Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Ondřej Kolafa

Empirické odhady ve stochastickém programování; závislá data

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Vlasta Kaňková, CSc.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická statistika a ekonometrie

Praha 2014

Rád bych na tomto místě poděkoval vedoucí mé práce, RNDr. Vlastě Kaňkové, CSc., za pomoc, díky které jsem ze zorientoval v dané problematice, za podnětné připomínky k vlastnímu textu a především za čas, který mně i mé práci věnovala.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 30. 7. 2014

Název práce: Empirické odhady ve stochastickém programování; závislá data

Autor: Ondřej Kolafa

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Vlasta Kaňková, CSc., Ústav teorie informace a automatizace, Akademie věd ČR

Abstrakt: Práce pojednává o úlohách stochastického programování založených na empirickém a teoretickém rozdělení a jejich vzájemném vztahu. Nejdříve se věnuje případu úloh, kdy empirické rozdělení odpovídá nezávislému náhodném výběru. Jsou ukázány některé základní vlastnosti a poté konvergence úlohy založené na empirickém rozdělení k úloze teoretické. Práce dále zavádí různé druhy závislosti – m-závislost, mixingy a také obecnější pojem slabé závislosti. Pro posloupnosti s některými z těchto závislostí jsou dokázány podobné vlastnosti, které platí pro posloupnosti nezávislé. V práci jsou na závěr teoretické poznatky demonstrovány na numerických příkladech, ve kterých jsou porovnávány posloupnosti závislé s nezávislými i posloupnosti s různou závislostí mezi sebou.

Klíčová slova: stochastické programování, empirické odhady, slabá závislost, mixing

Title: Empiciral Estimates in Stochastic Programming; Dependent Data

Author: Ondřej Kolafa

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Vlasta Kaňková, CSc., Institute of Information Theory and Automation, Czech Academy of Sciences

Abstract: This thesis concentrates on stochastic programming problems based on empirical and theoretical distributions and their relationship. Firstly, it focuses on the case where the empirical distribution is an independent random sample. The basic properties are shown followed by the convergence between the problem based on the empirical distribution and the same problem applied to the theoretical distribution. The thesis continues with an overview of some types of dependence – m-dependence, mixing, and also more general weak dependence. For sequences with some of these types of dependence, properties are shown to be similar to those holding for independent sequences. In the last section, the theory is demonstrated using numerical examples, and dependent and independent sequences, including sequences with different types of dependence, are compared.

Keywords: Stochastic Programming, Empiciral Estimates, Weak Dependence, Mixing

Obsah

Úvod

Práce se zabývá vztahem mezi stochastickými optimalizačními úlohami, které jsou dané teoretickým a empirickým rozdělením. Především se soustředí na „rychlost“, s jakou optimální hodnota úlohy z empirického rozdělení konverguje k optimální hodnotě úlohy z rozdělení teoretického se zvyšujícím se rozsahem výběru. Součástí práce je také přehled různých typů závislosti, které jsou dále srovnávány s nezávislým výběrem a mezi sebou v optimalizačních úlohách, opět především v souvislosti s optimální hodnotou.

Práce je rozdělena na dvě hlavní kapitoly. V první kapitole se nejdříve zavádí základní pojmy, které jsou používány dále v textu. To zahrnuje nejen definice optimalizačních úloh, ale také stručný úvod do stochastického programování, aby čtenář, který se v této oblasti příliš neorientuje, mohl problematice rychle porozumět. Pro lepší pochopení jsou uvedeny také jednoduché příklady.

V další části je zavedena optimalizační úloha založená na náhodném výběru. V takovém případě potom mluvíme o odhadu optimálního řešení v úloze stochastického programování, a je tedy možné uvést jeho základní vlastnosti – nestrannost, konzistenci a konvergenci k normálnímu rozdělení.

V poslední části kapitoly je zkoumána „rychlost“ konvergence odhadu založeném na náhodném výběru k teoretické úloze za pomoci pravděpodobnosti, s jakou se tento odhad odchýlí od úlohy teoretické. Jsou uvedeny podmínky, za kterých lze konvergenci zajistit a jaké je její limitní chování.

V druhé kapitole se práce věnuje závislým výběrům. Nejdříve je zavedena m-závislost jako příklad velmi jednoduché a snadno popsatelné závislosti. Dále je ukázáno, že v úlohách založených na výběru z m-závislých posloupností lze zajistit konvergenci také.

Další část kapitoly se věnuje jiným druhům závislosti. Nejdříve je uveden přehled tzv. mixingů a poté i obecnější pojem slabé závislosti. Následuje přehled několika známých posloupností, které tyto druhy závislosti splňují. Jedná se např. o Markovovy řetězce a ARMA nebo GARCH proces. Nakonec je stejně jako pro m-posloupnosti ukázána konvergence.

Uvedené vlastnosti jsou demonstrovány na numerických příkladech, ve kterých jsou srovnávány úlohy založené na nezávislých a závislých výběrech.

Kapitola je ukončena shrnutím tvrzení, která zajišťují konvergenci k normálnímu rozdělení i pro závislé výběry a tato vlastnost je také demonstrována na numerickém příkladu.

Použité značení

$\mathbb{R}^n$	značí n-rozměrný eukleidovský prostor, jehož prvkem je n-rozměrný vektor x z prostoru $\mathbb{R}^n$, tedy $x = (x_1, \ldots, x_n)$
$\mathbb{Z}$	značí množinu celých čísel
$\|\| \cdot \|\|_1$	značí L_1 normu na $\mathbb{R}^n$, tj. $\|\|x\|\|_1 = \sum_{i=1}^{n} \|x_i\|$
$\|\| \cdot \|\|_2$	značí L_2 normu na $\mathbb{R}^n$, tj. $\|\|x\|\|_2 = \sqrt{\sum_{i=1}^{n} (x_i)^2}$
$\|\| \cdot \|\|_\infty$	značí L_∞ normu na $\mathbb{R}^n$, tj. $\|\|x\|\|_\infty = \max_{i=\{1,\ldots,n\}} \|x_i\|$
$\|\| \cdot \|\|$	značí jakoukoli konzistentní normu na $\mathbb{R}^n$
(a,b)	značí otevřený interval
$[a,b]$	značí uzavřený interval
$f(x) = O(g(x))$	znamená, že f/g je omezená na nějakém okolí bodu x
s.j.	je zkratka pro „skoro jistě“ a $\xrightarrow{\text{s.j.}}$ značí konvergenci s.j.
$\xrightarrow{P}$	značí konvergenci v pravděpodobnosti
$\xrightarrow{D}$	značí konvergenci v distribuci
$E_\xi[h(x,\xi)]$	značí střední hodnotu funkce h pro každé x vzhledem k rozdělení náhodné veličiny ξ, kde h je funkce ve tvaru definovaném v definici 14
$\mathrm{Var}_\xi[h(x,\xi)]$	značí rozptyl funkce h pro každé x vzhledem k rozdělení náhodné veličiny ξ
$\xi \sim R[0,1]$	značí veličinu ξ s rovnoměrným rozdělením na intervalu $[0,1]$
$\xi \sim N(\mu,\sigma^2)$	značí veličinu ξ s normálním rozdělením se střední hodnotou μ a rozptylem σ^2
$\chi(d)$	značí d-okolí množiny $\chi \subset \mathbb{R}^n$ pro $d > 0$
$\lfloor c \rfloor$	značí dolní celou část čísla c

