,53	0,1253	0,1111	0,1111	0,1152	0,1152	0,1251	0,1251	0,1262	0,1262	0,1253
2	10,62	0,1062	0,1111	0,1111	0,1130	0,1130	0,0990	0,0990	0,1058	0,1058	0,1062
3	8,07	0,0806	0,1111	0,1111	0,1094	0,1094	0,0862	0,0862	0,0807	0,0807	0,0806
4	12,17	0,1217	0,1111	0,1111	0,1045	0,1045	0,1204	0,1204	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,57125	2,57125	0,20884	0,20884	0,00247	0,00247	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 11: Výsledky pro normální rozdělení N(0; 9)

N(0;4)		Hellinger										
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)	
-4	4,01	0,0400	0,1111	0,1111	0,0393	0,0393	0,0392	0,0392	0,0400	0,0400	0,0400	
-3	6,56	0,0656	0,1111	0,1111	0,0689	0,0689	0,0690	0,0690	0,0658	0,0658	0,0656	
-2	12,10	0,1210	0,1111	0,1111	0,1167	0,1167	0,1168	0,1168	0,1204	0,1204	0,1210	
-1	17,47	0,1747	0,1111	0,1111	0,1739	0,1739	0,1738	0,1738	0,1758	0,1758	0,1747	
0	19,74	0,1974	0,1111	0,1111	0,2026	0,2026	0,2023	0,2023	0,1960	0,1960	0,1974	
1	17,47	0,1747	0,1111	0,1111	0,1739	0,1739	0,1738	0,1738	0,1758	0,1758	0,1747	
2	12,10	0,1210	0,1111	0,1111	0,1167	0,1167	0,1168	0,1168	0,1204	0,1204	0,1210	
3	6,56	0,0656	0,1111	0,1111	0,0689	0,0689	0,0690	0,0690	0,0658	0,0658	0,0656	
4	4,01	0,0400	0,1111	0,1111	0,0393	0,0393	0,0392	0,0392	0,0400	0,0400	0,0400	
		Krit.:	26,9861	26,9861	0,0802	0,0802	0,07866	0,07866	0,00309	0,00309	0	
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0	

N(0;4)		Pearson										
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)	
-4	4,01	0,0400	0,1111	0,1111	0,0487	0,0487	0,0399	0,0399	0,0399	0,0399	0,0400	
-3	6,56	0,0656	0,1111	0,1111	0,0642	0,0642	0,0665	0,0665	0,0662	0,0662	0,0656	
-2	12,10	0,1210	0,1111	0,1111	0,0930	0,0930	0,1186	0,1186	0,1187	0,1187	0,1210	
-1	17,47	0,1747	0,1111	0,1111	0,1596	0,1596	0,1787	0,1787	0,1792	0,1792	0,1747	
0	19,74	0,1974	0,1111	0,1111	0,2690	0,2690	0,1927	0,1927	0,1917	0,1917	0,1974	
1	17,47	0,1747	0,1111	0,1111	0,1596	0,1596	0,1787	0,1787	0,1792	0,1792	0,1747	
2	12,10	0,1210	0,1111	0,1111	0,0930	0,0930	0,1186	0,1186	0,1187	0,1187	0,1210	
3	6,56	0,0656	0,1111	0,1111	0,0642	0,0642	0,0665	0,0665	0,0662	0,0662	0,0656	
4	4,01	0,0400	0,1111	0,1111	0,0487	0,0487	0,0399	0,0399	0,0399	0,0399	0,0400	
		Krit.:	26,9861	26,9861	4,19495	4,19495	0,04157	0,04157	0,04974	0,04974	0	
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0	

N(0;4)		Shannon										
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)	
-4	4,01	0,0400	0,1111	0,1111	0,0349	0,0349	0,0394	0,0394	0,0400	0,0400	0,0400	
-3	6,56	0,0656	0,1111	0,1111	0,0734	0,0734	0,0681	0,0681	0,0657	0,0657	0,0656	
-2	12,10	0,1210	0,1111	0,1111	0,1248	0,1248	0,1178	0,1178	0,1207	0,1207	0,1210	
-1	17,47	0,1747	0,1111	0,1111	0,1716	0,1716	0,1743	0,1743	0,1752	0,1752	0,1747	
0	19,74	0,1974	0,1111	0,1111	0,1908	0,1908	0,2007	0,2007	0,1967	0,1967	0,1974	
1	17,47	0,1747	0,1111	0,1111	0,1716	0,1716	0,1743	0,1743	0,1752	0,1752	0,1747	
2	12,10	0,1210	0,1111	0,1111	0,1248	0,1248	0,1178	0,1178	0,1207	0,1207	0,1210	
3	6,56	0,0656	0,1111	0,1111	0,0734	0,0734	0,0681	0,0681	0,0657	0,0657	0,0656	
4	4,01	0,0400	0,1111	0,1111	0,0349	0,0349	0,0394	0,0394	0,0400	0,0400	0,0400	
		Krit.:	26,9861	26,9861	0,37128	0,37128	0,0431	0,0431	0,00066	0,00066	0	
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0	

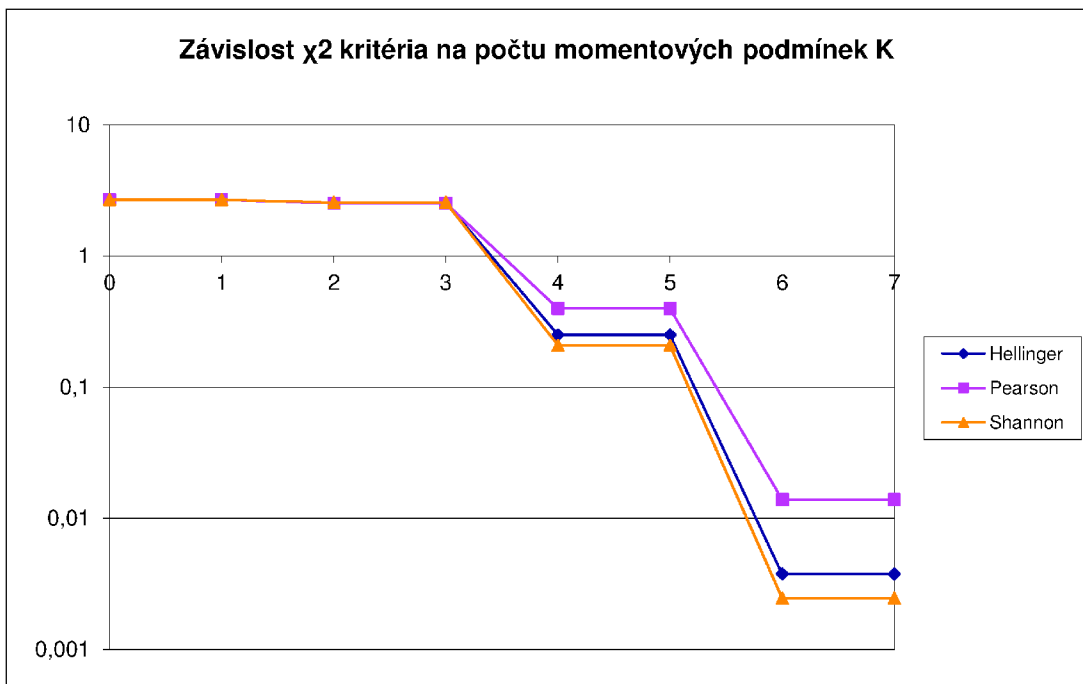
Obrázek 12: Výsledky pro normální rozdělení $N(0; 4)$

N(0;2,25) Hellinger											
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	0,98	0,0098	0,1111	0,1111	0,1045	0,1045	0,1203	0,1203	0,1217	0,1217	0,1217
-3	3,80	0,0380	0,1111	0,1111	0,1093	0,1093	0,0867	0,0867	0,0808	0,0808	0,0806
-2	11,09	0,1109	0,1111	0,1111	0,1130	0,1130	0,0984	0,0984	0,1057	0,1057	0,1062
-1	21,08	0,2108	0,1111	0,1111	0,1152	0,1152	0,1247	0,1247	0,1264	0,1264	0,1253
0	26,11	0,2610	0,1111	0,1111	0,1160	0,1160	0,1398	0,1398	0,1311	0,1311	0,1324
1	21,08	0,2108	0,1111	0,1111	0,1152	0,1152	0,1247	0,1247	0,1264	0,1264	0,1253
2	11,09	0,1109	0,1111	0,1111	0,1130	0,1130	0,0984	0,0984	0,1057	0,1057	0,1062
3	3,80	0,0380	0,1111	0,1111	0,1093	0,1093	0,0867	0,0867	0,0808	0,0808	0,0806
4	0,98	0,0098	0,1111	0,1111	0,1045	0,1045	0,1203	0,1203	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,56312	2,56312	0,25103	0,25103	0,00378	0,00378	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

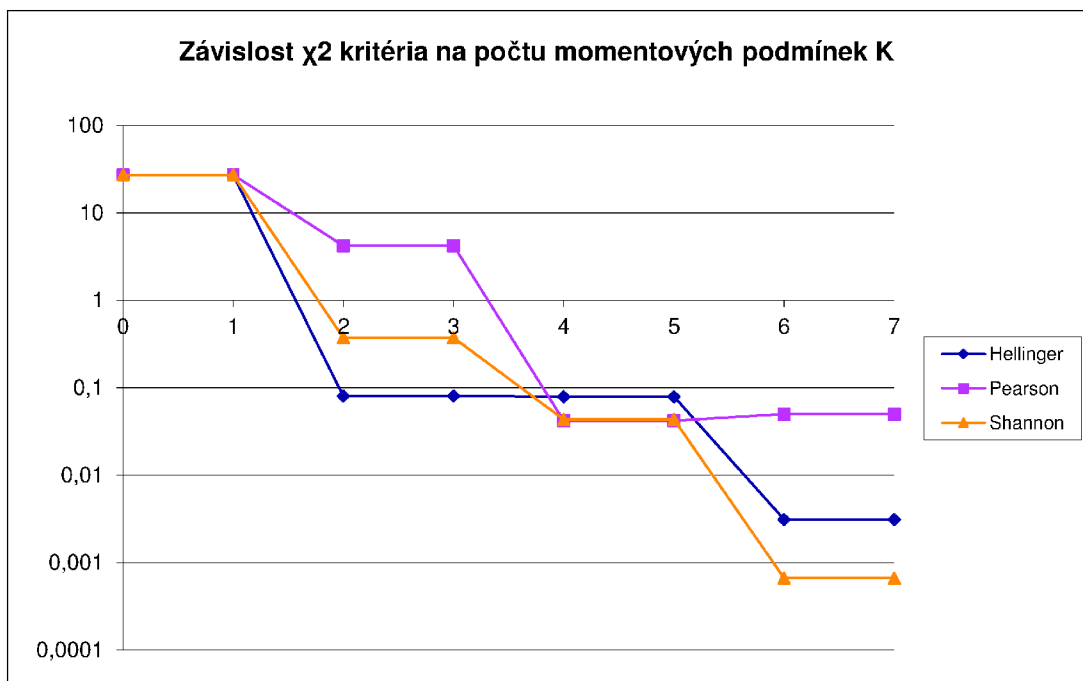
N(0;2,25) Pearson											
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	0,98	0,0098	0,1111	0,1111	0,1046	0,1046	0,1199	0,1199	0,1217	0,1217	0,1217
-3	3,80	0,0380	0,1111	0,1111	0,1092	0,1092	0,0881	0,0881	0,0809	0,0809	0,0806
-2	11,09	0,1109	0,1111	0,1111	0,1129	0,1129	0,0971	0,0971	0,1052	0,1052	0,1062
-1	21,08	0,2108	0,1111	0,1111	0,1153	0,1153	0,1230	0,1230	0,1274	0,1274	0,1253
0	26,11	0,2610	0,1111	0,1111	0,1161	0,1161	0,1437	0,1437	0,1298	0,1298	0,1324
1	21,08	0,2108	0,1111	0,1111	0,1153	0,1153	0,1230	0,1230	0,1274	0,1274	0,1253
2	11,09	0,1109	0,1111	0,1111	0,1129	0,1129	0,0971	0,0971	0,1052	0,1052	0,1062
3	3,80	0,0380	0,1111	0,1111	0,1092	0,1092	0,0881	0,0881	0,0809	0,0809	0,0806
4	0,98	0,0098	0,1111	0,1111	0,1046	0,1046	0,1199	0,1199	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,53889	2,53889	0,39914	0,39914	0,01393	0,01393	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

N(0;2,25) Shannon											
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
-4	0,98	0,0098	0,1111	0,1111	0,1045	0,1045	0,1204	0,1204	0,1217	0,1217	0,1217
-3	3,80	0,0380	0,1111	0,1111	0,1094	0,1094	0,0862	0,0862	0,0807	0,0807	0,0806
-2	11,09	0,1109	0,1111	0,1111	0,1130	0,1130	0,0990	0,0990	0,1058	0,1058	0,1062
-1	21,08	0,2108	0,1111	0,1111	0,1152	0,1152	0,1251	0,1251	0,1262	0,1262	0,1253
0	26,11	0,2610	0,1111	0,1111	0,1160	0,1160	0,1387	0,1387	0,1313	0,1313	0,1324
1	21,08	0,2108	0,1111	0,1111	0,1152	0,1152	0,1251	0,1251	0,1262	0,1262	0,1253
2	11,09	0,1109	0,1111	0,1111	0,1130	0,1130	0,0990	0,0990	0,1058	0,1058	0,1062
3	3,80	0,0380	0,1111	0,1111	0,1094	0,1094	0,0862	0,0862	0,0807	0,0807	0,0806
4	0,98	0,0098	0,1111	0,1111	0,1045	0,1045	0,1204	0,1204	0,1217	0,1217	0,1217
		Krit.:	2,69119	2,69119	2,57125	2,57125	0,20884	0,20884	0,00247	0,00247	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

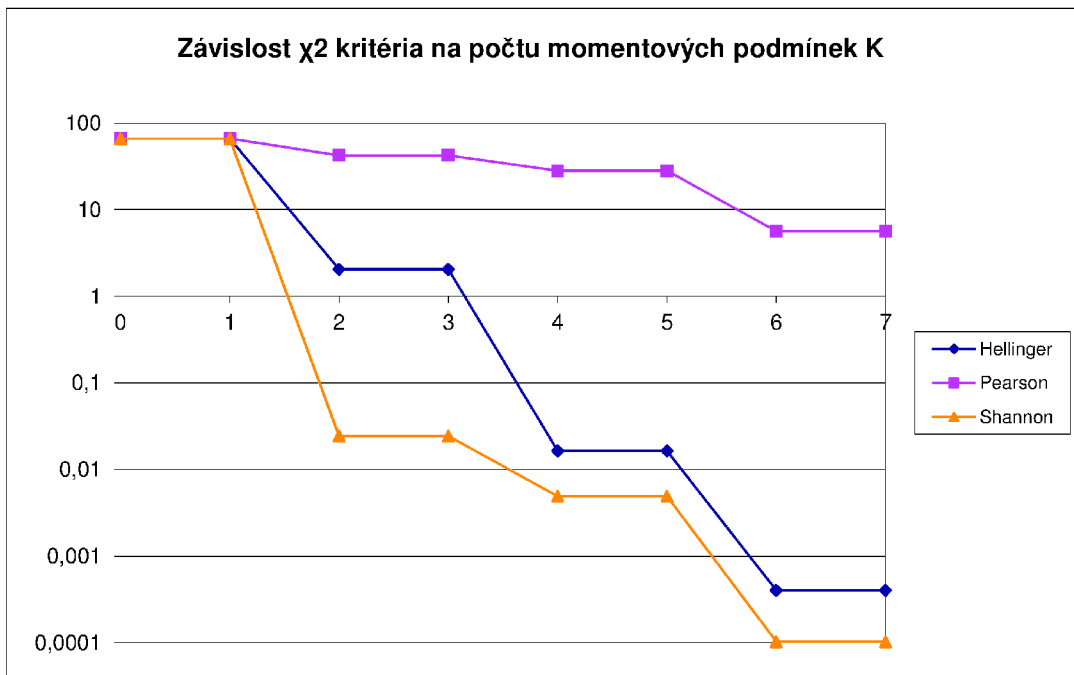
Obrázek 13: Výsledky pro normální rozdělení $N(0; 2,25)$