V citacích vět zkratky L., P., T., C. a D. značí Lemma, Proposition, Theorem, Corollary a Definition

Použité definice

Definice 1 (Pravděpodobnostní prostor, Anděl (2011), str. 13)**.** Nechť Ω je prostor elementárních jevů $\omega \in \Omega$ a $\mathcal{A}$ je σ-algebra podmnožin Ω. Jednotlivým množinám z $\mathcal{A}$ je dána pravděpodobnost pomocí pravděpodobností míry P. Trojici $(\Omega, \mathcal{A}, P)$ potom říkáme pravděpodobnostní prostor.

Elementárním jevem ω rozumíme takový náhodný jev, který obsahuje pouze jeden prvek z Ω, tedy neexistují jiné dva jevy ω_1, ω_2 takové, že $\omega = \omega_1 \cup \omega_2$.

Definice 2 (Náhodná veličina, náhodný vektor, Anděl (2011), str. 13)**.** Nechť $\mathcal{B}$ je systém borelovských množin $\mathbb{R}^1$ a $\xi(\omega)$ je měřitelná funkce z $(\Omega, \mathcal{A}, P)$ do $(R, \mathcal{B})$. Potom se $\xi(\omega)$ nazývá náhodná veličina.

Je-li $\mathcal{B}$ je systém borelovských množin $\mathbb{R}^n$ a $\xi(\omega)$ je měřitelná funkce z $(\Omega, \mathcal{A}, P)$ do $(\mathbb{R}^n, \mathcal{B})$, potom $\xi(\omega)$ nazýváme náhodným vektorem.

Borelovskou množinou rozumíme takovou množinu, kterou můžeme získat z otevřených, resp. uzavřených množin pomocí spočetného sjednocení, průniku a doplňku.

Definice 3 (slabá konvergence, Anděl (2011), str. 19)**.** Posloupnost $\{F_i, i = 1, 2, \dots\}$ distribučních funkcí konverguje slabě k F, pokud $F_i(x) \to F(x)$ v každém bodě x, který je bodem spojitosti F.

Definice 4 (konvergence v pravděpodobnosti, Anděl (2011), str. 329)**.** Posloupnost náhodných veličin $\{\xi_i, i = 1, 2, \dots\}$ konverguje v pravděpodobnosti k náhodné veličině ξ, pokud pro každé $\epsilon > 0$ platí

$$\lim_{i\to\infty} P[|\xi_i - \xi| > \epsilon] = 0.$$

Definice 5 (konvergence s.j., Anděl (2011) str. 329)**.** Posloupnost náhodných veličin $\{\xi_i, i = 1, 2, \dots\}$ konverguje s.j. k náhodné veličině ξ, pokud platí

$$P[\lim_{i\to\infty} \xi_i = \xi] = 1.$$

Definice 6 (konvergence v distribuci, Anděl (2011) str. 329)**.** Nechť posloupnost náhodných veličin $\{\xi_i, i = 1, 2, \dots\}$ má distribuční funkci F_i a $\{\xi\}$ má distribuční funkci F. Jestliže F_i slabě konvergují k F, říkáme, že $\{\xi_i\}$ konverguje v distribuci k $\{\xi\}$.

Definice 7 (stejnoměrná konvergence, Zelený (2009), str. 1)**.** Říkáme, že posloupnost funkcí $\{f_i(x), i = 1, 2, \dots\}$ konverguje stejnoměrně k $f(x)$, pokud pro každé $\epsilon > 0$ existuje N_0 přirozené takové, že

$$\forall x, \forall N \text{ přirozené}, N > N_0 : |f_N(x) - f(x)| < \varepsilon.$$

Definice 8 (polospojitost, Shapiro a kol. (2009), str. 333)**.** Funkce $f : x \to R \cup \{-\infty, +\infty\}$ je zdola polospojitá v bodě $x_0 \in \mathbb{R}^n$, pokud platí

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

Definice 9 (nestrannost, Anděl (2011), str. 101)**.** Odhad $\hat{g}(\theta)$ parametrické funkce $g(\theta)$ na základě náhodného výběru veličiny ξ se nazývá nestranný, pokud platí $E_\xi[\hat{g}(\theta)] = g(\theta)$.

Definice 10 (konzistence, Anděl (2011), str. 105)**.** Předpokládejme, že θ je jednorozměrný parametr a $\xi_1, \dots, \xi_N$ je výběr z rozdělení, které závisí na θ. Pro každé přirozené N mějme definován odhad $\hat{g_N}(\xi_1, \dots, \xi_N)$, který se nazývá konzistentní, pokud platí $\hat{g_N}(\xi_1, \dots, \xi_N) \xrightarrow{P} \theta$.

Řekneme, že je silně konzistentní, pokud platí $\hat{g_N}(\xi_1, \dots, \xi_N) \xrightarrow{\text{s.j.}} \theta$.

Definice 11 (momentová vytvořující funkce, Mandl a Mazurová (1999), str. 15)**.** Nechť ξ je náhodná veličina. Pokud existuje $h > 0$ takové, že platí

$$M(t) = E_\xi[\mathrm{e}^{\xi t}], |t| < h,$$

kde střední hodnota existuje a je konečná, potom $M(t)$ značí momentovou vytvořující funkci veličiny ξ.

Definice 12 (Lipschitzovská funkce, Searcóid (2006), D. 9.4.1, str. 154)**.** Nechť $\chi \subseteq \mathbb{R}^n$ je otevřená množina. Řekneme, že funkce $f : \chi \to \mathbb{R}^1$ je Lipschitzovská nebo Lipschitzovsky spojitá na χ s parametrem $L \geq 0$, pokud platí

$$|f(x) - f(y)| \leq L||x - y||$$

pro všechna $x, y \in \chi$.

Kapitola 1

Nezávislé posloupnosti

1.1 Úvod do stochastického programování

Matematické programování se zabývá řešením úloh, které se na základě vstupních parametrů snaží najít optimální řešení ve smyslu minima nebo maxima nějaké funkce.

Těmito úlohami může být například minimalizace nákladů na rozvoz zboží vozidlem do prodejen (varianta této úlohy je známá pod pojmem dopravní problém), maximalizace výnosů výběrem optimálního výrobního plánu nebo i kombinace minimalizace rizika a maximalizace výnosu portfolia finančních aktiv (např. Markowitzův model).

Obvykle můžeme úlohu matematického programování včetně zmíněných výše zapsat ve tvaru:

Definice 13 (Úloha matematického programování, Kall a Wallace (1994), str. 2, rovnice (1.3))**.**

$$
\begin{aligned}
&\min h(x) \\
&\text{za podmínek } g_i(x) \leq 0, i = 1, \ldots, m \\
&x \in X \subset \mathbb{R}^n,
\end{aligned}
\tag{1.1}
$$

kde $h(x), g_i(x) : \mathbb{R}^n \to \mathbb{R}^1$. Funkci $h(x)$ říkáme účelová funkce, nerovnicím $g_i(x) \leq 0, i = 1, \ldots, m$, omezující podmínky a množinu určenou omezujícími podmínkami nazýváme množina přípustných řešení.

Úlohou je tedy nalézt minimum účelové funkce na dané množině za platnosti omezujících podmínek. Nadále budeme používat zkrácený zápis této úlohy:

$$\min_{x \in \mathcal{X}} h(x),$$

kde $\chi = \{x \in X, g_i(x) \leq 0, i = 1, \ldots, m\}$ je množina přípustných řešení definovaná stejně jako v definici 1.1 a o níž předpokládáme, že je neprázdná. Řešení úlohy budeme značit x^*.

Zmíněnou úlohu samozřejmě můžeme snadno převést na úlohu maximalizační, v textu se ale budeme věnovat úloze v minimalizačním tvaru. Uveďme si nyní velmi jednoduchý příklad úlohy (1.1).

Příklad 1. Prodejce nakupuje zboží za cenu p za jednotku a prodává za cenu c za jednotku (předpokládejme $p < c$). Je mu známo, že poptávka po zboží bude D. Cílem prodejce je maximalizovat zisk, tedy nakoupit zboží tak, aby ho mohl co nejvíce prodat a zároveň nenakoupil žádné navíc. Úlohu lze tedy zapsat takto:

$$\min_{x \in \chi} \left[x \cdot p - c \cdot \min(D, x)\right],$$

kde χ je interval $[0,\infty)$ a $-(x \cdot p - c \cdot \min(D, x))$ je dosažený zisk. Řešením úlohy je zjevně $x = D$. □

Nalezení řešení je tedy pro prodejce elementární úvahou. Do složitejší situace se ale dostane v případě, že poptávku po zboží nezná, resp. je-li poptávka D náhodnou veličinou. Předpokládáme, že prodejce musí nákup zboží provést dříve, než zná realizaci veličiny D.

V takovém případě není jasné, co přesně by znamenalo minimalizovat účelovou funkci, ve které je obsažena náhodná veličina. Řešením je převést úlohu na deterministickou, např. nahradit účelovou funkci střední hodnotou, samozřejmě je-li to v dané úloze možné.

Pokud toto aplikujeme na příklad 1, dostáváme známou úlohu prodavače novin, podobně jak je uvedena např. v Ovchinnikov a Raz (2011), str. 6.

Příklad 2. Nechť je zadání stejné jako v příkladu 1, ale poptávka D je náhodná veličina s nějakým známým rozdělením. Potom je možné optimalizační úlohu zapsat takto:

$$\min_{x \in \chi} E_D(x \cdot p - c \cdot \min(D, x)).$$

□

Vidíme, že spočtením střední hodnoty opět dostaneme deterministickou úlohu, i když tentokrát její tvar závisí na rozdělení veličiny D.

Poměrně obecný případ úlohy, kterou jsme uvedli v předchozím příkladu, můžeme zapsat jako úlohu nalézt minimum ze střední hodnoty funkce:

Definice 14 (Shapiro a kol. (2009), str. 155, rovnice (5.1))**.**

$$\min_{x \in \chi} E_\xi[h(x,\xi)] \tag{1.2}$$

kde $x \in \mathbb{R}^n$, $\chi \subset \mathbb{R}^n$ je neprázdná, uzavřená množina přípustných řešení, ξ je náhodný vektor nabývající hodnot z množiny $\Xi \subset \mathbb{R}^m$, $h(\cdot,\cdot)$: $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^1$ a střední hodnota $h(x,\xi)$ existuje pro každé $x \in \chi$. Tato úloha je jedním z typů úloh stochastického programování.

I když se v dalším textu budeme zabývat především úlohou typu (1.2), uveďme pro pořádek ještě alespoň jednu z dalších možností, jak se vypořádat s náhodnými veličinami v optimalizačních úlohách. Prodavač novin z příkladu 2 předpokládá, že se mu tento přístup vyplatí přes nějaké dlouhé období. Zároveň se mu ale může stát, že v některý den vyprodá poslední kus novin, ale ještě zdaleka neuspokojil poptávku zákazníků (záleží samozřejmě na rozdělení veličiny D). To může vést k tomu, že si zákaznici najdou jiného prodavače, u kterého mají větší jistotu, že bude mít noviny k dispozici.

Prodavač novin tedy může být motivován k tomu, aby se snažil uspokojit poptávku většiny zákazníků, kteří k němu přijdou, i za cenu toho, že mu nějaké noviny zbudou. Tento požadavek, kterému se říká pravděpodobnostní omezení, můžeme zapsat jako $P[x \geq D] \geq 1 - \alpha$, pokud lze pro danou úlohu tuto pravděpodobnost vyjádřit. Prodavač tedy chce uspokojit poptávku zákazníků alespoň s pravděpodobností $1 - \alpha$, kde α zpravidla uvažujeme malé. Dostáváme tak následující úlohu.