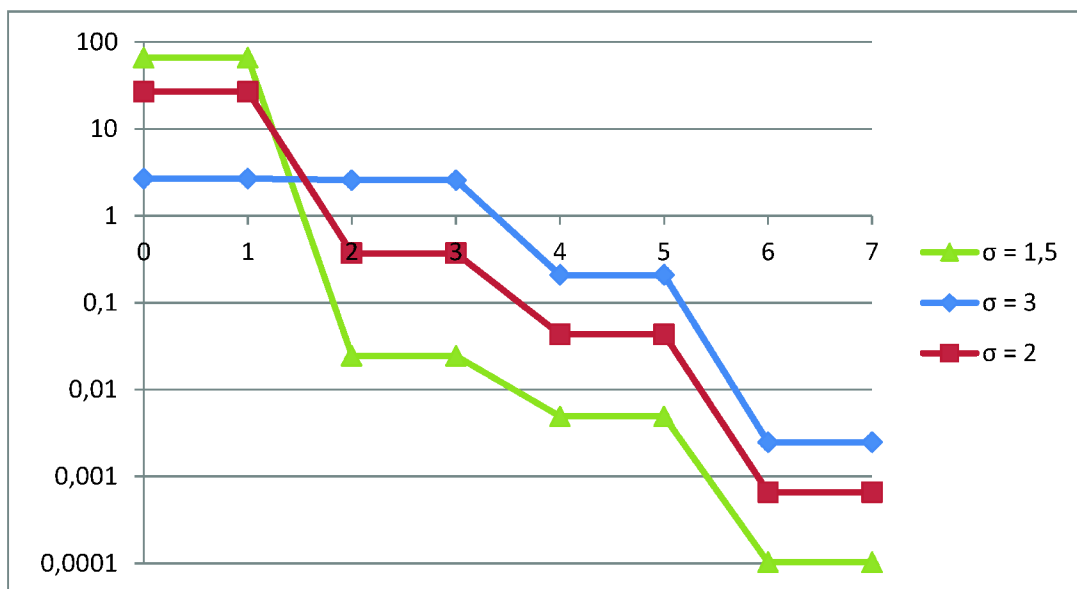
Obrázek 14: Graf závislosti χ^2 na K pro rozdělení $N(0; 9)$



Obrázek 15: Graf závislosti χ^2 na K pro rozdělení $N(0; 4)$



Obrázek 16: Graf závislosti χ^2 na K pro rozdělení $N(0; 4)$



Obrázek 17: Graf závislosti χ^2 na K pro Shannonovu kvazinormu

8.4 Weibullovo rozdělení

Posledním typem rozdělení, kterým se budeme zabývat je Weibullovo. Zvolíme-li hodnotu druhého parametru rovnu jedné, přejde Weibullovo rozdělení na exponenciální. Z tohoto důvodu je exponenciální rozdělení zařazeno do této podkapitoly jako speciální případ Weibullova. Odhad rozdělení budeme provádět na čtyřech rozděleních $W(10; 1) = E(0,1)$, $W(4; 1) = E(0,25)$, $W(6; 1,5)$ a $W(8; 2)$. Pro všechny zástupce bylo provedeno stejné třídění do 9 tříd a to takové že hranice tříd jsou sudá čísla od 0 do 16. Třídy jsou pak ve výsledkové tabulce reprezentovány lichými čísly od 1 do 17 (třída reprezentovaná hodnotou 17 odpovídá intervalu $(16; \infty)$). Pravděpodobnosti P_i a z nich pak teoretické četnosti f_i jsou vypočteny z distribučních funkcí zvolených rozdělení.

Odhadované pravděpodobnosti byly vypočteny opět Estimátorem pro počty momentových podmínek $K = 0 \dots 8$. Podrobné výsledky výpočtů jsou v tabulkách, uvedených z důvodu převedení z Excelu jako obrázky 18, 19, 20 a 21. V posledních dvou řádcích tabulky je testová statistika chí-kvadrát testu a příslušná kritická hodnota. Testové kritérium je vyznačeno barevně v případě, že chí-kvadrát test zamítá shodu pozorování a odhadu.

V grafech na obrázcích 22, 23, 24 a 25 je znázorněna závislost testového kritéria χ^2 na počtu momentových podmínek K . Z obrázků je názorně vidět srovnání našich tří kvazinorem z hlediska kvality aproximace.

Pro aproximaci exponenciálního rozdělení $E(0,1)$ dopadají všechny kvazinormy téměř shodně. Pro další zástupce jsou výsledky velmi podobné pro Hellingerovu a Shannonovu kvazinormu, k dobré shodě stačí jedna nebo dvě momentové omezení. Pearsonova kvazinorma dopadá o něco hůře, nicméně v porovnání se zkoumanými diskrétními rozděleními, tato kvazinorma dopadá přece jenom výrazně lépe.

E(0,1)		Hellinger									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	18,13	0,1813	0,1111	0,1404	0,2161	0,1670	0,1854	0,1804	0,1814	0,1813	0,1813
3	14,84	0,1484	0,1111	0,1314	0,1298	0,1705	0,1369	0,1524	0,1475	0,1485	0,1484
5	12,15	0,1215	0,1111	0,1233	0,0953	0,1302	0,1260	0,1162	0,1235	0,1212	0,1215
7	9,95	0,0995	0,1111	0,1159	0,0798	0,0913	0,1082	0,0995	0,0974	0,1001	0,0995
9	8,14	0,0814	0,1111	0,1092	0,0738	0,0674	0,0813	0,0859	0,0815	0,0807	0,0814
11	6,67	0,0667	0,1111	0,1030	0,0747	0,0564	0,0589	0,0665	0,0686	0,0673	0,0667
13	5,46	0,0546	0,1111	0,0973	0,0828	0,0567	0,0486	0,0495	0,0526	0,0543	0,0546
15	4,47	0,0447	0,1111	0,0921	0,1021	0,0758	0,0571	0,0487	0,0456	0,0448	0,0447
17	20,19	0,2019	0,1111	0,0873	0,1455	0,1847	0,1976	0,2010	0,2018	0,2019	0,2019
		Krit.:	22,7352	22,9867	8,57942	2,45808	0,64765	0,14512	0,02283	0,0016	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

E(0,1)		Pearson									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	18,13	0,1813	0,1111	0,1440	0,2295	0,1593	0,1868	0,1801	0,1815	0,1813	0,1813
3	14,84	0,1484	0,1111	0,1315	0,1125	0,1908	0,1314	0,1537	0,1473	0,1485	0,1484
5	12,15	0,1215	0,1111	0,1218	0,0902	0,1195	0,1329	0,1136	0,1241	0,1212	0,1215
7	9,95	0,0995	0,1111	0,1139	0,0815	0,0841	0,1076	0,1018	0,0965	0,1002	0,0995
9	8,14	0,0814	0,1111	0,1074	0,0785	0,0681	0,0771	0,0859	0,0823	0,0805	0,0814
11	6,67	0,0667	0,1111	0,1019	0,0795	0,0611	0,0594	0,0648	0,0682	0,0674	0,0667
13	5,46	0,0546	0,1111	0,0971	0,0852	0,0608	0,0519	0,0509	0,0527	0,0543	0,0546
15	4,47	0,0447	0,1111	0,0930	0,0994	0,0707	0,0549	0,0482	0,0456	0,0448	0,0447
17	20,19	0,2019	0,1111	0,0894	0,1435	0,1857	0,1980	0,2010	0,2018	0,2019	0,2019
		Krit.:	22,7352	21,7508	10,3472	3,00725	0,71893	0,16065	0,02816	0,0024	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

E(0,1)		Shannon									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	18,13	0,1813	0,1111	0,1394	0,2107	0,1683	0,1854	0,1803	0,1814	0,1813	0,1813
3	14,84	0,1484	0,1111	0,1313	0,1360	0,1668	0,1376	0,1524	0,1475	0,1485	0,1484
5	12,15	0,1215	0,1111	0,1237	0,0982	0,1321	0,1247	0,1165	0,1235	0,1212	0,1215
7	9,95	0,0995	0,1111	0,1165	0,0793	0,0938	0,1086	0,0989	0,0975	0,1001	0,0995
9	8,14	0,0814	0,1111	0,1098	0,0717	0,0670	0,0828	0,0862	0,0812	0,0807	0,0814
11	6,67	0,0667	0,1111	0,1035	0,0724	0,0540	0,0584	0,0672	0,0689	0,0673	0,0667
13	5,46	0,0546	0,1111	0,0975	0,0818	0,0551	0,0465	0,0484	0,0524	0,0543	0,0546
15	4,47	0,0447	0,1111	0,0918	0,1035	0,0796	0,0591	0,0492	0,0456	0,0448	0,0447
17	20,19	0,2019	0,1111	0,0865	0,1463	0,1833	0,1970	0,2009	0,2017	0,2019	0,2019
		Krit.:	22,7352	23,4697	8,12305	2,74963	0,80543	0,18106	0,0262	0,00167	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 18: Výsledky pro exponenciální rozdělení E(0,1)

E(0,25)		Hellinger									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	39,35	0,3935	0,1111	0,4536	0,3998	0,3859	0,3962	0,3929	0,3936	0,3935	0,3935
3	23,86	0,2386	0,1111	0,2022	0,2347	0,2539	0,2300	0,2413	0,2381	0,2387	0,2386
5	14,47	0,1447	0,1111	0,1139	0,1392	0,1433	0,1516	0,1404	0,1460	0,1445	0,1447
7	8,78	0,0878	0,1111	0,0730	0,0850	0,0812	0,0905	0,0897	0,0862	0,0882	0,0878
9	5,33	0,0533	0,1111	0,0507	0,0538	0,0492	0,0507	0,0550	0,0539	0,0529	0,0533
11	3,23	0,0323	0,1111	0,0373	0,0352	0,0322	0,0294	0,0312	0,0329	0,0327	0,0323
13	1,96	0,0196	0,1111	0,0285	0,0238	0,0228	0,0193	0,0182	0,0188	0,0194	0,0196
15	1,19	0,0119	0,1111	0,0226	0,0166	0,0174	0,0155	0,0134	0,0123	0,0120	0,0119
17	1,83	0,0183	0,1111	0,0183	0,0119	0,0142	0,0168	0,0179	0,0182	0,0183	0,0183
		Krit.:	120,648	3,44822	0,62782	0,53248	0,21068	0,057	0,01083	0,00113	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

E(0,25)		Pearson									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	39,35	0,3935	0,1111	0,5951	0,3035	0,3901	0,4045	0,3905	0,3939	0,3935	0,3935
3	23,86	0,2386	0,1111	0,0920	0,4142	0,2406	0,1982	0,2537	0,2359	0,2388	0,2386
5	14,47	0,1447	0,1111	0,0654	0,0945	0,1563	0,1917	0,1161	0,1518	0,1439	0,1447
7	8,78	0,0878	0,1111	0,0535	0,0551	0,0794	0,0782	0,1103	0,0781	0,0894	0,0878
9	5,33	0,0533	0,1111	0,0464	0,0390	0,0471	0,0434	0,0511	0,0599	0,0513	0,0533
11	3,23	0,0323	0,1111	0,0415	0,0302	0,0316	0,0292	0,0273	0,0315	0,0339	0,0323
13	1,96	0,0196	0,1111	0,0379	0,0246	0,0230	0,0220	0,0185	0,0178	0,0188	0,0196
15	1,19	0,0119	0,1111	0,0351	0,0208	0,0177	0,0178	0,0153	0,0131	0,0121	0,0119
17	1,83	0,0183	0,1111	0,0329	0,0180	0,0142	0,0152	0,0173	0,0181	0,0183	0,0183
		Krit.:	120,648	45,3599	15,7359	0,62571	2,668	1,44383	0,26182	0,02277	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

E(0,25)		Shannon									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	39,35	0,3935	0,1111	0,3842	0,4011	0,3896	0,3948	0,3932	0,3935	0,3935	0,3935
3	23,86	0,2386	0,1111	0,2386	0,2327	0,2455	0,2346	0,2400	0,2383	0,2386	0,2386
5	14,47	0,1447	0,1111	0,1482	0,1390	0,1458	0,1471	0,1427	0,1454	0,1446	0,1447
7	8,78	0,0878	0,1111	0,0920	0,0855	0,0847	0,0900	0,0883	0,0871	0,0880	0,0878
9	5,33	0,0533	0,1111	0,0572	0,0542	0,0501	0,0525	0,0545	0,0535	0,0531	0,0533
11	3,23	0,0323	0,1111	0,0355	0,0354	0,0312	0,0303	0,0320	0,0328	0,0325	0,0323
13	1,96	0,0196	0,1111	0,0220	0,0238	0,0213	0,0188	0,0184	0,0190	0,0195	0,0196
15	1,19	0,0119	0,1111	0,0137	0,0165	0,0166	0,0146	0,0129	0,0121	0,0119	0,0119
17	1,83	0,0183	0,1111	0,0085	0,0118	0,0152	0,0173	0,0180	0,0183	0,0183	0,0183
		Krit.:	120,648	1,28435	0,65304	0,26756	0,09037	0,02296	0,00392	0,00035	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

Obrázek 19: Výsledky pro exponenciální rozdělení E(0,25)

W(6;1,5)		Hellinger									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	17,50	0,1750	0,1111	0,3108	0,1803	0,1787	0,1780	0,1745	0,1751	0,1750	0,1750
3	24,47	0,2447	0,1111	0,1969	0,2315	0,2332	0,2347	0,2467	0,2439	0,2448	0,2447
5	21,24	0,2124	0,1111	0,1359	0,2198	0,2212	0,2209	0,2095	0,2143	0,2122	0,2124
7	15,34	0,1534	0,1111	0,0993	0,1576	0,1571	0,1562	0,1540	0,1510	0,1539	0,1534
9	9,82	0,0982	0,1111	0,0758	0,0955	0,0946	0,0945	0,1003	0,0993	0,0976	0,0982
11	5,72	0,0572	0,1111	0,0597	0,0545	0,0540	0,0543	0,0562	0,0578	0,0577	0,0572
13	3,08	0,0308	0,1111	0,0483	0,0312	0,0311	0,0314	0,0295	0,0297	0,0306	0,0308
15	1,55	0,0155	0,1111	0,0398	0,0184	0,0186	0,0186	0,0168	0,0160	0,0156	0,0155
17	1,28	0,0128	0,1111	0,0334	0,0113	0,0115	0,0114	0,0125	0,0127	0,0128	0,0128
		Krit.:	56,0746	18,4067	0,21463	0,20551	0,18684	0,02972	0,01314	0,00113	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

W(6;1,5)		Pearson									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	17,50	0,1750	0,1111	0,4143	0,0968	0,2149	0,1693	0,1734	0,1757	0,1749	0,1750
3	24,47	0,2447	0,1111	0,1302	0,3225	0,1425	0,2688	0,2522	0,2403	0,2453	0,2447
5	21,24	0,2124	0,1111	0,0944	0,3077	0,2589	0,1768	0,1995	0,2242	0,2105	0,2124
7	15,34	0,1534	0,1111	0,0778	0,0955	0,2061	0,1699	0,1606	0,1371	0,1573	0,1534
9	9,82	0,0982	0,1111	0,0676	0,0571	0,0745	0,1056	0,1017	0,1096	0,0934	0,0982
11	5,72	0,0572	0,1111	0,0607	0,0408	0,0412	0,0508	0,0529	0,0553	0,0611	0,0572
13	3,08	0,0308	0,1111	0,0555	0,0317	0,0272	0,0284	0,0294	0,0282	0,0289	0,0308
15	1,55	0,0155	0,1111	0,0514	0,0259	0,0197	0,0180	0,0181	0,0173	0,0161	0,0155
17	1,28	0,0128	0,1111	0,0482	0,0220	0,0151	0,0124	0,0120	0,0124	0,0127	0,0128
		Krit.:	56,0746	53,6041	19,0688	11,8046	1,30241	0,23498	0,43329	0,07571	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

W(6;1,5)		Shannon									
Třída	f _i	P _i	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	17,50	0,1750	0,1111	0,2786	0,2013	0,1792	0,1773	0,1749	0,1751	0,1750	0,1750
3	24,47	0,2447	0,1111	0,2061	0,2140	0,2336	0,2372	0,2449	0,2443	0,2447	0,2447
5	21,24	0,2124	0,1111	0,1525	0,1972	0,2182	0,2184	0,2125	0,2134	0,2123	0,2124
7	15,34	0,1534	0,1111	0,1129	0,1575	0,1583	0,1560	0,1527	0,1523	0,1536	0,1534
9	9,82	0,0982	0,1111	0,0835	0,1091	0,0968	0,0956	0,0990	0,0986	0,0980	0,0982
11	5,72	0,0572	0,1111	0,0618	0,0654	0,0540	0,0546	0,0574	0,0577	0,0574	0,0572
13	3,08	0,0308	0,1111	0,0457	0,0340	0,0299	0,0310	0,0300	0,0301	0,0307	0,0308
15	1,55	0,0155	0,1111	0,0338	0,0154	0,0178	0,0183	0,0160	0,0158	0,0155	0,0155
17	1,28	0,0128	0,1111	0,0250	0,0060	0,0123	0,0116	0,0127	0,0127	0,0128	0,0128
		Krit.:	56,0746	10,7501	1,92532	0,14734	0,12339	0,00533	0,00391	0,00012	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