Příklad 3.

$$\min_{x \in \chi_1} \left[x \cdot p - c \cdot E_D[\min(D, x)] \right],$$

kde $\chi_1 = \{x \in \chi : P[x \geq D] \geq 1 - \alpha\}$. Má-li rozdělení veličiny D distribuční funkci F, můžeme množinu χ definovat takto:

$$\chi = \{x \geq F_D^{-1}(1 - \alpha), x \in [0,\infty)\},$$

kde F_D^{-1} značí kvantilovou funkci veličiny D. Kvantilovou funkci zde chápeme jako funkci určenou pomocí následujícího předpisu:

$$F_D^{-1}(u) = \inf\{d : F(d) \geq u\}, 0 < u < 1,$$

viz Anděl (2011), str. 17. □

1.2 Základní vlastnosti

Uvažujme úlohu stochastického programování (1.2), kde množina přípustných řešení χ je neprázdná uzavřená podmnožina $\mathbb{R}^n$. Existuje-li střední hodnota funkce $h(x,\xi)$, kde $x \in \chi$ a ξ nabývá hodnot z Ξ, můžeme označit:

$$H(x) = E_\xi[h(x,\xi)]. \tag{1.3}$$

Dále v textu předpokládejme, že funkce $H(x)$ je omezená pro všechna $x \in \chi$. Existuje-li $\min\limits_{x\in\chi} H(x)$, můžeme dále psát

$$H(x^*) = \min_{x\in\chi} H(x),$$

kde $x^* \in \chi$ je optimálním řešením úlohy (1.2) a kde označíme $\vartheta = H(x^*)$ jako optimimální hodnotu úlohy (1.2).

Ukázali jsme zápis i příklady úlohy, kdy známe rozdělení náhodné veličiny, která se v úloze vyskytuje. V praxi se ale často může stát, že rozdělení známé není a k dispozici jsou pouze naměřená historická data, případně hodnoty získané simulací.

Věnujme se tedy nyní modifikaci úlohy (1.2) pro náhodný výběr. Předpokládejme, že máme náhodný výběr $\{\xi_i, i = 1, \ldots, N\}$, který obsahuje N realizací náhodného vektoru ξ. Pokud nebude uvedeno jinak, budeme v této kapitole uvažovat tento výběr nezávislý.

Pro libovolné $x \in \chi$ lze nyní snadno odhadnout střední hodnotu funkce $H(x)$ pomocí průměru z hodnot $h(x,\xi_i), i = 1, \ldots, N$. Tím nám vzniká následující aproximace úlohy (1.2):

$$\min_{x\in\chi} \frac{1}{N} \sum_{i=1}^{N} h(x,\xi_i). \tag{1.4}$$

Analogicky k úloze (1.2) můžeme zavést i následující značení. Řešení úlohy budeme značit jako x_N^* a empirický odhad funkce $H(x)$ budeme značit jako $H_N(x)$, tedy

$$H_N(x) = \frac{1}{N} \sum_{i=1}^{N} h(x,\xi_i).$$

Dále zaveďme i $\vartheta_N = H_N(x_N^*)$ jako optimální hodnotu.

Na úlohu (1.4) lze, především pokud neznáme rozdělení vektoru ξ, pohlížet i jako na úlohu stochastického programování se scénáři $\{\xi_i, i = 1, 2, \ldots N\}$, kdy každý nastane s pravděpodobností $1/N$, nebo také jako na úlohu, kde ξ má empirické rozdělení dané $\{\xi_i, i = 1, 2, \ldots N\}$.

Připomeňme, že pokud by $\{\xi_i, i = 1, \ldots, N\}$ byla posloupnost náhodných veličin a úloha byla vhodného tvaru (např. úloha z příkladu 3), mohli bychom podobně jako úlohu (1.4) zavést i úlohu s pravděpodobnostním omezením, kdy bychom konstruovali omezení pomocí výběrových kvantilů.

Ukažme si nyní dvě základní vlastnosti odhadu $H_N(x)$. Jelikož $\{\xi_i, i = 1, \ldots, N\}$ je výběr z náhodného vektoru ξ, zjevně platí pro každé $x \in \chi$, že

$$\begin{aligned} E_\xi[H_N(x)] &= E_\xi\left[\frac{1}{N}\sum_{i=1}^{N} h(x,\xi_i)\right] = \frac{1}{N}\sum_{i=1}^{N} E_\xi[h(x,\xi_i)] = \\ &= \frac{1}{N}\sum_{i=1}^{N} E_\xi[h(x,\xi)] = N \cdot \frac{1}{N} E_\xi[h(x,\xi)] = H(x). \end{aligned}$$

Odhad je tedy nestranný pro každé $x \in \chi$. Podívejme se nyní na konzistenci odhadu ϑ_N, pro který si ukážeme slabší tvrzení.

Věta 1 (Shapiro a kol. (2009), str. 163)**.** Pro odhad ϑ_N parametru ϑ platí nerovnost

$$\vartheta \geq E_\xi[\vartheta_N]$$

za předpokladu, že $E_\xi[\vartheta_N]$ existuje.

Důkaz. Zjevně platí nerovnost $H_N(x') \geq \inf_{x\in\chi} H_N(x)$ pro všechna $x' \in \chi$. Pokud na tuto nerovnost použijeme střední hodnotu, dostaneme

$$E_\xi[H_N(x')] \geq E_\xi\left[\inf_{x\in\chi} H_N(x)\right].$$

Jelikož z nestrannosti odhadu $H_N(x)$ pro všechna x, který jsme ukázali dříve, vyplývá, že $E_\xi[H_N(x')] = H(x')$, můžeme nerovnost přepsat následovně:

$$H(x') \geq E_\xi\left[\inf_{x\in\chi} H_N(x)\right] = E_\xi[\vartheta_N].$$

Předchozí nerovnost platí pro libovolné $x' \in \chi$, bude tedy platit i pro $H(x^*) = \vartheta$, čímž je věta dokázána. □

Pro pořádek si ještě uveďme tvrzení, které nám říká, že tato vychýlenost odhadu ϑ_N klesá s rostoucím rozsahem výběru.

Věta 2 (Shapiro a kol. (2009), P. 5.6, str. 163)**.** Pro každé N přirozené platí

$$E_\xi[\vartheta_N] \leq E_\xi[\vartheta_{N+1}] \leq \vartheta.$$

1.2.1 Konzistence odhadů

Ukažme si ještě konzistenci odhadu $H_N(x)$. Pokud je výběr $\{\xi_i, i = 1, \ldots, N\}$ nezávislý pro jakékoli N, potom podle Silného zákona velkých čísel (Anděl (2011), V. B.4), platí

$$H_N(x) \xrightarrow{\text{s.j.}} H(x) \text{ pro každé } x \in \chi.$$

Ukázali jsme si konvergenci funkce $H_N(x)$ v každém bodě $x \in \chi$, pokračujme nyní s konvergencí optimálních hodnot.

Věta 3 (Shapiro a kol. (2009), P. 5.2, str. 157)**.** Předpokládejme, že $H_N(x) \xrightarrow{\text{s.j.}} H(x)$ stejnoměrně pro $N \to \infty$. Potom $\vartheta_N \xrightarrow{\text{s.j.}} \vartheta$.

Povšimněme si, že pro konvergenci optimálních hodnot nám už nestačí pouze nezávislý výběr, ale musíme přidat i podmínku stejnoměrné konvergence. Pro praktické řešení optimalizačních úloh většinou nestačí znát pouze optimální hodnotu, ale je nutné znát zároveň i optimální řešení, resp. množinu optimálních řešení. Ukážeme si, že i množina optimálních řešení za splnění určitých předpokladů „konverguje“.

Pro tento účel zaveďme odchylku $D(A,B)$ množiny A od množiny B, pomocí které můžeme vyčíslit „podobnost“ množin:

$$D(A,B) = \sup_{x \in A} \operatorname{dist}(x,B), \text{ kde}$$

$$\operatorname{dist}(x,B) = \inf_{x' \in B} ||x - x'||_2.$$

Je nutné si uvědomit, že nemusí platit $D(A,B) = D(B,A)$, záleží tedy na pořadí, v jakém odchylku množin zjišťujeme.

Věta 4 (Shapiro a kol. (2009), T. 5.3, str. 158)**.** Předpokládejme, že existuje kompaktní množina $C \subset \mathbb{R}^n$ taková, že:

(i) množina S optimálních řešení úlohy (1.2) je neprázdná a je obsažena v C,

(ii) funkce $H(x)$ je omezená a spojitá na C,

(iii) $H_N(x) \xrightarrow{\text{s.j.}} H(x)$ stejnoměrně pro $N \to \infty$,

(iv) pro dostatečně velké N, množina S_N optimálních řešení úlohy (1.4) je neprázdná a $S_N \subset C$,

potom $\vartheta_N \xrightarrow{\text{s.j.}} \vartheta$ a $D(S_N,S) \xrightarrow{\text{s.j.}} 0$ pro $N \to \infty$.

Ověření stejnoměrné konvergence může být v optimalizačních úlohách problematické. Uveďme si tedy větu, jejíž předpoklady jsou ověřitelné snadněji. Je ale nutné počítat s omezením, že jednou z podmínek je i konvexita úlohy, která sice pro mnohé známé optimalizační úlohy platí, ale některé úlohy z praxe ji samozřejmě splňovat nemusí.

Věta 5 (Shapiro a kol. (2009), T. 5.4, str. 159)**.** Předpokládejme, že

(i) integrand funkce h z rovnice (1.3) je náhodný zdola polospojitý,

(ii) pro skoro všechna $\xi \in \Xi$ je funkce $h(\cdot, \xi)$ konvexní,

(iii) množina χ je uzavřená a konvexní,

(iv) střední hodnota funkce H je zdola polospojitá a existuje bod $\overline{x} \in \chi$ takový, že $H(x) < +\infty$ pro všechna x v okolí $\overline{x}$,

(v) množina S je neprázdná a omezená,

(vi) zákon velkých čísel platí pro každé $x \in \chi$,

potom $\vartheta_N \xrightarrow{\text{s.j.}} \vartheta$ a $D(S_N,S) \xrightarrow{\text{s.j.}} 0$ pro $N \to \infty$.

Velká třída náhodných polospojitých funkcí jsou tzv. Carathéodoryho funkce, tj. reálné funkce $F : \mathbb{R}^n \times \Omega \to \mathbb{R}^1$ takové, že $F(x,\cdot)$ je měřitelná pro každé $x \in \mathbb{R}^n$ a $F(\cdot,\omega)$ je spojitá pro skoro všechna $\omega \in \Omega$. Mohli bychom tedy upravit předpoklad (i) a požadovat integrand h s touto, lépe ověřitelnou vlastností. (Pro vysvětlení pojmu náhodné polospojitosti je možné nahlédnout do Shapiro a kol. (2009), str. 366.)

Povšimněme si dále, že podmínka (iv) nepožaduje omezenost funkce H(x) pro všechna $x \in \chi$. Integrand funkce $h(x,\xi)$ tedy může nabývat i hodnot $\pm\infty$ pro nějaké x. Připomeňme také, že polospojitost je slabší vlastností než spojitost, takže platnost věty se nemění, pokud v ní nahradíme polospojitost zdola za spojitost.

Příklad 4. Ukažme si nyní použití věty 5 v příkladu 2. Můžeme použít zavedené značení, tedy

$$h(x,D) = x \cdot p - c \cdot \min(D, x).$$

Předpoklady věty 5 jsou splněny pro běžné optimalizační úlohy, což platí i pro úlohu naši. Věnujme se ale nyní podrobněji bodu (v).

Poptávka může mít například diskrétní rovnoměrné rozdělení na intervalu $[a,b]$, kde $a,b \in \mathbb{R}^1$, $a < b$, takže nabývá celočíselných hodnot, což je vzhledem k povaze úlohy logické. Nedopustíme se ale velké chyby tím, že budeme pro názornější výpočet předpokládat, že poptávka má rozdělení spojité. To navíc pro snažší výpočet transformujeme na interval [0,1].

Pro jednoduchost uvažujme, že poptávka D má rovnoměrné rozdělení: $D \sim R[0,1]$. Funkci $H(x)$ můžeme přepsat následovně:

$$H(x) = x \cdot p - c \cdot E_D[\min(D, x)], \text{kde}$$

$$E_D[\min(D, x)] = \int_0^x D \mathrm{d}D + x \int_x^1 1 \mathrm{d}D = x - \frac{x^2}{2}.$$

Z toho už snadno dopočítáme tvar funkce $H(x)$:

$$H(x) = x(p - c) + c\frac{x^2}{2}.$$

Analýzou této funkce zjistíme minimum, které bude odpovídat jedinému optimálnímu řešení úlohy, tedy $S = \{x^*\} = \{\frac{c-p}{c}\}$. Jelikož množina obsahuje právě jeden prvek, je zjevně neprázdná a omezená a bod (v) je tak splněn. Dosazením můžeme také snadno získat i optimální hodnotu $\vartheta = -\frac{(c-p)^2}{2c}$. □

Příklad 5. Uvažujme nyní drobnou modifikaci úlohy 4. Předpokládejme, že se na prodavače novin vztahuje 100% daň z příjmu nad určitou hranici a. Pokud dosáhne přes nějaké dlouhé období (např. rok) průměrného zisku vyššího, než je tato hranice a, musí celý zisk nad tuto hranici odevzdat. Jinými slovy nemůže dosáhnout průměrného zisku vyššího, než je právě a.