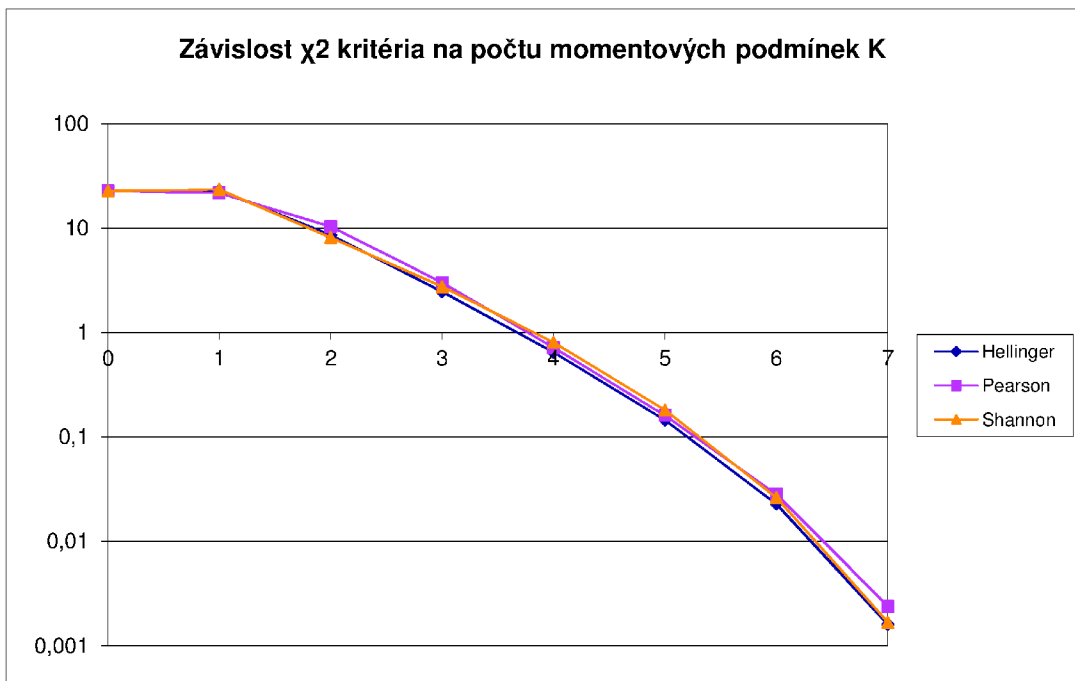
Obrázek 20: Výsledky pro Weibullovo rozdělení W(6; 1,5)

W(8;2)		Hellinger									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	6,06	0,0606	0,1111	0,1949	0,0772	0,0715	0,0646	0,0607	0,0608	0,0606	0,0606
3	16,06	0,1606	0,1111	0,1613	0,1354	0,1370	0,1477	0,1598	0,1594	0,1607	0,1606
5	20,90	0,2090	0,1111	0,1357	0,2006	0,2103	0,2187	0,2113	0,2119	0,2088	0,2090
7	20,19	0,2019	0,1111	0,1158	0,2175	0,2207	0,2078	0,1987	0,1984	0,2024	0,2019
9	15,83	0,1583	0,1111	0,0999	0,1663	0,1604	0,1526	0,1600	0,1597	0,1577	0,1583
11	10,42	0,1042	0,1111	0,0871	0,1004	0,0954	0,0996	0,1051	0,1054	0,1047	0,1042
13	5,86	0,0586	0,1111	0,0766	0,0552	0,0540	0,0599	0,0567	0,0568	0,0584	0,0586
15	2,85	0,0285	0,1111	0,0679	0,0302	0,0314	0,0328	0,0295	0,0293	0,0286	0,0285
17	1,83	0,0183	0,1111	0,0606	0,0171	0,0192	0,0163	0,0181	0,0182	0,0183	0,0183
		Krit.:	38,9663	29,0215	1,06572	0,88815	0,32347	0,02061	0,02183	0,00064	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

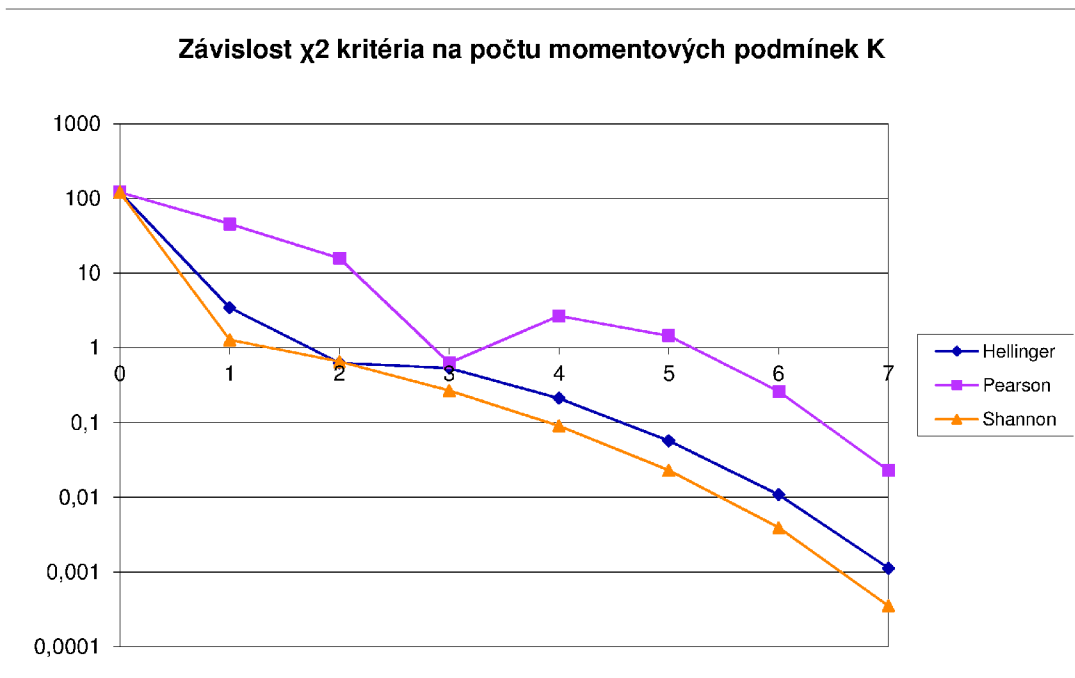
W(8;2)		Pearson									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	6,06	0,0606	0,1111	0,2246	0,0644	0,0938	0,0612	0,0629	0,0612	0,0605	0,0606
3	16,06	0,1606	0,1111	0,1512	0,1042	0,1060	0,1542	0,1482	0,1569	0,1612	0,1606
5	20,90	0,2090	0,1111	0,1216	0,2514	0,1584	0,2297	0,2353	0,2187	0,2067	0,2090
7	20,19	0,2019	0,1111	0,1045	0,2864	0,3138	0,1738	0,1762	0,1891	0,2064	0,2019
9	15,83	0,1583	0,1111	0,0931	0,1110	0,1516	0,1718	0,1662	0,1662	0,1526	0,1583
11	10,42	0,1042	0,1111	0,0847	0,0670	0,0738	0,1087	0,1098	0,1044	0,1087	0,1042
13	5,86	0,0586	0,1111	0,0783	0,0478	0,0459	0,0525	0,0536	0,0552	0,0563	0,0586
15	2,85	0,0285	0,1111	0,0731	0,0372	0,0323	0,0294	0,0295	0,0304	0,0291	0,0285
17	1,83	0,0183	0,1111	0,0688	0,0304	0,0244	0,0186	0,0184	0,0179	0,0182	0,0183
		Krit.:	38,9663	39,33	11,2817	11,4232	0,86621	0,89875	0,21024	0,06366	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

W(8;2)		Shannon									
Třída	fi	Pi	pi (K=0)	pi (K=1)	pi (K=2)	pi (K=3)	pi (K=4)	pi (K=5)	pi (K=6)	pi (K=7)	pi (K=8)
1	6,06	0,0606	0,1111	0,1874	0,0843	0,0662	0,0633	0,0607	0,0607	0,0606	0,0606
3	16,06	0,1606	0,1111	0,1615	0,1403	0,1479	0,1524	0,1601	0,1601	0,1606	0,1606
5	20,90	0,2090	0,1111	0,1391	0,1870	0,2114	0,2143	0,2102	0,2103	0,2089	0,2090
7	20,19	0,2019	0,1111	0,1199	0,1995	0,2106	0,2067	0,2005	0,2005	0,2020	0,2019
9	15,83	0,1583	0,1111	0,1033	0,1705	0,1592	0,1554	0,1587	0,1587	0,1581	0,1583
11	10,42	0,1042	0,1111	0,0890	0,1166	0,0995	0,1001	0,1050	0,1050	0,1043	0,1042
13	5,86	0,0586	0,1111	0,0767	0,0639	0,0560	0,0588	0,0576	0,0576	0,0585	0,0586
15	2,85	0,0285	0,1111	0,0661	0,0280	0,0310	0,0324	0,0289	0,0289	0,0285	0,0285
17	1,83	0,0183	0,1111	0,0569	0,0098	0,0183	0,0166	0,0182	0,0182	0,0183	0,0183
		Krit.:	38,9663	26,0713	2,21339	0,24843	0,16474	0,00492	0,00492	5,6E-05	0
		Kvantil:	16,746	14,0692	12,5936	11,0726	9,49002	7,81688	5,99146	3,84315	0

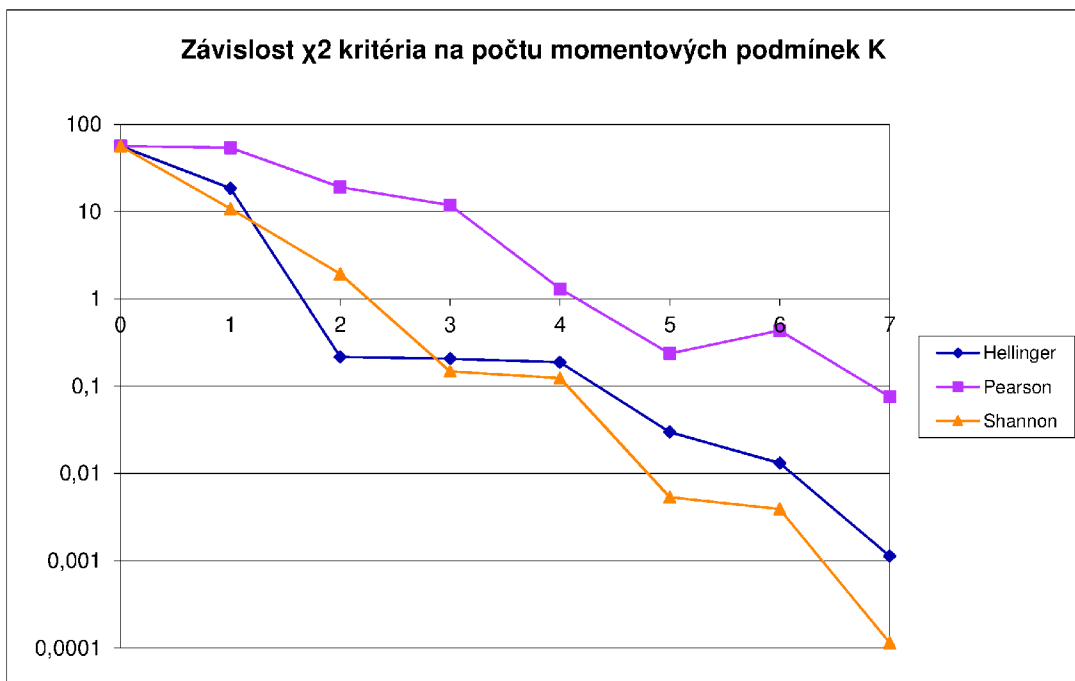
Obrázek 21: Výsledky pro Weibullovo rozdělení W(8; 2)



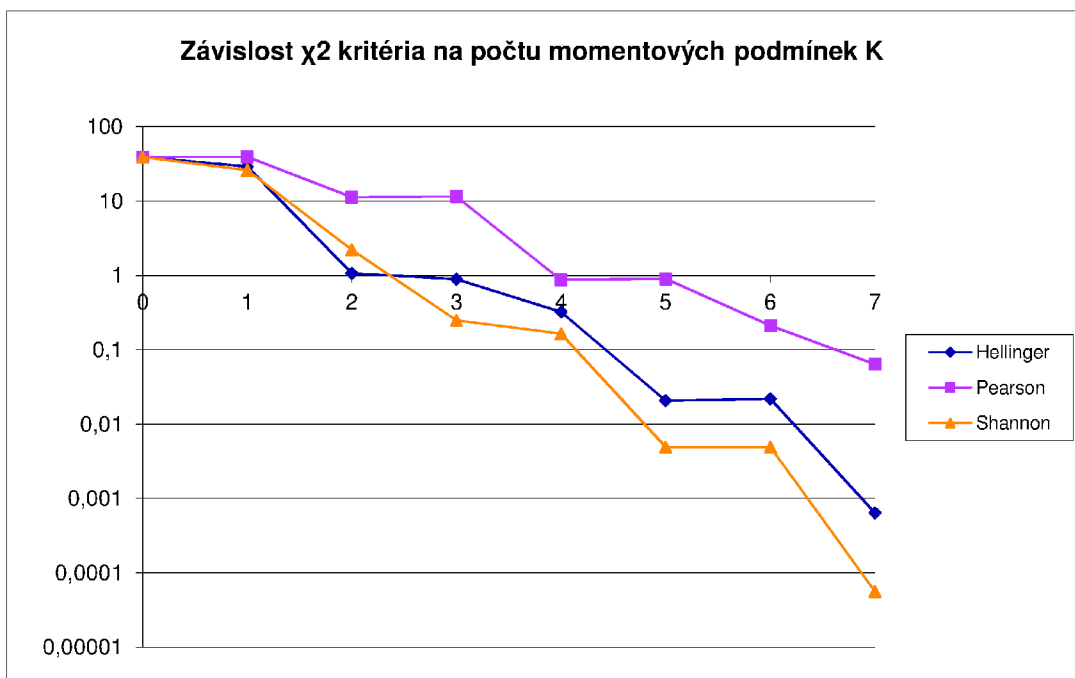
Obrázek 22: Graf závislosti χ^2 na K pro rozdělení $E(0,1)$



Obrázek 23: Graf závislosti χ^2 na K pro rozdělení $E(0,25)$



Obrázek 24: Graf závislosti χ^2 na K pro rozdělení $W(6; 1,5)$



Obrázek 25: Graf závislosti χ^2 na K pro rozdělení $W(8; 2)$

9 Odhad rozdělení s obecnými lineárními podmínkami

Předpokládáme stejně jako v kapitole 5, že pozorovaná diskrétní náhodná veličina X na (Ω, Σ, P) , jejíž rozdělení pravděpodobnosti chceme odhadnout (fitovat), nabývá nejvýše konečně mnoha různých hodnot $x_j^* \in \mathbb{R}$, (tj. $\Omega = \{x_1^*, \dots, x_m^*\} \subset \mathbb{R}$) s neznámými pravděpodobnostmi

$$p_j = P(X = x_j^*), \quad j = 1, \dots, m, \quad m > 1.$$

Pozorováním náhodné veličiny X získáme statistický soubor (x_1, \dots, x_n) a jeho roztríděním dostaneme roztríděný statistický soubor

$$\left(\left(x_1^*, \frac{f_1}{n} \right), \dots, \left(x_m^*, \frac{f_m}{n} \right) \right),$$

kde f_j je absolutní četnost pozorované hodnoty x_j^* . Dále předpokládáme, že $n > m$ a $f_j > 0$ pro všechna $j = 1, \dots, m$ (jestliže $f_j = 0$, pak j -tou třídu vynecháme). Pro odhad rozdělení \mathbf{p} požadujeme, aby toto rozdělení navíc splňovalo nějaké zadané podmínky, jejichž počet je $K \geq 1$. Mezi tyto podmínky zařazujeme zřejmou podmínku $\sum_{j=1}^m p_j = 1$. Hledáme pak takové rozdělení \mathbf{p} , které má minimální kvazinormu $D_f(\mathbf{p}, \mathbf{p}_0)$ za těchto dodatečných podmínek. V kapitole 5 jsme se zabývali **momentovými podmínkami**

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad k = 0, \dots, K.$$

Jedná se o lineární podmínky v proměnných p_j s konkrétními koeficienty na levých stranách, a sice s mocninami pozorovaných hodnot x_j^{*k} . V této kapitole se zaměřujeme na **obecné lineární podmínky**

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K,$$

zapsáno maticově

$$\mathbf{A}\mathbf{p} = \mathbf{b},$$

kde

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \dots & a_{Km} \end{pmatrix}$$

je matice typu $K + 1$ krát m a

$$\mathbf{b} = \begin{pmatrix} 1 \\ b_1 \\ b_2 \\ \vdots \\ b_K \end{pmatrix}$$

je sloupcový vektor pravých stran omezení. Hledáme pak takové rozdělení \mathbf{p} , které má minimální kvazinormu $D_f(\mathbf{p}, \mathbf{p}_0)$.

9.1 Hellingerova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Hellingerovu kvazinormu za K počátečních obecných podmínek**

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K,$$

jestliže jeho Hellingerova kvazinorma

$$H(\mathbf{p}, \mathbf{p}_0) = 2 - \frac{2}{\sqrt{m}} \sum_{j=1}^m \sqrt{p_j}$$

je minimální pro

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme

$$p_j(\boldsymbol{\lambda}) = \frac{1}{m \left(\sum_{k=0}^K \lambda_k a_{kj} \right)^2}, \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$ jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = H(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j a_{kj} - b_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce, anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $H_K = \min H(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Hellingerovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$H_K = 2 - 2 \sum_{j=1}^m \left| \sum_{k=0}^K \lambda_k a_{kj} \right|.$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $H_0 = 0$.