Uvědomme si, že je vhodné zavést parametr a následovně: $0 > -a > H(x^*)$. Pokud by $-a \leq H(x^*)$, daň by našemu prodavači jeho příjmy nijak neovlivnila a úloha by se tak prakticky nelišila od úlohy v příkladu 4. Úloha je tedy tvaru

$$\min_{x \in \chi} E_D[\max(h(x,D), -a)] = \min_{x \in \chi} E_D[\max(x \cdot p - c \cdot \min(D, x), -a)]$$

tedy $h(x,D)$ použijeme z příkladu 4. Nezapomínejme na záporné znaménko u hranice a, protože máme úlohu zapsánu jako minimalizační. Snadno nahlédneme, že optimální řešení nebude nyní jediné, ale bude jím množina hodnot určená vztahem

$$-a \geq H(x),$$

kde opět $H(x)$ vezmeme z příkladu 4. Množina optimálních řešení je tedy interval $[x_1,x_2]$, kde $x_{1,2} = \frac{c-p\pm\sqrt{-2ac+(c-p)^2}}{c}$. Interval je ale stále konvexní množina, a tak máme zaručenu platnost bodu (v). Ostatní body lze odvodit podobně jako v úloze 4, stejně tak vypočítat i optimální hodnotu. □

Zatím jsme uvažovali, že množina χ je fixní, tedy nezávislá na výběru $\{\xi_i, i = 1, \dots, N\}$. V některých případech tomu ale tak být nemusí a potom dostaneme úlohu ve tvaru:

$$\min_{x \in \chi_N} H_N(x), \tag{1.5}$$

kde χ_N je podmnožina $\mathbb{R}^n$ závisející na výběru $\{\xi_i, i = 1, \dots, N\}$. Volba x tedy závisí na množině, jejíž tvar závisí na pravděpodobnostní míře.

Příklad 6. Úlohu tohoto typu dostaneme opět modifikací příkladu 2, kdy množina χ bude záviset na náhodné veličině D. Toho dosáhneme tak, že si do definice množiny zařadíme omezení funkcí $g(x)$, která závisí na veličině D. Například definujme funkci, v níž se vyskytuje α-kvantil rozdělení veličiny D, který označíme q_α:

$$g(x) = x - q_\alpha.$$

Pomocí této funkce dále definujme množinu přípustných řešení χ_1:

$$\chi_1 = \{x \in \chi : g(x) \leq 0\},$$

kde χ je množina přípustných řešení stejná jako příkladu 2. Toto omezení si můžeme představit tak, že prodavač nemůže nakoupit více novin, než je určitá hranice, která ale závisí na rozdělení poptávky. To může nastat například v případě, že na stejném místě operuje i další prodavač novin, kterému musí první prodavač přenechat nějakou část prodejů. Tato část ale závisí na rozdělení veličiny D. $\square$

Množinu přípustných řešení χ_N úlohy sestavené z náhodného výběru můžeme obecně definovat takto:

$$\chi_N = \{x \in \chi : g_N(x) \leq 0\},$$

kde $g_N(x)$ bude mít v našem případě tvar

$$\chi_N = \{x \in \chi : x - q_{\alpha N} \leq 0\},$$

kde $q_{\alpha N}$ je výběrový α-kvantil z náhodného výběru $\{D_i, i = 1, \ldots, N\}$.

Množina χ_N sice závisí na daném výběru, ale se zvyšujícím se rozsahem výběru zjevně „konverguje" k χ, jelikož empirická kvantilová funkce konverguje ke kvantilové funkci veličiny D. Dá se očekávat, že k sobě budou konvergovat i optimální řešení. Uveďme si úpravu věty 5, která toto očekávání potvrzuje.

Věta 6 (Shapiro a kol. (2009), T. 5.5, str. 160)**.** Předpokládejme, že kromě platnosti podmínek (i) až (iv) z věty 4, platí i následující:

(v) pokud $\{x_N\} \in \chi_N$ a $x_N \xrightarrow{\text{s.j.}} x$, potom $x \in \chi$,

(vi) pro nějaký bod $x \in S$ existuje posloupnost $x_N \in \chi_N$ taková, že $x_N \xrightarrow{\text{s.j.}} x$.

Potom $\vartheta_N \xrightarrow{\text{s.j.}} \vartheta$ a $D(S_N,S) \xrightarrow{\text{s.j.}} 0$ pro $N \to \infty$.

1.2.2 Interval spolehlivosti

Konzistence odhadů, které jsme se v této kapitole věnovali, nám zajišťuje, že chyba odhadu, tj. odlišnost odhadu od odhadované hodnoty, se s rostoucím rozsahem výběru blíží nule.

V praxi ale nemáme k dispozici nekonečně či dostatečně velký výběr a bylo by dobré mít k dispozici znalost o velikosti chyby, kterou může odhad mít. Odvodíme si interval spolehlivosti známým postupem pomocí centrální limitní věty.

Předpokládejme, že $H_N(x)$ má rozptyl $\sigma^2(x)/N$, kde $\sigma^2(x) = \mathrm{Var}_\xi[h(x,\xi)]$. Potom podle centrální limitní věty platí:

$$\sqrt{N}\,[H_N(x) - H(x)] \xrightarrow{D} Z$$

kde Z je veličina s normálním rozdělením $N(0,\sigma^2(x))$. To znamená, že odhad $H_N(x)$ má pro dostatečně velké N přibližně normální rozdělení se střední hodnotou $H(x)$ a rozptylem $\sigma^2(x)/\sqrt{N}$.