9.2 Shannonova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Shannonovu kvazinormu za K počátečních obecných podmínek**

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K,$$

jestliže jeho Shannonova kvazinorma

$$S(\mathbf{p}, \mathbf{p}_0) = \sum_{j=1}^m p_j \ln p_j + \ln m$$

je minimální pro

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme

$$p_j(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \lambda_k a_{kj}\right), \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$ jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = S(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j a_{kj} - M_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce

$$\sum_{j=1}^m \exp\left(-1 + \sum_{k=0}^K \lambda_k a_{kj}\right) a_{\nu j} = b_\nu, \quad \nu = 1, \dots, K,$$

anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $S_K = \min S(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Shannonovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$S_K = \ln m - \sum_{j=1}^m \left(1 + \sum_{k=0}^K \lambda_k a_{kj}\right) \exp\left(-1 - \sum_{k=0}^K \lambda_k a_{kj}\right).$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $S_0 = 0$.

9.3 Pearsonova kvazinorma

Rozdělení pravděpodobnosti $\mathbf{p} = (p_1, \dots, p_m)$ pozorované diskrétní náhodné veličiny X má na pravděpodobnostním prostoru (Ω, Σ, P) , kde $\Omega = \{x_1^*, \dots, x_m^*\}$, $m > 0$ a Σ je množina všech podmnožin Ω , **minimální Pearsonovu kvazinormu za K počátečních momentových podmínek**

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K,$$

jestliže jeho Pearsonova kvazinorma

$$P(\mathbf{p}, \mathbf{p}_0) = \frac{1}{m^2} \sum_{j=1}^m \frac{1}{p_j} - 1$$

je minimální pro

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K.$$

Pro $K < m - 1$ obdržíme

$$p_j(\boldsymbol{\lambda}) = \frac{1}{m \sqrt{\sum_{k=0}^K \lambda_k a_{kj}}}, \quad j = 1, \dots, m,$$

kde λ_k , $k = 0, \dots, K$ jsou Lagrangeovy multiplikátory pro Lagrangeovu funkci

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = P(\mathbf{p}, \mathbf{p}_0) + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j a_{kj} - b_k \right)$$

a $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$.

Lagrangeovy multiplikátory λ_k je možno určit pomocí nelineární soustavy rovnic odpovídající nulovému gradientu Lagrangeovy funkce, anebo přímo aplikovat některou metodu nelineární optimalizace pro určení jejího minima.

Jestliže označíme $P_K = \min P(\mathbf{p}(\boldsymbol{\lambda}), \mathbf{p}_0)$, kde $\mathbf{p}(\boldsymbol{\lambda}) = (p_1(\boldsymbol{\lambda}), \dots, p_m(\boldsymbol{\lambda}))$ je odhad rozdělení pravděpodobnosti s minimální Shannonovou kvazinormou za daných $K < m - 1$ momentových podmínek, pak

$$P_K = \frac{1}{m} \sum_{j=1}^m \sqrt{\sum_{k=0}^K \lambda_k a_{kj}} - 1.$$

Pro $K = 0$ je $p_j = \frac{1}{m}$, $j = 1, \dots, m$ a $P_0 = 0$.

9.4 Příklad

Předpokládejme, že první podmínkou ($k = 1$) v zápise

$$\sum_{j=1}^m p_j a_{kj} = b_k, \quad k = 0, \dots, K$$

chceme zaručit rovnost geometrických průměrů pro pozorování a odhad, tj.

$$\prod_{i=1}^m x_i^{*p_i} = \prod_{i=1}^m x_i^{*\frac{f_i}{n}}.$$

Tuto nelineární podmínku můžeme snadno linearizovat logaritmováním

$$\sum_{i=1}^m p_i \ln(x_i^*) = \sum_{i=1}^m \frac{f_i}{n} \ln(x_i^*) = G_1.$$

Spolu se samozřejmou podmínkou řádu 0 můžeme tuto podmínku zapsat maticovým zápisem $\mathbf{A}\mathbf{p} = \mathbf{b}$ volbou

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \ln(x_1^*) & \ln(x_2^*) & \dots & \ln(x_m^*) \end{pmatrix}$$

a

$$\mathbf{b} = \begin{pmatrix} 1 \\ G_1 \end{pmatrix}.$$

Pokud bychom chtěli zaručit ještě navíc stejnou variabilitu pro pozorování a odhad, přidáme ještě další podmínku

$$\sum_{i=1}^m p_i \ln^2(x_i^*) = \sum_{i=1}^m \frac{f_i}{n} \ln^2(x_i^*) = G_2$$

a v maticovém zápisu $\mathbf{A}\mathbf{p} = \mathbf{b}$ volíme

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \ln(x_1^*) & \ln(x_2^*) & \dots & \ln(x_m^*) \\ \ln^2(x_1^*) & \ln^2(x_2^*) & \dots & \ln^2(x_m^*) \end{pmatrix}$$

a

$$\mathbf{b} = \begin{pmatrix} 1 \\ G_1 \\ G_2 \end{pmatrix}.$$

Jedná se rovnost "geometrických" momentů druhého řádu.

Mějme nyní konkrétní datový soubor o rozsahu $n = 100$, jehož roztříděním dostaneme četnostní tabulku 12. Předpokládejme, že tato data vyjadřují koeficienty růstu nějaké

Tabulka 12: Tabulka četností

x_i^*	0,8	1	1,2	1,4	1,6
f_i	17	35	27	14	7

veličiny za 100 časových období, např. měsíců. Je tedy věcně správnější pracovat s geometrickým průměrem (příp. obecně s geometrickými momenty vyšších řádů) než aritmetickým průměrem (příp. obecně s aritmetickými momenty vyšších řádů). Vypočítejme a srovnáme tedy odhady pomocí minimalizace Pearsonovy kvazinormy v prvním případě při použití klasických aritmetických momentových podmínek, jejichž použití není pro taková data věcně správné a v druhém případě při použití správnějších geometrických momentových podmínek. V obou případech volíme rovnost momentů až do řádu $K = 1$, příp. $K = 2$.

Počet tříd je $m = 5$ rozsah $n = 100$, takže

$$M_1 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^* = 1,301,$$

$$M_2 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*2} = 1,096,$$

$$G_1 = \frac{1}{100} \sum_{i=1}^5 f_i \ln(x_i^*) = 0,0913,$$

$$G_2 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^2(x_i^*) = 0,0488.$$

Pomocí optimalizačního nástroje Řešitel z Excelu pro určení minima Pearsonovy kvazinormy jsme získali výsledky v tabulkách 13 a 14. První tabulka 13 odpovídá omezením ve tvaru momentových podmínek pro $K = 1$, příp. $K = 2$ a druhá tabulka 14 odpovídá omezením ve tvaru "geometrických" momentových podmínek, popsaných výše.

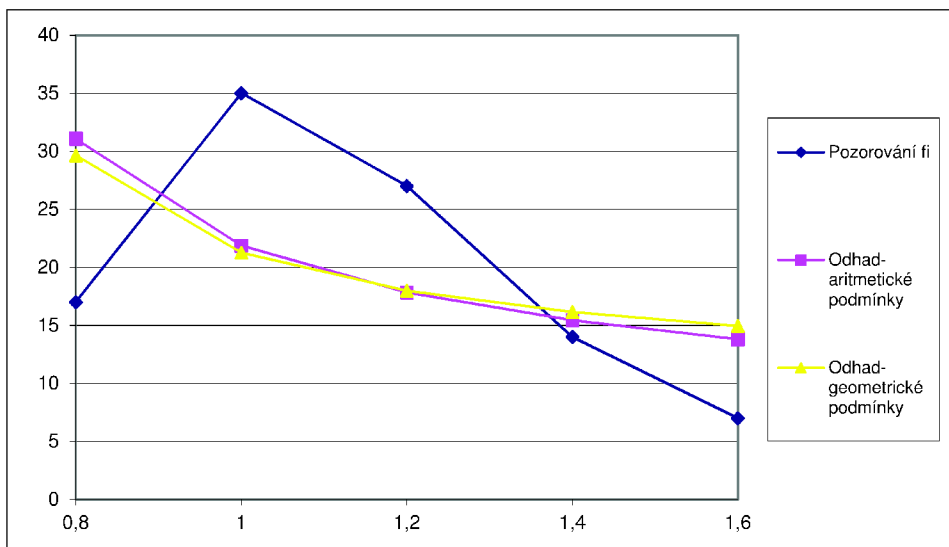
Tabulka 13: Výsledky pro aritmetické momentové podmínky

x_i	0,8	1	1,2	1,4	1,6
f_i	17	35	27	14	7
odhad $K = 1$	31,1	21,9	17,8	15,4	13,8
odhad $K = 2$	15,6	38,1	25,9	12,4	8,0

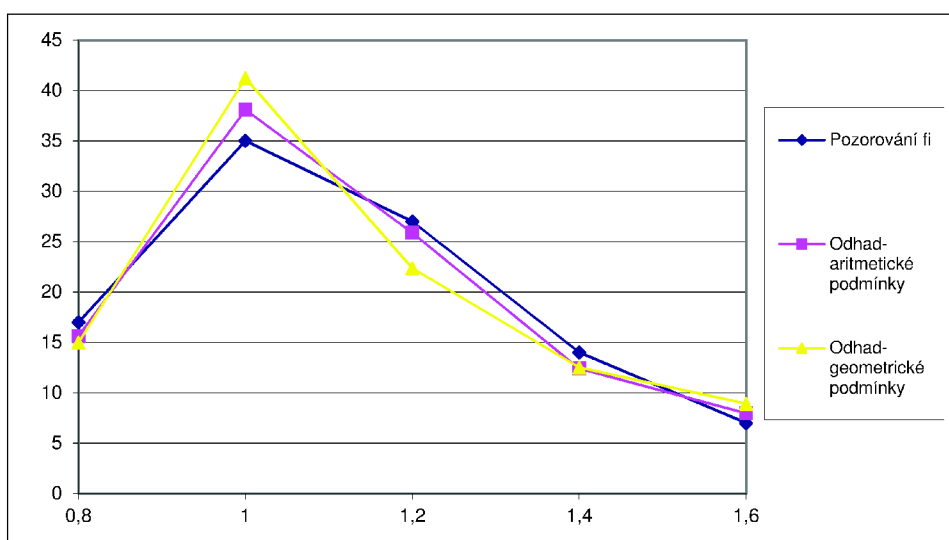
Pro přehled uveďme ještě grafické znázornění výsledků. Na obrázku 26 je srovnáno pozorování s odhady při použití aritmetických a geometrických momentových podmínek pro $K = 1$ (rovnost průměrů). Na obrázku 27 je srovnání téhož pro $K = 2$ (rovnost průměrů a druhých momentů).

Tabulka 14: Výsledky pro geometrické momentové podmínky

x_i	0,8	1	1,2	1,4	1,6
f_i	17	35	27	14	7
odhad 1	29,6	21,3	18,0	16,2	15,0
odhad 2	15,0	41,2	22,4	12,5	8,9



Obrázek 26: Znázornění výsledků pro $K = 1$



Obrázek 27: Znázornění výsledků pro $K = 2$

10 Intervalové odhady pravděpodobností p_j pro Shannonovu kvazinormu

10.1 Odvození rozdělení pro vektor Lagrangeových multiplikátorů

V případě Shannonovy kvazinormy jsou odhady parametrů λ_k maximálně věrohodné [5], [6], neboť jde současně o odhady parametrů modifikovanou metodou minimálního chí-kvadrát [2]. Odhad vektoru $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$ má asymptoticky $(K+1)$ -rozměrné normální rozdělení $N(\boldsymbol{\lambda}_0, n^{-1}[\mathbf{J}(\boldsymbol{\lambda}_0)]^{-1})$, kde $\boldsymbol{\lambda}_0$ je teoretický vektor parametrů rozdělení náhodné veličiny X a $\mathbf{J}(\boldsymbol{\lambda}_0)$ je Fisherova informační matice [3].

Speciálně odhad λ_k teoretické hodnoty jednoho pevně zvoleného multiplikátoru λ_{0k} , $k = 0, \dots, K$, má asymptoticky normální rozdělení $N\left(\lambda_{0k}, \frac{1}{nJ(\lambda_{0k})}\right)$. Označme $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$ odhad vektoru Lagrangeových multiplikátorů, který dostaneme řešením minimalizační úlohy pro konkrétní pozorování. Pro výpočet intervalového odhadu λ_{0k} pomocí tohoto normálního rozdělení klademe $\lambda_{0k} = \bar{\lambda}_k$. Fisherova míra informace je

$$J(\lambda_{0k}) = E\left(\left[\frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0k}}\right]^2\right),$$

kde E značí střední hodnotu, $J(\lambda_{0k}) = J(\bar{\lambda}_k)$ a $x = x_j^*$. Po úpravách dostáváme

$$J(\bar{\lambda}_k) = \sum_{j=1}^m p_j x_j^{*2k} \approx \sum_{j=1}^m \frac{f_j}{n} x_j^{*2k} = M_{2k},$$

kde pravděpodobnosti p_j jsou určeny pomocí minimalizace Shannonovy kvazinormy.

Obdobně pro prvek Fisherovy informační matice dostáváme J_{ts} dostáváme pro $t, s = 0, \dots, K$

$$J_{ts} = E\left(\frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0t}} \cdot \frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0s}}\right).$$

Použijeme-li jako apriorní odhad $p_j \approx \frac{f_j}{n}$ dostaneme

$$J_{ts} = \sum_{j=1}^m \frac{f_j}{n} x_j^{*(t+s)} = M_{t+s}.$$

Předchozí úvahy můžeme shrnout do následující věty.

Věta 10.1 Vektor Lagrangeových multiplikátorů $\boldsymbol{\lambda}$ má asymptoticky $(K+1)$ -rozměrné normální rozdělení $N(\bar{\boldsymbol{\lambda}}, [n\mathbf{J}]^{-1})$, kde vektor středních hodnot $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$

dostaneme řešením minimalizační úlohy a odhad Fischerovy informační matice je

$$\mathbf{J} = \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_K \\ M_1 & M_2 & M_3 & \dots & M_{K+1} \\ M_2 & M_3 & M_4 & \dots & M_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_K & M_{K+1} & M_{K+2} & \dots & M_{2K} \end{pmatrix}.$$

10.2 Odvození rozdělení pro vektor pravděpodobností

Nyní se budeme zabývat intervalovými odhady pravděpodobností p_j . K tomu je potřeba zjistit tvar rozdělení vektoru $\mathbf{p} = (p_1, \dots, p_m)$. Protože vektor \mathbf{p} je závislý na vektoru $\boldsymbol{\lambda}$, který má asymptoticky normální rozdělení, je zřejmé, že rozdělení vektoru \mathbf{p} je opět asymptoticky normální. Závislost má tvar

$$p_j(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \lambda_k x_j^{*k}\right), \quad j = 1, \dots, m.$$

Uvedený nelineární vztah linearizujeme na tvar $\mathbf{p} = \mathbf{a} + \mathbf{B}\boldsymbol{\lambda}$ pomocí Taylorova polynomu prvního stupně. Jako střed tohoto polynomu zvolíme bod $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$. Položíme

$$\bar{p}_j = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k x_j^{*k}\right), \quad j = 1, \dots, m.$$

V okolí bodu $\bar{\mathbf{p}} = (\bar{p}_1, \dots, \bar{p}_m)$ nás rozdělení pravděpodobnosti vektoru \mathbf{p} bude nejvíce zajímat. Taylorův polynom má tvar

$$T(p_j)(\boldsymbol{\lambda}) = p_j(\bar{\boldsymbol{\lambda}}) + \sum_{i=0}^K \frac{\partial p_j(\bar{\lambda}_i)}{\partial \lambda_i} (\lambda_i - \bar{\lambda}_i).$$

Dosazením za

$$\frac{\partial p_j(\bar{\lambda}_i)}{\partial \lambda_i} = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k x_j^{*k}\right) x_j^{*i}, \quad j = 1, \dots, m$$

dostáváme

$$T(p_j)(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k x_j^{*k}\right) + \sum_{i=0}^K x_j^{*i} \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k x_j^{*k}\right) (\lambda_i - \bar{\lambda}_i).$$

Po úpravě

$$T(p_j)(\boldsymbol{\lambda}) = \bar{p}_j \left(1 - \sum_{i=0}^K \bar{\lambda}_i x_j^{*i}\right) + \bar{p}_j \sum_{i=0}^K \lambda_i x_j^{*i}.$$

Můžeme tedy psát

$$\mathbf{p} \approx \mathbf{a} + \mathbf{B}\boldsymbol{\lambda},$$

kde

$$\mathbf{a} = \begin{pmatrix} \bar{p}_1 \left(1 - \sum_{i=0}^K \bar{\lambda}_i x_1^{*i} \right) \\ \vdots \\ \bar{p}_m \left(1 - \sum_{i=0}^K \bar{\lambda}_i x_m^{*i} \right) \end{pmatrix}$$

a

$$\mathbf{B} = \begin{pmatrix} \bar{p}_1 x_1^{*0} & \bar{p}_1 x_1^{*1} & \dots & \bar{p}_1 x_1^{*K} \\ \bar{p}_2 x_2^{*0} & \bar{p}_2 x_2^{*1} & \dots & \bar{p}_2 x_2^{*K} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{p}_m x_m^{*0} & \bar{p}_m x_m^{*1} & \dots & \bar{p}_m x_m^{*K} \end{pmatrix}.$$

Využitím věty o lineární transformaci normálně rozděleného náhodného vektoru [3] dostaneme parametry normálního rozdělení vektoru \mathbf{p} . Je zřejmé, že střední hodnota vektoru $E(\mathbf{p}) = \mathbf{a} + \mathbf{B}\boldsymbol{\lambda} = \mathbf{p} = (\bar{p}_1, \dots, \bar{p}_m)$. Předcházející úvahy můžeme shrnout do následující věty.