To nám dovolí sestrojit následující $100(1-\alpha)$ procentní interval spolehlivosti pro $H(x)$:

$$\left[H_N(x) - \frac{u_{(\alpha/2)}\sigma_N(x)}{\sqrt{N}}, H_N(x) + \frac{u_{(\alpha/2)}\sigma_N(x)}{\sqrt{N}}\right],$$

kde $u_{(\alpha/2)}$ značí $1-\alpha/2$ kvantil normálního rozdělení $N(0,1)$, tj. $u_{(\alpha/2)} = \Phi^{-1}(1-\alpha/2)$ a $\sigma_N^2(x)$ je odhad $\sigma^2(x)$, tj.:

$$\sigma_N^2(x) = \frac{1}{N-1}\sum_{i=1}^{N}\left[h(x,\xi_i) - H_N(x)\right]^2.$$

Ukažme si nyní podobnou vlastnost pro optimální hodnoty.

Věta 7 (Shapiro a kol. (2009), T. 5.7., str. 165)**.** Nechť platí:

(i) existuje nějaké $\tilde{x} \in \chi$, pro které je $E_\xi[h(\tilde{x},\xi)^2]$ konečná,

(ii) existuje měřitelná funkce $C : \Xi \to R_+$ taková, že $E_\xi[C(\xi)^2]$ je konečná a

$$|h(x,\xi) - h(x',\xi)| \leq C(\xi)||x - x'||$$

pro všechna $x, x' \in \chi$,

(iii) χ je kompaktní a S obsahuje pouze jeden prvek, tj. $S = \{x^*\}$. Potom

$$\sqrt{N}\left[\vartheta_N - \vartheta\right] \xrightarrow{D} Z,$$

kde Z je veličina s normálním rozdělením $N(0,\sigma^2(x^*))$.

Podobně jako pro $H(x)$ můžeme tedy za pomoci x_N^* (x^* zpravidla neznáme) sestrojit interval spolehlivosti pro ϑ:

$$\left[\vartheta_N - \frac{u_{(\alpha/2)}\sigma_N(x_N^*)}{\sqrt{N}}, \vartheta_N + \frac{u_{(\alpha/2)}\sigma_N(x_N^*)}{\sqrt{N}}\right].$$

1.3 Konvergence odhadů

Ukázali jsme si, za jakých podmínek k sobě konvergují optimální hodnoty (resp. množiny optimálních řešení). Nyní se budeme zabývat tím, jak „rychlá“ je jejich konvergence.

V dalším textu se budeme věnovat výrazu $P[H(x_N^*) - H(x^*) \geq \epsilon]$ a zda ho lze omezit nějakou funkcí. Ukáže se, že za určitých předpokladů lze k omezení využít exponenciální funkci závislou na N, tedy na velikosti výběru, což tak naznačuje i to, jak „rychlá“ je konvergence.

Pro začátek si uveďme základní větu, ze které jsou odvozeny další výsledky.

Věta 8 (Dai a kol. (2000), L. 3.1., str. 493, Chernoffova mez)**.** Nechť ξ nabývající hodnot z $\mathbb{R}^1$ je libovolná náhodná veličina s konečnou momentovou vytvořující funkcí $E_\xi[\mathrm{e}^{\theta\xi}]$ pro $0 \leq \ \theta \leq \theta_0$, kde $\theta_0 > 0$ je konstanta. Potom platí:

$$P[\xi \geq x] \leq E_\xi[\mathrm{e}^{\theta\xi}]\mathrm{e}^{-\theta x}$$

pro všechna $x \in \mathbb{R}^1$ a $0 \leq \ \theta \leq \theta_0$.

Nyní se budeme zabývat konvergencí pro optimální hodnoty.

Věta 9 (Dai a kol. (2000), T. 3.1., str. 493)**.** Předpokládejme, že

(A1) existuje $a > 0$, $\theta_0 > 0$, $\eta(\xi) : \mathbb{R}^m \to \mathbb{R}^1$ takové že
$|h(x,\xi_i)| \leq a\eta(\xi)$, $E_\xi[\mathrm{e}^{\theta\xi}] < \infty$,
pro všechna $x \in \chi$ a $0 \leq \ \theta \leq \theta_0$.

Potom pro jakékoli $\epsilon > 0$ existují $\alpha > 0$, $\beta > 0$ taková, že

$$P[H(x_N^*) - H(x^*) \geq \epsilon] \leq \alpha\mathrm{e}^{-\beta N}$$

pro všechna $N > 0$.

Pokud existuje optimální řešení a je jediné, lze si ukázat podobné tvrzení jako pro optimální hodnotu.

Věta 10 (Dai a kol. (2000), T. 3.2., str. 497)**.** Předpokládejme (A1) a že x^* je jediným řešením úlohy (1.2). Potom pro každé $\epsilon > 0$ existují $\alpha > 0$, $\beta > 0$ taková, že

$$P\left[||x_N^* - x^*||_2 \geq \epsilon\right] \leq \alpha\mathrm{e}^{-\beta N}$$

pro všechna $N \geq 1$.

Pokud máme úlohu, ve které x nabývá diskrétních hodnot, lze samozřejmě modifikovat větu 10 a ukázat si, že i pro tyto úlohy platí podobné tvrzení.

Důsledek 1 (Dai a kol. (2000), C. 3.1., str. 497)**.** Nechť χ je diskrétní. Za předpokladů věty 10 existuje $\alpha > 0$, $\beta > 0$ takové, že

$$P[x_N^* = x^*] \geq 1 - \alpha\mathrm{e}^{-\beta N}$$

pro všechna $N \geq 1$.

Víme, že $P[H(x_N^*) - H(x^*) \geq \epsilon]$ je omezena exponenciální funkcí závislou na N. Existence takové funkce je sice zaručena, ale už je složitější takovou funkci pro obecnou úlohu najít. Pro vhodně zvolenou funkci to ale možné je. V následujícím příkladu použijeme funkci kvadratickou.