Věta 10.2 *Vektor pravděpodobností \mathbf{p} má asymptoticky m -rozměrné normální rozdělení $N(\bar{\mathbf{p}}, \mathbf{B} [n\mathbf{J}]^{-1} \mathbf{B}^T)$, kde vektor středních hodnot $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$ a vektor $\bar{\mathbf{p}} = (\bar{p}_1, \dots, \bar{p}_m)$ dostaneme minimalizací Shannonovy kvazinormy za daných podmínek,*

$$\mathbf{J} = \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_K \\ M_1 & M_2 & M_3 & \dots & M_{K+1} \\ M_2 & M_3 & M_4 & \dots & M_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_K & M_{K+1} & M_{K+2} & \dots & M_{2K} \end{pmatrix}$$

a

$$\mathbf{B} = \begin{pmatrix} \bar{p}_1 x_1^{*0} & \bar{p}_1 x_1^{*1} & \dots & \bar{p}_1 x_1^{*K} \\ \bar{p}_2 x_2^{*0} & \bar{p}_2 x_2^{*1} & \dots & \bar{p}_2 x_2^{*K} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{p}_m x_m^{*0} & \bar{p}_m x_m^{*1} & \dots & \bar{p}_m x_m^{*K} \end{pmatrix}.$$

Pomocí této věty můžeme snadno zkonstruovat intervalové odhady (individuální i sdružené) pravděpodobností p_j pro Shannonovu kvazinormu. Demonstrujeme to v následujícím příkladu.

10.3 Příklad

Pozorováním diskrétní náhodné veličiny X nabývající hodnot $x = 1, \dots, 7$ jsme získali statistický soubor o rozsahu $n = 180$. Po jeho rozřídění jsme obdrželi diskrétní empirické rozdělení náhodné veličiny X dané tabulkou 15, kde jsou x_j^* středy tříd a f_j pozorované absolutní četnosti.

Tabulka 15: Četnostní tabulka

x_j^*	1	2	3	4	5	6	7
f_j	15	36	21	15	27	42	24

Stanovme 95% individuální intervalové odhady pravděpodobností p_j splňující minimální Shannonovu kvazinormu za vedlejších podmínek daných prvními obecnými momenty

$$M_0 = \frac{1}{n} \sum_{j=1}^m f_j = 1,$$

$$M_1 = \frac{1}{n} \sum_{j=1}^m f_j x_j^* = 12,75,$$

$$M_2 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*2} = 65,85,$$

$$M_3 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*3} = 375,15,$$

$$M_4 = \frac{1}{n} \sum_{j=1}^m f_j x_j^{*4} = 2251,05.$$

Jedná se o nelineární optimalizační optimalizační úlohu, jejíž existenci a jednoznačnost řešení zaručuje konvexnost účelové funkce a linearita v pravděpodobnostech p_j omezujících momentových podmínek [4]. Pro řešení byl použit klasický analytický přístup pomocí Lagrangeových multiplikátorů. Řešení příslušné nelineární soustavy rovnic a další numerické výpočty byly provedeny v programu Maple. Volme postupně počet momentových omezení $K = 0, 1, 2, 3, 4$. Výsledky pro jednotlivé odhady jsou uvedené v následujícím přehledu, který je doplněn grafickým znázorněním 95% konfidenčních intervalů na obrázcích 28, 29, 30 a 31.

- $K = 0$

Odhad Lagrangeových multiplikátorů a pravděpodobností:

$$\boldsymbol{\lambda} = (0,94591),$$

$$\mathbf{p} = (0,14286; 0,14286; 0,14286; 0,14286; 0,14286; 0,14286; 0,14286).$$

Fischerova informační matice a kovarianční matice vektoru $\boldsymbol{\lambda}$:

$$\mathbf{J} = (1), \quad \text{var } \boldsymbol{\lambda} = \left(\frac{1}{180} \right).$$

Kovarianční matice vektoru \mathbf{p} :

$$\text{var } \mathbf{p} = \begin{pmatrix} 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \\ 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 & 0,1134 \end{pmatrix} \cdot 10^{-3}.$$

95% intervalové odhady pro pravděpodobnosti p_j :

$$P(p_j \in (0,14263; 0,14308)) = 0,95$$

pro všechna $j = 1, 2, \dots, 7$.

- $K = 1$

Odhad Lagrangeových multiplikátorů a pravděpodobností:

$$\boldsymbol{\lambda} = (-1,20458; 0,06270),$$

$$\mathbf{p} = (0,11743; 0,12503; 0,13313; 0,14174; 0,15091; 0,16068; 0,17108).$$

Fischerova informační matice a kovarianční matice vektoru $\boldsymbol{\lambda}$:

$$\mathbf{J} = \begin{pmatrix} 1 & \frac{17}{4} \\ \frac{17}{4} & \frac{439}{20} \end{pmatrix}, \quad \text{var } \boldsymbol{\lambda} = \begin{pmatrix} \frac{439}{13995} & -\frac{17}{2799} \\ -\frac{17}{2799} & \frac{4}{2799} \end{pmatrix}.$$

Kovarianční matice vektoru \mathbf{p} :

$$\mathbf{p} = \begin{pmatrix} 0,2848 & 0,2350 & 0,1776 & 0,1118 & 0,0367 & -0,0485 & -0,1450 \\ 0,2350 & 0,2000 & 0,1594 & 0,1127 & 0,0593 & -0,0014 & -0,0703 \\ 0,1776 & 0,1594 & 0,1380 & 0,1133 & 0,0847 & 0,0520 & 0,0146 \\ 0,1118 & 0,1127 & 0,1133 & 0,1134 & 0,1131 & 0,1123 & 0,1109 \\ 0,0367 & 0,0593 & 0,0847 & 0,1131 & 0,1448 & 0,1802 & 0,2195 \\ -0,0485 & -0,0014 & 0,0520 & 0,1123 & 0,1802 & 0,2564 & 0,3418 \\ -0,1450 & -0,0703 & 0,0146 & 0,1109 & 0,2195 & 0,3418 & 0,4789 \end{pmatrix} \cdot 10^{-3}.$$

95% intervalové odhady pro pravděpodobnosti p_j :

$$P(p_1 \in (0,11688; 0,11799)) = 0,95,$$

$$P(p_2 \in (0,12464; 0,12543)) = 0,95,$$

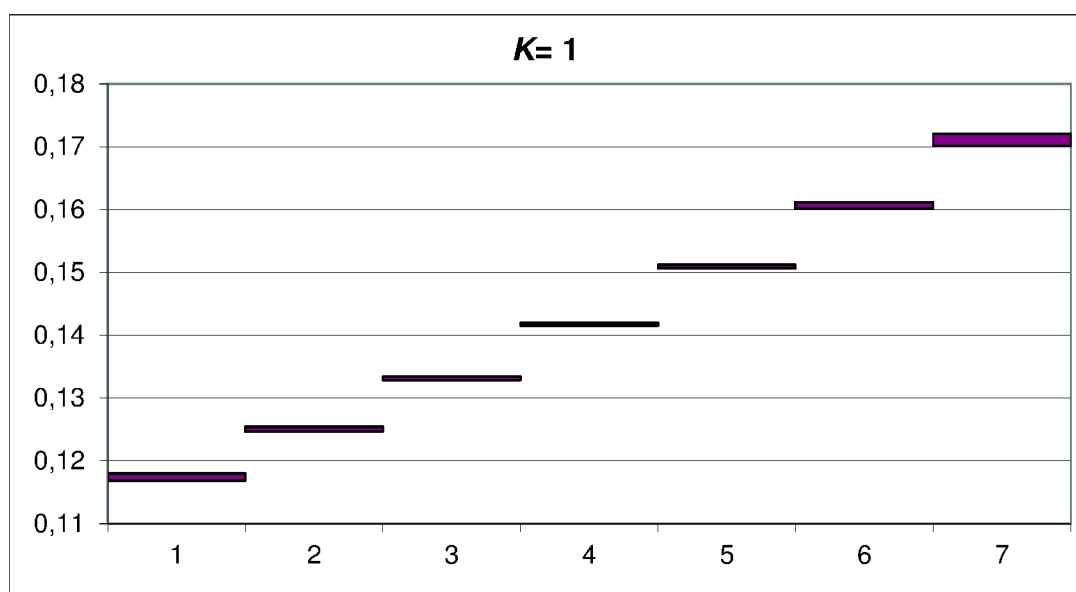
$$P(p_3 \in (0,13285; 0,13340)) = 0,95,$$

$$P(p_4 \in (0,14152; 0,14196)) = 0,95,$$

$$P(p_5 \in (0,15063; 0,15120)) = 0,95,$$

$$P(p_6 \in (0,16018; 0,16118)) = 0,95,$$

$$P(p_7 \in (0,17014; 0,17201)) = 0,95.$$



Obrázek 28: 95% intervaly spolehlivosti pro $k = 1$

- $K = 2$

Odhad Lagrangeových multiplikátorů a pravděpodobností:

$$\boldsymbol{\lambda} = (-1,28410; 0,11356; -0,00621),$$

$$\mathbf{p} = (0,11340; 0,12469; 0,13542; 0,14525; 0,15387; 0,16100; 0,16637).$$

Fischerova informační matice a kovarianční matice vektoru $\boldsymbol{\lambda}$:

$$\mathbf{J} = \begin{pmatrix} 1 & \frac{17}{4} & \frac{439}{20} \\ \frac{17}{4} & \frac{439}{20} & \frac{2501}{20} \\ \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} \end{pmatrix}, \quad \text{var } \boldsymbol{\lambda} = \begin{pmatrix} \frac{2313}{17561} & -\frac{22207}{316098} & \frac{2483}{316098} \\ -\frac{22207}{316098} & \frac{107419}{2528784} & -\frac{4235}{842928} \\ \frac{2483}{316098} & -\frac{4235}{842928} & \frac{1555}{2528784} \end{pmatrix}.$$

Kovarianční matice vektoru \mathbf{p} :

$$\mathbf{p} = \begin{pmatrix} 0,5139 & 0,2474 & 0,0296 & -0,1111 & -0,1489 & -0,0626 & 0,1617 \\ 0,2474 & 0,2007 & 0,1494 & 0,0963 & 0,0446 & -0,0028 & -0,0429 \\ 0,0296 & 0,1494 & 0,2273 & 0,2474 & 0,1958 & 0,0621 & -0,1592 \\ -0,1111 & 0,0963 & 0,2474 & 0,3172 & 0,2835 & 0,1292 & -0,1555 \\ -0,1489 & 0,0446 & 0,1958 & 0,2835 & 0,2885 & 0,1957 & -0,0043 \\ -0,0626 & -0,0028 & 0,0621 & 0,1292 & 0,1957 & 0,2584 & 0,3143 \\ 0,1617 & -0,0429 & -0,1592 & -0,1555 & -0,0043 & 0,3143 & 0,8102 \end{pmatrix} \cdot 10^{-3}.$$

95% intervalové odhady pro pravděpodobnosti p_j :

$$P(p_1 \in (0,11239; 0,11441)) = 0,95,$$

$$P(p_2 \in (0,12430; 0,12508)) = 0,95,$$

$$P(p_3 \in (0,13497; 0,13586)) = 0,95,$$

$$P(p_4 \in (0,14463; 0,14587)) = 0,95,$$

$$P(p_5 \in (0,15331; 0,15444)) = 0,95,$$

$$P(p_6 \in (0,16049; 0,16150)) = 0,95,$$

$$P(p_7 \in (0,16479; 0,16796)) = 0,95.$$

- $K = 3$

Odhad Lagrangeových multiplikátorů a pravděpodobností

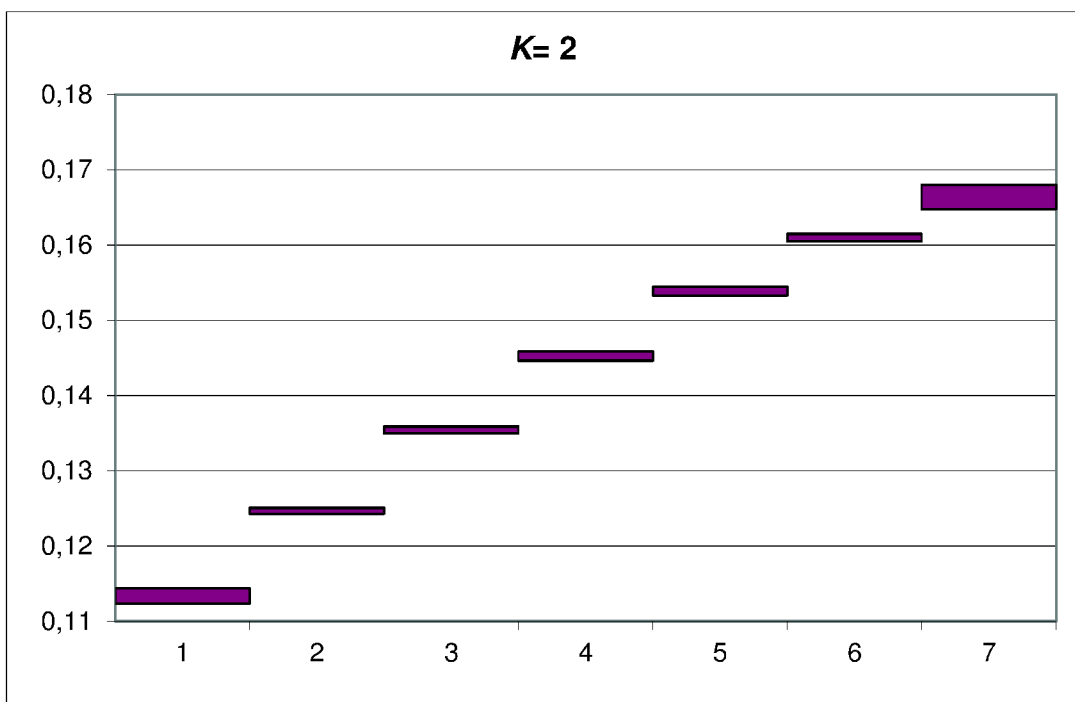
$$\boldsymbol{\lambda} = (-1,17237; 0,01058; -0,02944; -0,293),$$

$$\mathbf{p} = (0,11574; 0,12255; 0,13288; 0,14496; 0,15634; 0,16378; 0,16376).$$

Fischerova informační matice a kovarianční matice vektoru $\boldsymbol{\lambda}$:

$$\mathbf{J} = \begin{pmatrix} 1 & \frac{17}{4} & \frac{439}{20} & \frac{2501}{20} \\ \frac{17}{4} & \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} \\ \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} & \frac{18577}{4} \\ \frac{2501}{20} & \frac{15007}{20} & \frac{18577}{4} & \frac{587119}{20} \end{pmatrix},$$

$$\text{var } \boldsymbol{\lambda} = \begin{pmatrix} \frac{247543}{529011} & -\frac{73592}{176337} & \frac{166006}{1587033} & -\frac{12497}{1587033} \\ -\frac{73592}{176337} & \frac{942283}{2351160} & -\frac{554909}{5290110} & \frac{172009}{21160440} \\ \frac{166006}{1587033} & -\frac{554909}{5290110} & \frac{5418727}{190443960} & -\frac{431513}{190443960} \\ -\frac{12497}{1587033} & \frac{172009}{21160440} & -\frac{431513}{190443960} & \frac{17561}{95221980} \end{pmatrix}.$$



Obrázek 29: 95% intervaly spolehlivosti pro $k = 2$

Kovarianční matice vektoru \mathbf{p} :

$$\mathbf{p} = \begin{pmatrix} 0,7777 & 0,1365 & -0,1377 & -0,1449 & -0,0068 & 0,0921 & -0,0738 \\ 0,1365 & 0,2453 & 0,2212 & 0,1091 & -0,0235 & -0,0751 & 0,0674 \\ -0,1377 & 0,2212 & 0,3345 & 0,2641 & 0,0933 & -0,0465 & 0,0094 \\ -0,1449 & 0,1091 & 0,2641 & 0,3200 & 0,2681 & 0,1107 & -0,1219 \\ -0,0068 & -0,0235 & 0,0933 & 0,2681 & 0,3877 & 0,2979 & -0,1481 \\ 0,0921 & -0,0751 & -0,0465 & 0,1107 & 0,2979 & 0,3692 & 0,1617 \\ -0,0738 & 0,0674 & 0,0094 & -0,1219 & -0,1481 & 0,1617 & 1,0152 \end{pmatrix} \cdot 10^{-3}.$$

95% intervalové odhady pro pravděpodobnosti p_j :

$$P(p_1 \in (0,11421; 0,11726)) = 0,95,$$

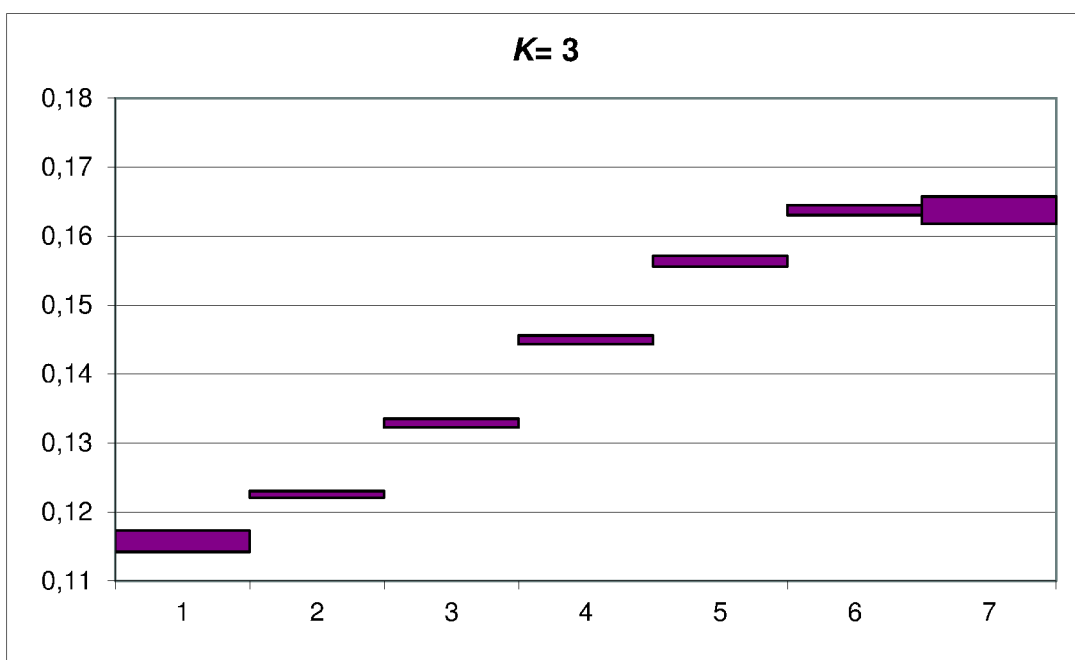
$$P(p_2 \in (0,12207; 0,12303)) = 0,95,$$

$$P(p_3 \in (0,13222; 0,13353)) = 0,95,$$

$$P(p_4 \in (0,14433; 0,14559)) = 0,95,$$

$$P(p_5 \in (0,15558; 0,15710)) = 0,95,$$

$$P(p_6 \in (0,16306; 0,16450)) = 0,95,$$



Obrázek 30: 95% intervaly spolehlivosti pro $k = 3$

$$P(p_7 \in (0,16177; 0,16575)) = 0,95.$$

- $K = 4$

Odhad Lagrangeových multiplikátorů a pravděpodobností

$$\lambda = (-6,01720; 7,47108; -3,52618; 0,64806; -0,04032),$$

$$\mathbf{p} = (0,08505; 0,19390; 0,12147; 0,09085; 0,13391; 0,24395; 0,13087).$$

Fischerova informační matice a kovarianční matice vektoru λ :

$$\mathbf{J} = \begin{pmatrix} 1 & \frac{17}{4} & \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} \\ \frac{17}{4} & \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} & \frac{18577}{4} \\ \frac{439}{20} & \frac{2501}{20} & \frac{15007}{20} & \frac{18577}{4} & \frac{587119}{20} \\ \frac{2501}{20} & \frac{15007}{20} & \frac{18577}{4} & \frac{587119}{20} & \frac{3769781}{20} \\ \frac{15007}{20} & \frac{18577}{4} & \frac{587119}{20} & \frac{3769781}{20} & \frac{24508447}{20} \end{pmatrix},$$

$$\text{var } \lambda = \begin{pmatrix} \frac{2516161}{1530387} & -\frac{2252755}{1020258} & \frac{17578729}{18364644} & -\frac{754283}{4591161} & \frac{59387}{6121548} \\ -\frac{2252755}{1020258} & \frac{15951643}{5101290} & -\frac{57254041}{40810320} & \frac{111667}{453448} & -\frac{602753}{40810320} \\ \frac{17578729}{18364644} & -\frac{57254041}{40810320} & \frac{1424704813}{2203757280} & -\frac{63715379}{550939320} & \frac{5165819}{734585760} \\ -\frac{754283}{4591161} & \frac{111667}{453448} & -\frac{63715379}{550939320} & \frac{5780801}{275469660} & -\frac{118463}{91823220} \\ \frac{59387}{6121548} & -\frac{602753}{40810320} & \frac{5165819}{734585760} & -\frac{118463}{91823220} & \frac{2177}{27206880} \end{pmatrix}.$$

Kovarianční matice vektoru \mathbf{p} :

$$\mathbf{p} = \begin{pmatrix} 0,4713 & 0,0422 & -0,0603 & 0,0102 & 0,0392 & -0,0397 & 0,0095 \\ 0,0422 & 0,8781 & 0,2469 & -0,0661 & -0,1305 & 0,1413 & -0,0346 \\ -0,0603 & 0,2469 & 0,2997 & 0,1996 & 0,1003 & -0,1514 & 0,0400 \\ 0,0101 & -0,0661 & 0,1996 & 0,2407 & 0,2091 & -0,1068 & 0,0181 \\ 0,0392 & -0,1305 & 0,1003 & 0,2091 & 0,3213 & 0,2611 & -0,0566 \\ -0,0397 & 0,1413 & -0,1514 & -0,1068 & 0,2611 & 1,2031 & 0,0478 \\ 0,0095 & -0,0346 & 0,0400 & 0,0181 & -0,0566 & 0,0478 & 0,7028 \end{pmatrix} \cdot 10^{-3}.$$

95% intervalové odhady pro pravděpodobnosti p_j :

$$P(p_1 \in (0,08412; 0,08597)) = 0,95,$$

$$P(p_2 \in (0,19218; 0,19562)) = 0,95,$$

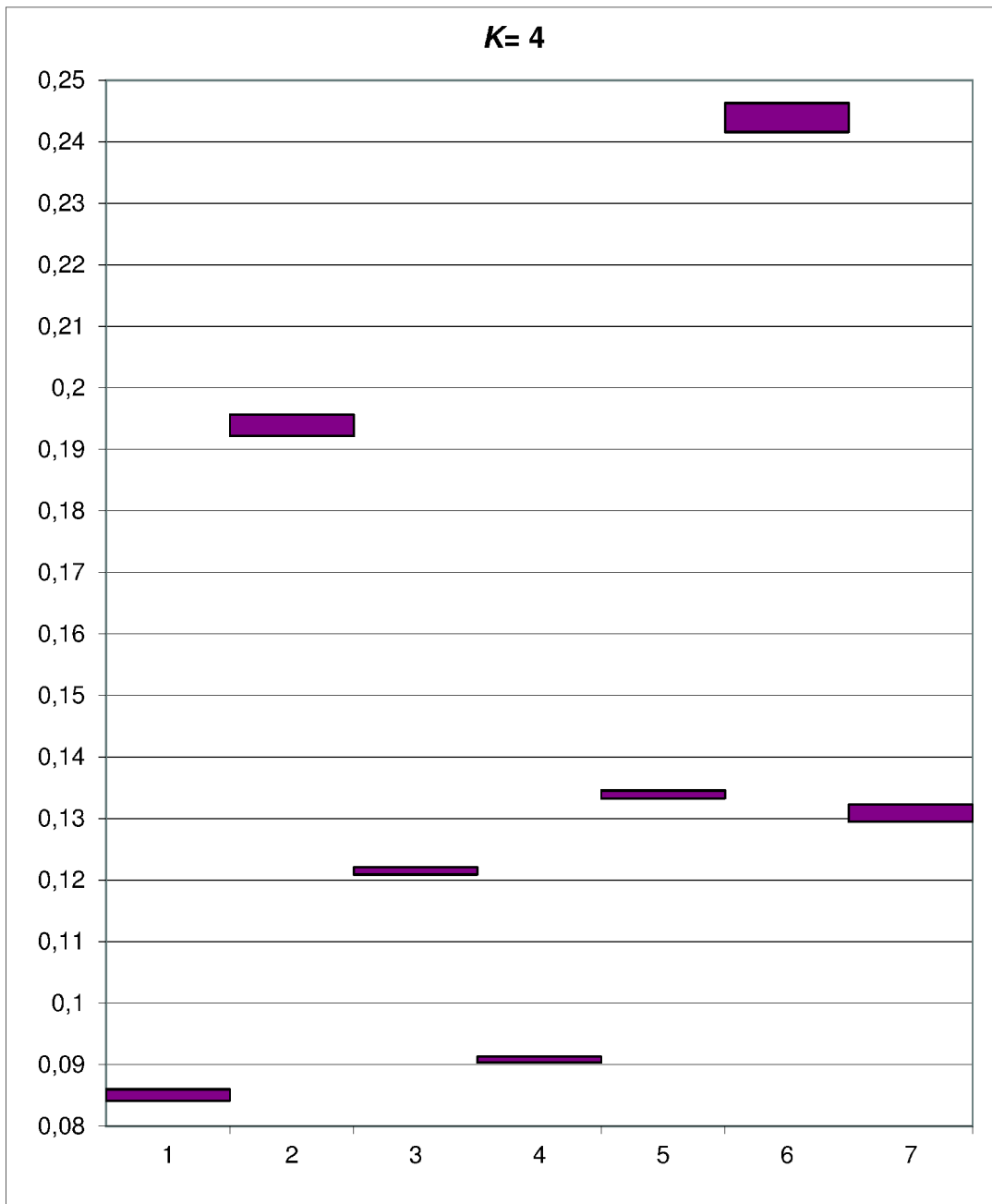
$$P(p_3 \in (0,12089; 0,12206)) = 0,95,$$

$$P(p_4 \in (0,09038; 0,09132)) = 0,95,$$

$$P(p_5 \in (0,13328; 0,13454)) = 0,95,$$

$$P(p_6 \in (0,24159; 0,24630)) = 0,95,$$

$$P(p_7 \in (0,12949; 0,13225)) = 0,95.$$



Obrázek 31: 95% intervaly spolehlivosti pro $k = 4$

11 Intervalové odhady pravděpodobností p_j pro Shannonovu kvazinormu a obecné podmínky

11.1 Odvození rozdělení pro vektor Lagrangeových multiplikátorů

Jak již bylo popsáno v kapitole 10, jsou odhady parametrů λ_k pomocí Shannonovy kvazinormy maximálně věrohodné, odhad vektoru $\boldsymbol{\lambda} = (\lambda_0, \dots, \lambda_K)$ má asymptoticky $(K + 1)$ -rozměrné normální rozdělení $N(\boldsymbol{\lambda}_0, n^{-1} [\mathbf{J}(\boldsymbol{\lambda}_0)]^{-1})$, kde $\boldsymbol{\lambda}_0$ je teoretický vektor parametrů rozdělení náhodné veličiny X a $\mathbf{J}(\boldsymbol{\lambda}_0)$ je Fisherova informační matice.

Speciálně odhad λ_k teoretické hodnoty jednoho pevně zvoleného multiplikátoru λ_{0k} , $k = 0, \dots, K$, má asymptoticky normální rozdělení $N\left(\lambda_{0k}, \frac{1}{nJ(\lambda_{0k})}\right)$. Označme $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$ odhad vektoru Lagrangeových multiplikátorů, který dostaneme řešením minimalizační úlohy pro konkrétní pozorování. Pro výpočet intervalového odhadu λ_{0k} pomocí tohoto normálního rozdělení klademe $\lambda_{0k} = \bar{\lambda}_k$. Protože Fisherova míra informace

$$J(\lambda_{0k}) = E \left(\left[\frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0k}} \right]^2 \right),$$

kde E značí střední hodnotu a pro $x = x_j^*$ je

$$p(x_j^*, \boldsymbol{\lambda}_0) = \exp \left(-1 - \sum_{k=0}^K \lambda_k a_{kj} \right),$$

klademe $J(\lambda_{0k}) = J(\bar{\lambda}_k)$. Po úpravách je

$$J(\bar{\lambda}_k) = \sum_{j=1}^m p_j a_{kj}^2 \approx \sum_{j=1}^m \frac{f_j}{n} a_{kj}^2,$$

kde pravděpodobnosti p_j jsou určeny pomocí minimalizace Shannonovy kvazinormy.

Obdobně pro prvek Fisherovy informační matice dostáváme J_{ts} dostáváme pro $t, s = 0, \dots, K$

$$J_{ts} = E \left(\frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0t}} \cdot \frac{\partial \ln p(x, \boldsymbol{\lambda}_0)}{\partial \lambda_{0s}} \right) = E(a_{tj} a_{sj}).$$

Použijeme-li jako apriorní odhad $p_j \approx \frac{f_j}{n}$ dostáváme

$$J_{ts} = \frac{f_j}{n} a_{tj} a_{sj}.$$

Předchozí úvahy můžeme shrnout do věty 11.1.

Věta 11.1 Vektor Lagrangeových multiplikátorů $\boldsymbol{\lambda}$ má v případě obecných lineárních podmínek asymptoticky $(K + 1)$ -rozměrné normální rozdělení $N(\bar{\boldsymbol{\lambda}}, [n\mathbf{J}]^{-1})$, kde vektor středních hodnot $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$ dostaneme řešením minimalizační úlohy a

$$\mathbf{J} = \begin{pmatrix} \frac{f_j}{n} & \frac{f_j}{n} a_{1j} & \cdots & \frac{f_j}{n} a_{Kj} \\ \frac{f_j}{n} a_{1j} & \frac{f_j}{n} a_{1j}^2 & \cdots & \frac{f_j}{n} a_{1j} a_{Kj} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{f_j}{n} a_{Kj} & \frac{f_j}{n} a_{Kj} a_{2j} & \cdots & \frac{f_j}{n} a_{Kj}^2 \end{pmatrix}.$$

11.2 Odvození rozdělení pro vektor pravděpodobností

Nyní se budeme zabývat intervalovými odhady pravděpodobností p_j . K tomu je potřeba zjistit tvar rozdělení vektoru $\mathbf{p} = (p_1, \dots, p_K)$. Protože \mathbf{p} je závislé na $\boldsymbol{\lambda}$, které má asymptoticky normální rozdělení, je zřejmé, že rozdělení vektoru \mathbf{p} je opět asymptoticky normální. Závislost má tvar

$$p_j(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \lambda_k a_{kj}\right), \quad j = 1, \dots, m.$$

K určení jeho parametrů tuto závislost linearizujeme na tvar $\mathbf{p} = \mathbf{a} + \mathbf{B}\boldsymbol{\lambda}$. Sestavíme tedy Taylorův polynom prvního stupně k této transformaci. Jako střed polynomu zvolíme bod $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$. Jeho dosazením do transformačních rovnic dostaneme bod

$$\bar{p}_j = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k a_{kj}\right), \quad j = 1, \dots, m,$$

který je řešením minimalizační úlohy pro Shannonovu kvazinormu, V okolí tohoto bodu nás rozdělení \mathbf{p} bude nejvíce zajímat. Vektor \mathbf{a} je zřejmě roven středu Taylorova polynomu a budeme jej značit $\bar{\mathbf{p}} = (\bar{p}_1, \dots, \bar{p}_m)$. Taylorův polynom má tvar

$$T(p_j)(\boldsymbol{\lambda}) = p_j(\bar{\boldsymbol{\lambda}}) + \sum_{i=0}^K \frac{\partial p_j(\bar{\lambda}_i)}{\partial \lambda_i} (\lambda_i - \bar{\lambda}_i).$$

Dosazením za

$$\frac{\partial p_j(\bar{\lambda}_i)}{\partial \lambda_i} = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k a_{kj}\right) a_{ij}, \quad j = 1, \dots, m$$

dostáváme

$$T(p_j)(\boldsymbol{\lambda}) = \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k a_{kj}\right) + \sum_{i=0}^K a_{ij} \exp\left(-1 - \sum_{k=0}^K \bar{\lambda}_k a_{kj}\right) (\lambda_i - \bar{\lambda}_i).$$

Po úpravách je

$$T(p_j)(\boldsymbol{\lambda}) = \bar{p}_j \left(1 - \sum_{i=0}^K \bar{\lambda}_i a_{ij}\right) + \bar{p}_j \sum_{i=0}^K \lambda_i a_{ij}.$$

Můžeme tedy psát

$$\mathbf{p} \approx \mathbf{a} + \mathbf{B}\boldsymbol{\lambda},$$

kde

$$\mathbf{a} = \begin{pmatrix} \bar{p}_1 \left(1 - \sum_{i=0}^K \bar{\lambda}_i a_{i1} \right) \\ \vdots \\ \bar{p}_m \left(1 - \sum_{i=0}^K \bar{\lambda}_i a_{im} \right) \end{pmatrix}$$

a

$$\mathbf{B} = \begin{pmatrix} \bar{p}_1 & \bar{p}_1 a_{11} & \dots & \bar{p}_1 a_{K1} \\ \bar{p}_2 & \bar{p}_2 a_{12} & \dots & \bar{p}_2 a_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{p}_m & \bar{p}_m a_{1m} & \dots & \bar{p}_m a_{Km} \end{pmatrix}.$$

Využijeme nyní věty o lineární transformaci normálně rozděleného náhodného vektoru a dostaneme tak parametry normálního rozdělení pro vektor \mathbf{p} [3]. Je zřejmé, že střední hodnota vektoru $E(\mathbf{p}) = \mathbf{a} + \mathbf{B}\boldsymbol{\lambda} = \mathbf{p} = (\bar{p}_1, \dots, \bar{p}_m)$. Tyto úvahy můžeme shrnout do věty 11.2.

Věta 11.2 *Vektor pravděpodobností \mathbf{p} má asymptoticky m -rozměrné normální rozdělení $N(\bar{\mathbf{p}}, \mathbf{B} [n\mathbf{J}]^{-1} \mathbf{B}^T)$, kde vektor středních hodnot $\bar{\boldsymbol{\lambda}} = (\bar{\lambda}_0, \dots, \bar{\lambda}_K)$ a vektor $\bar{\mathbf{p}} = (\bar{p}_1, \dots, \bar{p}_m)$ dostaneme řešením minimalizační úlohy,*

$$\mathbf{J} = \begin{pmatrix} \frac{f_j}{n} & \frac{f_j}{n} a_{1j} & \dots & \frac{f_j}{n} a_{Kj} \\ \frac{f_j}{n} a_{1j} & \frac{f_j}{n} a_{1j}^2 & \dots & \frac{f_j}{n} a_{1j} a_{Kj} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{f_j}{n} a_{Kj} & \frac{f_j}{n} a_{Kj} a_{2j} & \dots & \frac{f_j}{n} a_{Kj}^2 \end{pmatrix}$$

a

$$\mathbf{B} = \begin{pmatrix} \bar{p}_1 & \bar{p}_1 a_{11} & \dots & \bar{p}_1 a_{K1} \\ \bar{p}_2 & \bar{p}_2 a_{12} & \dots & \bar{p}_2 a_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{p}_m & \bar{p}_m a_{1m} & \dots & \bar{p}_m a_{Km} \end{pmatrix}.$$

Pomocí této věty můžeme snadno zkonstruovat intervalové odhady (individuální i sdružené) pravděpodobností p_j pro Shannonovu kvazinormu. Demonstrujeme to v následujícím příkladu.

11.3 Příklad

Navažme nyní na příklad z kapitoly 10 a řešme úlohu minimalizace Shannonovy kvazinormy za podmínek zaručující rovnost "aritmetických" resp. "geometrických" momentů pro pozorování a odhad až do řádu $k = 0, 1, \dots, m - 1$. Empirické četnosti jsou v tabulce 16.

Tabulka 16: Tabulka četností

x_i^*	0,8	1	1,2	1,4	1,6
f_i	17	35	27	14	7

Matematický model této úlohy vypadá

$$\min \sum_{i=0}^K p_j \ln p_j$$

za podmínek

$$\sum_{j=1}^m p_j x_j^{*k} = M_k, \quad \text{resp.} \quad \sum_{j=1}^m p_j \ln^k x_j^* = G_k, \quad k = 0, \dots, K,$$

kde momenty M_k a G_k vychází

$$M_1 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^* = 1,118,$$

$$G_1 = \frac{1}{100} \sum_{i=1}^5 f_i \ln(x_i^*) = 0,0913,$$

$$M_2 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*2} = 1,301,$$

$$G_2 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^2(x_i^*) = 0,0488,$$

$$M_3 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*3} = 1,574,$$

$$G_3 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^3(x_i^*) = 0,0123,$$

$$M_4 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*4} = 1,976,$$

$$G_4 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^4(x_i^*) = 0,00593,$$

$$M_5 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*5} = 2,565,$$

$$G_5 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^5(x_i^*) = 0,00217,$$

$$M_6 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*6} = 3,429,$$

$$G_6 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^6(x_i^*) = 0,000989,$$

$$M_7 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*7} = 4,708,$$

$$G_7 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^7(x_i^*) = 0,000420,$$

$$M_8 = \frac{1}{100} \sum_{i=1}^5 f_i x_i^{*8} = 6,612,$$

$$G_8 = \frac{1}{100} \sum_{i=1}^5 f_i \ln^8(x_i^*) = 0,000191.$$

Lagrangeova funkce

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = p_j \ln p_j + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j x_j^{*k} - M_k \right),$$

resp.

$$\Lambda(\mathbf{p}, \boldsymbol{\lambda}) = p_j \ln p_j + \sum_{k=0}^K \lambda_k \left(\sum_{j=1}^m p_j \ln^k x_j^* - G_k \right).$$

Bodové odhady Lagrangeových multiplikátorů $\bar{\lambda}_k$ jsme dostali řešením soustavy rovnic

$$\sum_{j=1}^m \exp \left(-1 + \sum_{k=0}^K \bar{\lambda}_k x_j^{*k} \right) x_j^{*\nu} = M_\nu, \quad \nu = 1, \dots, K,$$

resp.

$$\sum_{j=1}^m \exp \left(-1 + \sum_{k=0}^K \bar{\lambda}_k \ln^k x_j^* \right) \ln^\nu x_j^* = G_\nu, \quad \nu = 1, \dots, K$$

v programu Maple a bodové odhady pravděpodobností \bar{p}_j dostaneme jako

$$\bar{p}_j = \exp \left(-1 - \sum_{k=0}^K \lambda_k x_j^{*k} \right),$$

resp.

$$\bar{p}_j = \exp \left(-1 - \sum_{k=0}^K \lambda_k x_j^{*k} \right), \quad j = 1, \dots, m.$$

Určeme dále 95% intervalové odhady pravděpodobností p_j .
Fischerova informační matice vychází

$$\mathbf{J} = \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_K \\ M_1 & M_2 & M_3 & \dots & M_{K+1} \\ M_2 & M_3 & M_4 & \dots & M_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M_K & M_{K+1} & M_{K+2} & \dots & M_{2K} \end{pmatrix}$$

pro "aritmetické" momenty, resp.

$$\mathbf{J} = \begin{pmatrix} G_0 & G_1 & G_2 & \dots & G_K \\ G_1 & G_2 & G_3 & \dots & G_{K+1} \\ G_2 & G_3 & G_4 & \dots & G_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ G_K & G_{K+1} & G_{K+2} & \dots & G_{2K} \end{pmatrix}$$

pro "geometrické" momenty.

Další numerické výpočty byly provedeny v programu Maple. Jedná se konkrétně o kovarianční matici pro vektor Lagrangeových multiplikátorů $\text{var } \boldsymbol{\lambda}$, matici lineární transformace \mathbf{B} , kovarianční matici pro vektor pravděpodobností $\text{var } \mathbf{p}$. Oproti předcházející kapitole je uveden pouze přehled výsledků, konkrétně odhad střední hodnoty \bar{p}_j , směrodatné odchylky, dolní d_j a horní h_j meze 95% konfidenčního intervalu pro jednotlivé pravděpodobnosti p_j . Tyto výsledky jsou pro aritmetické a geometrické momentové podmínky do řádu k , $k = 0, 1, \dots, m - 1$ jsou uvedeny v následujících tabulkách 17 a 18.

Tabulka 17: Aritmetické momenty

x_j^*		0,8	1,0	1,2	1,4	1,6
f_j		17	35	27	14	7
$K = 0$	\bar{p}_j	0,20000	0,20000	0,20000	0,20000	0,20000
	σ_j	0,02000	0,02000	0,02000	0,02000	0,02000
	d_j	0,19922	0,19922	0,19922	0,19922	0,19922
	h_j	0,20078	0,20078	0,20078	0,20078	0,20078
$K = 1$	\bar{p}_j	0,29085	0,23603	0,19154	0,15544	0,12614
	σ_j	0,05014	0,02662	0,02037	0,02483	0,02967
	d_j	0,28592	0,23464	0,19073	0,15423	0,12442
	h_j	0,29578	0,23742	0,19236	0,15665	0,12787
$K = 2$	\bar{p}_j	0,19055	0,30297	0,28779	0,16332	0,05538
	σ_j	0,04240	0,03671	0,04029	0,02609	0,01823
	d_j	0,18703	0,30032	0,28460	0,16199	0,05472
	h_j	0,19408	0,30561	0,29097	0,16465	0,05603
$K = 3$	\bar{p}_j	0,17071	0,34715	0,27427	0,13715	0,07071
	σ_j	0,04106	0,05441	0,03986	0,02886	0,02581
	d_j	0,16741	0,34135	0,27116	0,13552	0,06941
	h_j	0,17402	0,35295	0,27739	0,13878	0,07202
$K = 4$	\bar{p}_j	0,17000	0,35000	0,27000	0,14000	0,07000
	σ_j	0,04125	0,05921	0,05206	0,03749	0,02653
	d_j	0,16667	0,34313	0,26469	0,13724	0,06862
	h_j	0,17333	0,35687	0,27531	0,14276	0,07138

Tabulka 18: Geometrické momenty

x_j^*		0,8	1,0	1,2	1,4	1,6
f_j		17	35	27	14	7
$K = 0$	\bar{p}_j	0,20000	0,20000	0,20000	0,20000	0,20000
	σ_j	0,02000	0,02000	0,02000	0,02000	0,02000
	d_j	0,19922	0,19922	0,19922	0,19922	0,19922
	h_j	0,20078	0,20078	0,20078	0,20078	0,20078
$K = 1$	\bar{p}_j	0,28373	0,22636	0,18822	0,16103	0,14067
	σ_j	0,05267	0,02486	0,02066	0,02539	0,03000
	d_j	0,27829	0,22515	0,18738	0,15976	0,13890
	h_j	0,28916	0,22758	0,18905	0,16229	0,14243
$K = 2$	\bar{p}_j	0,17647	0,33131	0,28007	0,15051	0,06164
	σ_j	0,04096	0,04296	0,03712	0,02438	0,01884
	d_j	0,17318	0,32769	0,27737	0,14934	0,06095
	h_j	0,17976	0,33493	0,28277	0,15167	0,06234
$K = 3$	\bar{p}_j	0,16977	0,35128	0,26750	0,14212	0,06934
	σ_j	0,04107	0,05726	0,04151	0,02732	0,02497
	d_j	0,16646	0,34485	0,26412	0,14065	0,06812
	h_j	0,17307	0,35770	0,27087	0,14358	0,07057
$K = 4$	\bar{p}_j	0,17000	0,35000	0,27000	0,14000	0,07000
	σ_j	0,04123	0,05916	0,05196	0,03742	0,02646
	d_j	0,16667	0,34314	0,26471	0,13726	0,06863
	h_j	0,17333	0,35686	0,27529	0,14274	0,07137

Závěr

Obvykle používané metody pro odhad rozdělení pravděpodobnosti a jejich softwarová podpora jsou obvykle zaměřeny pouze na nejběžnější typy rozdělení. To však může být v některých oblastech použití omezující. Popsaný přístup nepreferuje žádný typ rozdělení. Vyžaduje pouze znalost odhadů vybraných charakteristik rozdělení (např. střední hodnoty, rozptylu, příp. vyšších momentů) a při jejich respektování hledá nejvíce neurčité rozdělení pravděpodobnosti \mathbf{p} na příslušném pravděpodobnostním prostoru. Díky této základní myšlence se vžilo označení "pesimistické odhady". Je rozumné hledat odhad rozdělení pravděpodobnosti tímto pesimistickým způsobem, neboť tím nebudeme do systému přidávat navíc nějaké informace, o kterých nevíme, a naše odhadnuté rozdělení bude respektovat jen známé informace. Efektivnost takových odhadů také dokládají dále citované publikace při odhadech neznámých vícemodálních rozdělení pravděpodobnosti.

Tato práce v jistém smyslu navazuje na publikace [5], [6], [7], [8], [9], [10], [11], [13], [19], [20], [22], [24], [34], [35] a [36], které otvírají a prezentují problematiku kvazinorem pro odhady rozdělení, a také na [21], kde je mj. popsán "entropický" základ přístupu k odhadu diskrétního rozdělení pravděpodobnosti. V [21] je problematika rozšířena o obdobný odhad náhodného vektoru ve 2D a práce obsahuje zajímavé příklady využití. Jejím přínosem je také implementace této metodiky do softwaru Estimator, jehož jádrem je optimalizační software GAMS a který je příjemným uživatelským prostředím pro praktické využití odhadu rozdělení pomocí kvazinorem. Software Estimator a také řešič z Excelu byl použit v této dizertační práci.

Tato práce rozšiřuje řešení úlohy odhadu rozdělení pravděpodobnosti pozorované diskrétní náhodné veličiny jak v oblasti teoretické, tak i aplikační. V úvodních kapitolách zavádí pojem duální f-divergence a kvazinorma a výčet nejpoužívanějších kvazinorem rozšiřuje o jejich duální protějšky. Dalším přínosem práce je zobecnění momentových podmínek na obecné lineární podmínky, které mají širší možnost použití. Jak je ukázáno v příkladě, hodí se např. pro data, u kterých je smysluplnější pracovat s geometrickým průměrem oproti aritmetickému.

Co se týče aplikační oblasti, práce podrobně zkoumá aproximace významných rozdělení pravděpodobnosti pomocí kvazinorem. Srovnává vhodnost použití tří nejvýznamnějších kvazinorem (Hellingerova, Pearsonova a Shannonova) pro aproximace čtyř v praxi důležitých rozdělení (binomické, Poissonovo, normální a Weibullovo) s různými parametry. Pro většinu z nich se dá souhrnně říct, že nejlépe dopadají odhady použitím Shannonovy kvazinormy, naopak nejhůře pro Pearsonovu kvazinormu. Tato část také ukazuje, že metodika primárně vytvořená pro diskrétní rozdělení je dobře použitelná i pro diskretizovaná spojitá rozdělení, což nabízí široká uplatnění např. ve spolehlivosti.

Možná největším přínosem práce je konstrukce intervalových odhadů pravděpodobností p_i pro Shannonovu kvazinormu. Je založena na maximální věrohodnosti odhadů Lagrangeových multiplikátorů λ_i příslušné optimalizační úlohy a tedy asymptotické normalitě rozdělení vektoru $\boldsymbol{\lambda}$. Jelikož vektor \mathbf{p} je nelineární funkcí vektoru $\boldsymbol{\lambda}$, je předpokládána normalita vektoru \mathbf{p} pouze přibližná. Pro větší rozsahy souborů je však tato nepřesnost zanedbatelná a jde o odhady asymptotické. V této části práce jsou také příklady výpočtu intervalových odhadů pravděpodobností p_i pro konkrétní datové

soubory. Práce se zabývá pouze individuálními konfidenčními intervaly pravděpodobností p_i a dalším možným rozšířením jsou sdružené intervalové odhady.

Další výzkum v oblasti aplikací kvazinorem je prováděn v pracích [42], [43], [44], [46], [48], [49]. Tyto publikace jdou trochu jiným směrem především v tom, že se zabývají odhadem kategoriálních veličin a odpadá problém minimalizace kvazinormy s vedlejšími, např. momentovými podmínkami. Je zde používána tzv. gradientní metoda, která vede na soustavu diferenciálních rovnic, a získané odhady mají charakter diskrétních jádrových odhadů.

Dalším směrem, kterým může jít výzkum v této oblasti fitování rozdělení pravděpodobnosti, je využití metod bootstrapu [47]. Jedná se o odhady pomocí kvazinorem z náhodných výběrů s opakováním ze souboru pozorovaných dat a nabízí široké možnosti především v aplikačních aspektech tohoto přístupu. Jako teoreticky zajímavá problematika se jeví souvislost v práci použitých kvazinorem se statistikami, které mají asymptoticky chí-kvadrát rozdělení [33].

Literatura

- [1] Vajda, I. *Theory of statistical inference and information*. London: Kluwer Academic Press, 1989, ISBN 9027727813.
- [2] Anděl, J. *Statistické metody*. Praha: Matfyzpress, 2007, ISBN 978-80-7378-003-6.
- [3] Anděl, J. *Matematická statistika*, Praha, Bratislava: SNTL -ALFA, 1978.
- [4] Bazaraa, M. - S., Sherali, H., D., Shetty, C., M. *Nonlinear programming: Theory and Algorithms* USA: Wiley-Interscience, 2006, ISBN 978-0-471-48600-8.
- [5] Karpíšek, Z. *Statistical Properties of Discrete Probability Distributions with Maximum Entropy*. Folia Fac. Sci. Nat. Univ. Masarykianae Brunensis, Mathematica 9, Brno, 2001, pp. 21-32, ISBN 80-210-2544-1.
- [6] Karpíšek, Z. - Jurák, P. *Modeling of Probability Distribution with Maximum Entropy*. MENDEL '01. 7th International Conference on Soft Computing. Brno, 2001, pp. 232-239, ISBN 80-214-1894-X.
- [7] Karpíšek, Z. - Jurák, P. *Estimate of Discrete Probability Distribution by Means of Hellinger Distance*. MENDEL'02. 8th International Conference on Soft Computing. Brno, 2002, pp. 301-306, ISBN 80-214-2135-5.
- [8] Karpíšek, Z. - Jurák, P. *Odhady diskrétních rozdělení pravděpodobnosti pomocí kvazinorem*. Sborník konference Mezinárodní matematický workshop - Brno 2002. FAST VUT, Brno 21.11.2002, p. 7, CD ROM, ISBN 80-86433-16-1.
- [9] Karpíšek, Z. - Jurák, P. *Netradiční metody fitování diskrétních rozdělení pravděpodobnosti*. Sborník 7. konference Inteligentní systémy pro praxi. Seč u Chrudimi 18. - 19. 2. 2003, p 6. ISBN 80-239-0201-6.
- [10] JURÁK, P. - Karpíšek, Z. *Applications of the Two-Dimensional Hellinger and Shannon Quasi-Norm*. Proceedings of ICSC 2004. Kunovice 29. - 30. 1. 2004, pp. 285-290, ISBN 80-7314-025-X.
- [11] JURÁK, P. - Karpíšek, Z. *Hellingerova a Shannonova kvazinorma v dvourozměrném prostoru*. 3rd International Conference APLIMAT 2004 (part II). Bratislava 2004, pp. 545-552, ISBN 80-227-1995-1.
- [12] Karpíšek, Z. - Šácha, J. *Nepříliš známý test shody diskrétního rozdělení pravděpodobnosti*. Sborník přednášek 8. konference Inteligentní systémy pro praxi. Seč u Chrudimi 2004, p. 2 anotace a p. 5 CD ROM, ISBN 80-239-3680-8.
- [13] Jurák, P. - Karpíšek, Z. *Hellinger Quasi-norm and Shannon Quasi-norm in N-dimensional space*. MENDEL '04. 10th International Conference on Soft Computing. Brno, 2004, pp. 210-215, ISBN 80-214-2676-4.

- [14] Karpíšek, Z. - Sadovský, Z. - Šácha, J. *Pitman - Hellinger Test of Fit*. 4th International Conference APLIMAT 2005 (part II). Bratislava, 2005, pp. 471- 478, ISBN 80-969264-2-X.
- [15] Karpíšek - Z., Sadovský, Z. *Fitování diskrétních rozdělení pravděpodobnosti*. Celostátní seminář Analýza dat 2005/II. Lázně Bohdaneč 2005, pp. 43-52, ISBN 80-239-6552-2.
- [16] Karpíšek, Z. - Jurák, P. *Estimate of Discrete Probability Distribution by Means of Pearson Quasi-norm*. MENDEL '05. 11th International Conference on Soft Computing. Brno, 2005, pp. 202-206, ISBN 80-214-2961-5.
- [17] Gray, R. M. *Entropy and information theory*. Berlin: Springer-Verlag, 1990.
- [18] Jurák, P. *Statistické vlastnosti diskrétních rozdělení pravděpodobnosti s maximální entropií*. Diplomová práce, FSI VUT v Brně, Brno, 2001.
- [19] Karpíšek, Z. - Sadovský, Z. - Jurák, P. *Pearson Quasi-norm for Probability Distribution Goodness of Fit*. 5th International Conference APLIMAT 2006. Bratislava, 7. - 10. 2. 2006, pp. 545-551, ISBN 80-967305-7-6.
- [20] Karpíšek, Z. *F-divergence a kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Analýza dat 2006/I - Sborník přednášek Statistické metody pro praxi. Lázně Bohdaneč, 25. - 28. 4. 2006, pp. 19-32, ISBN 978-80-239-8995-3.
- [21] Jurák, P. *Odhady diskrétních rozdělení pravděpodobnostipomoví kvazinorem* Dizertační práce, FSI VUT v Brně, Brno, 2008.
- [22] Karpíšek, Z. *F-divergence a kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Analýza dat 2006/I - Sborník přednášek Statistické metody pro praxi, Lázn Bohdaneč, 25. - 28. 4. 2006, s. 19-32, ISBN 978-80-239-8995-3.
- [23] Karpíšek, Z. - Šikulová, M. *Pravděpodobnost a matematická statistika*. Brno: PCDIR, páté vydání, 1996, 204 s., ISBN 80-214-0830-8.
- [24] Karpíšek, Z. - Jurák, P. *Minimální kvazinormy Hellingerova a Shannonova typu*. Sborník celostátního semináře Analýza dat 2002/II, Lázně Bohdaneč, 26. - 29. 11. 2002, s. 83-89, ISBN 80-239-0204-0.
- [25] Karpíšek, Z. - Jurák, P. - Sadovský, Z. *Quasi-norms for discrete probability distribution estimation*. 5th International conference on soft computing applied in computer and economic environments, Kunovice, 26. 1. 2007, s. 113-120, ISBN 80-7314-108-6.
- [26] Karpíšek, Z. - Sadovský, Z. - Jurák, P. *Pearson quasi-norm for probability distribution goodness of fit*. 5th International conference APLIMAT 2006, Bratislava, 7. - 10. 2. 2006, s. 545-551, ISBN 80-967305-7-6.
- [27] Klapka, J. - Dvořák, J. - Popela, P. *Metody operačního výzkumu*. Brno: PC-DIR, 1996.

- [28] Paulov, J. *Teória informácií a regionálna analýza*. PMFA, 1982.
- [29] Pitman, E. J. G. *Some basic theory for statistical inference*. New York: John Wiley & Sons, 1978.
- [30] Rektorys, K. *Přehled užité matematiky II*. Praha: Prometheus, s.r.o., šesté vydání, 1995, 874 s., ISBN 80-85849-62-3.
- [31] Österreicher, F. *F-divergences - representation theorem and metrization*. 22. 12. 2002, [online] URL <<http://www.sbg.ac.at/mat/staff/oesterreicher/vortrag2.pdf>> [rev. 2003-09-02], [cit. 2007-10-21].
- [32] Shannon, C. - Weaver, W. *The mathematical theory of communications*. Urbana, Illinois: University of Illinois Press, 1949.
- [33] Read, T., R., C. - Cressie, N. *Goodness-of-fit Statistics for Discrete Multivariate data*. Springer, New York, 1988.
- [34] Karpíšek, Z. - Jurák, P. *Quasi-norms and Discrete Probability Distribution Estimations*. MENDEL '06. 12th International Conference on Soft Computing. Brno, 31. 5. - 2. 6. 2006, pp. 203-208, ISBN 80-214-3195-4.
- [35] Karpíšek, Z. - Jurák, P. - Šácha, J. *Kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Sborník přednášek 8. národní konference Statistické dny v Brně. Brno, 27. - 28. 6. 2006, pp. 27-28 abstrakt a p. 13 na CD ROM, ISBN 80-214-3214-4.
- [36] Karpíšek, Z. - Jurák, P. - Šácha, J. *Divergences for Discrete Probability Distribution Estimations*. Summer School DATASTAT '06, Proceedings, Masaryk University, Brno, 2007, pp. 109-120, ISBN 978-80-210-4493-7.
- [37] Karpíšek, Z. - Jurák, P. - Sadovský, Z. *Quasi-norms for Discrete Probability Distribution Estimation*. 5th International Conference on Soft Computing Applied in Computer and Economic Environments. Kunovice, 26. 1. 2007, pp. 113 - 120, ISBN 80-7314-108-6.
- [38] Karpíšek, Z. - Jurák, P. - Neradová, V. *Divergences and Quasi-norms of Discrete Probability Distributions*. 6th International Conference APLIMAT 2007 (Part I). Bratislava, 6. - 9. 2. 2007, pp. 387-395, ISBN 978-80-969562-5-8.
- [39] Šácha, J. - Karpíšek, Z. *Kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Moderní matematické metody v inženýrství - 3mi. Ostrava, 2007, pp. 276-280, ISBN 978-80-248-1649-4.
- [40] Karpíšek, Z. - Šácha, J. - Jurák, P. *Discrete Probability Distribution Estimation by Means of Quasi-norms*. STOPTIMA 2007 (International Conference Stochastic Programming and Optimization: Modeling and Applications). Brno, 1. - 5. 9. 2007, p. 10 + poster, ISBN 978-80-214-3474-5.

- [41] Šácha, J. - Karpíšek, Z. - Jurák, P. *Quasi-norms for Discrete Probability Distribution Estimation*. International Conference Risk, Quality and Reliability RQR '07. Ostrava, 20. - 21. 9. 2007, pp. 179-186, ISBN 978-80-248-1575-6.
- [42] Karpíšek, Z. - Neradová, V. - Sadovský, Z. *Pesimistické odhady diskrétního rozdělení pravděpodobnosti*. Konference Inteligentní systémy pro praxi. Lázně Bohdaneč, 30. - 31. 1. 2008, anotace pp. 65-68, plný text 9 p. na CD ROM, ISBN 978-80-7399-354-2.
- [43] Karpíšek, Z. - Neradová, V. *Estimation of Categorical Variable Probability Distribution*. 7th International Conference APLIMAT 2008. Bratislava, 5. - 8. 2. 2008, Book of abstracts p. 101, ISBN 978-80-89313-02-0, Proceedings pp. 1145-1154, ISBN 978-80-89313-03-7.
- [44] Karpíšek, Z. - Neradová, V. - Žampachová, E. *A Contribution to the Estimation of Discrete Probability Distribution*. MENDEL 2008. 14th International Conference on Soft Computing. Brno, 18. - 20. 6. 2008, pp. 287-292, ISSN 1803-3814 (Mendel Journal Series on CD), ISBN 978-80-214-3675-6.
- [45] Šácha, J. - Karpíšek, Z. *Aproximace rozdělení pravděpodobnosti pomocí kvazinorem*. REQUEST '08. Sborník konference CQR. Brno, 25. - 28. 6. 2008, pp. 201-207, ISBN 978-80-214-3774-6.
- [46] Karpíšek, Z. - Neradová, V. - Žampachová, E. *Odhad rozdělení pravděpodobnosti kategoriální veličiny*. REQUEST '08. Sborník konference CQR. Brno, 25. - 28. 6. 2008, pp. 93-101, ISBN 978-80-214-3774-6.
- [47] Karpíšek, Z. - Lacinová, V. *Odhady diskrétního rozdělení pravděpodobnosti s použitím kvazinorem a bootstrapu*. Analýza dat 2010/II - Statistické metody pro technologii a výzkum. Pardubice: TriloByte, CQR, 2010, pp. 131-145. ISBN 978-80-904053-3-2.
- [48] Karpíšek, Z. - Lacinová, V. *Aplikace odhadů rozdělení pravděpodobnosti kategoriální veličiny v průzkumových šetřeních*. Sborník konference Daně - teorie a praxe 2010. Brno: Akademie Sting, 2010, pp. 41-47. ISBN 978-80-86342-91-7.
- [49] Lacinová, V. - Karpíšek, Z. - Sadovský, Z. *Pesimistické odhady rozdělení pravděpodobnosti kategoriální veličiny*. Informační bulletin České statistické společnosti, roč. 22 (2), Praha, 2011, pp. 138-145. ISSN 1210-8022.
- [50] Šácha, J. - Karpíšek, Z. *Odhad rozdělení pravděpodobnosti s obecnými lineárními podmínkami*. Informační bulletin České statistické společnosti, roč. 22 (2), Praha, 2011, pp. 192-199. ISSN 1210-8022.
- [51] Šácha, J. *Interval Estimation for Shannon Quasi-norm*. *Engineering Mechanics (International Journal for Theoretical and Applied Mechanics)*, Brno, 2012, ISSN 1802-1484. (v recenzním řízení)
- [52] Jurák, P. *Estimator 1.4.1*. Autorizovaný software, Brno: VUT CQR 2008.

Autorovy publikace

1. Šácha, J. *Netradiční statistické metody*. Diplomová práce. FSI VUT v Brně, Brno, 2004.
2. Karpíšek, Z. - Sadovský, Z. - Šácha, J. *Pitmanův - Hellingerův test dobré shody*. Sborník konference Aplimat 2005. Bratislava: STU Bratislava, 2005. s. 471-477. ISBN: 80-969264-2-X.
3. Karpíšek, Z. - Šácha, J. *Nepříliš známý test shody diskrétního rozdělení pravděpodobnosti*. Sborník přednášek k 8. ročníku konference "Inteligentní systémy pro praxi". Ostrava-Poruba: AD&M Ostrava, 2004. s. 47-53. ISBN: 80-239-3680-8.
4. Karpíšek, Z. - Jurák, P. - Šácha, J. *Kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Sborník přednášek 8. národní konference Statistické dny v Brně. Brno: Česká statistická společnost, 2006. s. 49-61. ISBN: 80-214-3214-4.
5. Karpíšek, Z. - Jurák, P. - Šácha, J. *Divergences for Discrete Probability Distribution Estimations*. Proceedings of Summer School DATASTAT '06. Brno: Masaryk University Brno, 2007. s. 109-120. ISBN: 978-80-210-4493-7.
6. Šácha, J. - Karpíšek, Z. *Kvazinormy pro odhady diskrétních rozdělení pravděpodobnosti*. Moderní matematické metody v inženýrství - 3mi 2007. Ostrava: JČMF, KMDG VŠB-TU, 2007. s. 276-280. ISBN: 978-80-248-1649-4.
7. Karpíšek, Z. - Šácha, J. - Jurák, P. *Discrete Probability Distribution Estimation by Means of Quasi-norms*. Stoptima - Mozaic 2007. Brno: VUT Brno, 2007. s. 65-74. ISBN: 978-80-214-3474-5.
8. Šácha, J. - Karpíšek, Z. - Jurák, P. *Qusi-Norms for Discrete Probability Distribution Estimation*. Risk, quality and reliability. Ostrava, Czech republic: VŠB - Technical University of Ostrava, 2007. s. 179-185. ISBN: 978-80-248-1575-6.
9. Šácha, J. - Karpíšek, Z. *Aproximace rozdělení pravděpodobnosti pomocí kvazinorem*. REQUEST '08. Sborník konference CQR. Brno, 25. - 28. 6. 2008, pp. 201-207, ISBN 978-80-214-3774-6.
10. Šácha, J. *Aproximace rozdělení pravděpodobnosti*. Firma a konkurenční prostředí 2009 - 1. část. 1. vyd. Brno: MSD, spol. s r. o., 2009, s. 318-323. ISBN 978-80-7392-084-5.
11. Šácha, J. *Odhad rozdělení pravděpodobnosti s lineárními podmínkami*. [CD-ROM]. Enterprise and Competitive Environment 2011. s. 922-930. ISBN 978-80-87106-40-2.
12. Šácha, J. - Karpíšek, Z. *Odhad rozdělení pravděpodobnosti s obecnými lineárními podmínkami*. Informační bulletin České statistické společnosti, roč. 22 (2), Praha, 2011, pp. 192-199. ISSN 1210-8022.
13. Šácha, J. *Kvazinormy s obecnými podmínkami*. Forum Statisticum Slovaca, 2011, roč. 2011, č. 2, s. 174-179. ISSN: 1336- 7420.

14. Šácha, J. *Interval Estimation for Shannon Quasi-norm. Engineering Mechanics (International Journal for Theoretical and Applied Mechanics)*, Brno, 2012, ISSN 1802-1484. (v recenzním řízení)

Autorovy ostatní publikace

1. Ryglová, K. - Vajčnerová, I. - Šácha, J. *Servqual Potential for Quality Management in Hotel Services. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. 2012. sv. LX, č. 2, s. 307-314. ISSN 1211-8516.
2. Vajčnerová, I. - Šácha, J. - Ryglová, K. *Using the principal component analysis for evaluating the quality of a tourist destination. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. 2012. sv. LX, č. 2, s. 449-458. ISSN 1211-8516.

Seznam použitých zkratek a symbolů

\mathbb{R}	množina reálných čísel
\mathbb{R}^*	množina reálných čísel rozšířená o nevlastní prvky ∞ a $-\infty$
Ω	pravděpodobnostní prostor
Σ	jevové pole
P	pravděpodobnostní míra
\mathbf{p}_0	referenční rozdělení pravděpodobnosti s maximální neurčitostí
f	generující funkce f -divergence
$D_f(\mathbf{p}, \mathbf{q})$	f -divergence hustot \mathbf{p} a \mathbf{q}
Λ	Lagrangeova funkce
λ	Lagrangeův multiplikátor
$E(X)$	střední hodnota náhodné veličiny X
$D_f(\mathbf{p}, \mathbf{p}_0)$	kvazinorma diskrétního rozdělení pravděpodobnosti \mathbf{p}
$H(\mathbf{p}, \mathbf{p}_0)$	Hellingerova kvazinorma
$P(\mathbf{p}, \mathbf{p}_0)$	Pearsonova kvazinorma
$S(\mathbf{p}, \mathbf{p}_0)$	Shannonova kvazinorma
$\chi^2(\mathbf{p}, \mathbf{p}_0)$	kvadratická kvazinorma
$L(\mathbf{p}, \mathbf{p}_0)$	logaritmická kvazinorma
f_j	absolutní četnost
M_k	k -tý obecný moment
$\text{Bi}(n; p)$	binomické rozdělení s parametry n a p
$\text{Po}(l)$	Poissonovo rozdělení s parametrem l
$N(\mu; \sigma)$	normální rozdělení s parametry μ a σ
$W(d; c)$	Weibullovo rozdělení s parametry d a c
$E(l)$	exponenciální rozdělení s parametrem l
J	Fischerova míra informace
\mathbf{J}	Fischerova informační matice
T	Taylorův polynom
$\text{var } \mathbf{X}$	kovarianční matice náhodného vektoru X