Příklad 7 (Dai a kol. (2000), str. 502)**.** Předpokládejme tedy optimalizační problém ve tvaru:

$$\min_{x \in \mathbb{R}^1} E_\xi[h(x,\xi)] = \min_{x \in \mathbb{R}^1} E_\xi[a_2(x - a_1\xi)^2 + a_0], \tag{1.6}$$

kde $a_2 > 0$. Náhodná veličina ξ může mít libovolné rozdělení, pro které platí

$$E_\xi[\xi] = \overline{\xi}, \quad \mathrm{Var}_\xi[\xi] = \sigma_\xi^2.$$

Spočítejme nyní optimální řešení úlohy (1.6).

$$\begin{aligned} H(x) &= E_\xi[h(x,\xi)] = \\ &= E_\xi[a_2x^2 - 2a_1a_2x\xi + a_1^2\xi^2 + a_0] = \\ &= a_2x^2 - 2a_1a_2x\overline{\xi} + a_1^2(\overline{\xi}^2 + \sigma_\xi^2) + a_0 \end{aligned}$$

a pokud si uvědomíme, že $H(x)$ je kvadratická, konvexní funkce, nalezneme řešení následovně:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}H(x) &= 2a_2x - 2a_1a_2\overline{\xi} = 0 \\ a_2x &= a_1a_2\overline{\xi}, \end{aligned}$$

tedy $x^* = a_1\overline{\xi}$. Dosazením do $H(x)$ pak získáme

$$H(x^*) = E_\xi[a_2(a_1\overline{\xi} - a_1\xi)^2 + a_0] = a_2a_1^2\sigma_\xi^2 + a_0.$$

Předpokládejme, že známe $\{\xi_i, i = 1, \ldots, N\}$ a pokračujme podobně s empirickou aproximací $H_N(x)$:

$$H_N(x) = \frac{1}{N}\sum_{i=1}^{N}[a_2(x - a_1\xi_i)^2 + a_0].$$

Pro nalezení x_N^* vyřešíme rovnici

$$\begin{aligned} \frac{\mathrm{d}}{dx}H_N(x) &= \frac{1}{N}\sum_{i=1}^{N}[2a_2(x - a_1\xi_i)] = 0 \\ \frac{1}{N}\sum_{i=1}^{N} x &= \frac{1}{N}\sum_{i=1}^{N} a_1\xi_i, \end{aligned}$$

což dává

$$x_N^* = \frac{a_1}{N}\sum_{i=1}^{N}\xi_i.$$

Pro dané x_N^* dále počítejme

$$\begin{aligned} H(x_N^*) - H(x^*) &= E_\xi[h(x_N^*, \xi)] - E_\xi[h(x^*, \xi)] = \\ &= a_2E_\xi[(x_N^* - a_1\xi)^2 - (a_1\overline{\xi} - a_1\xi)^2] = \\ &= a_2E_\xi[(x_N^* - a_1\xi - a_1\overline{\xi} + a_1\xi)(x_N^* - a_1\xi + a_1\overline{\xi} - a_1\xi)] = \\ &= a_2(x_N^* - a_1\overline{\xi})E_\xi[x_N^* - 2a_1\xi + a_1\overline{\xi}] = \\ &= a_2(x_N^* - a_1\overline{\xi})^2. \end{aligned}$$

Víme, že pro dostatečně velké N dle centrální limitní věty platí

$$x_N^* = \frac{a_1}{N}\sum_{i=1}^{N}\xi_i \sim \frac{a_1\sigma_\xi}{\sqrt{N}}U + a_1\bar{\xi}$$

kde U má normální rozdělení $N(0{,}1)$. Potom platí:

$$\begin{aligned} P[H(x_N^*) - H(x^*) \geq \epsilon] &= P[a_2(x_N^* - a_1\bar{\xi})^2 \geq \epsilon] \cong \\ &\cong P[a_2(\frac{a_1\sigma_\xi}{\sqrt{N}}U + a_1\bar{\xi} - a_1\bar{\xi})^2 \geq \epsilon] = \\ &= P\left[U^2 \geq \frac{N\epsilon}{a_2 a_1^2\sigma_\xi^2}\right], \end{aligned}$$

kde U^2 má χ^2 rozdělení o jednom stupni volnosti. Nechť dále $M(\theta) = \frac{1}{\sqrt{1-2\theta}}$ je jeho momentová vytvořující funkce. Potom aplikací Chernoffovy meze (8) dostaneme

$$P\left[U^2 \geq \frac{N\epsilon}{a_2a_1^2\sigma_\xi^2}\right] \leq \exp\left\{-\frac{N\theta\epsilon}{a_2a_1^2\sigma_\xi^2}\right\}M(\theta) \leq \exp\left\{-\frac{N\theta\epsilon}{a_2a_1^2\sigma_\xi^2}\right\}\frac{1}{\sqrt{1-2\theta}}$$

pro $0 < \theta < 0.5$. $P[H(x_N^*) - H(x^*) \geq \epsilon]$ je tedy omezena funkcí $\exp\left\{-\frac{N\theta\epsilon}{a_2a_1^2\sigma_\xi^2}\right\}\frac{1}{\sqrt{1-2\theta}}$, která pro $N \to \infty$ klesá k 0. Pokud označíme

$$\alpha = \frac{1}{\sqrt{1-2\theta}}, \quad \beta = \frac{\theta\epsilon}{a_2a_1^2\sigma_{xi}^2},$$

máme omezení tvaru $\alpha \mathrm{e}^{-\beta N}$. Dále můžeme najít hranici pro $P[||x_N^* - x^*||_2 \geq \epsilon]$:

$$\begin{aligned} P[||x_N^* - x^*||_2 \geq \epsilon] &= P\left[\left|\frac{a_1\sigma_\xi}{\sqrt{N}}U\right| \geq \epsilon\right] = \\ &= 2P\left[U \geq \frac{\epsilon\sqrt{N}}{a_1\sigma_\xi}\right] = 2\exp\left\{-\frac{\epsilon^2 N}{2a_1^2\sigma_\xi^2}\right\}, \end{aligned}$$

kde

$$\alpha = 2, \quad \beta = \frac{\epsilon^2}{2a_1^2\sigma_\xi^2}.$$

□

Zatím jsme si pouze ukázali, že existují parametry α a β, které vystupují v exponenciální funkci, ale neznáme postup, který by výpočet těchto parametrů umožňoval v obecném případě. Věnujme se tomu tedy nyní.

Věta 11 (Shapiro a kol. (2009), str. 387–388)**.** Předpokládejme posloupnost $\{\xi_i, i = 1, \ldots, N\}$ nezávislých, stejně rozdělených veličin s $E_\xi[\xi_i] = \mu, i = 1, \ldots, N$ a momentovou vytvořující funkcí $M(t)$. Označme

$$\bar{\xi} = \frac{1}{N}\sum_{i=1}^{N}\xi_i$$

jejich aritmetický průměr. Potom pro $c \geq \mu$ platí

$$P[\overline{\xi} \geq c] \leq \mathrm{e}^{-I(c)N}, \tag{1.7}$$

kde výraz

$$I(c) = \sup_{t \in \mathbb{R}^1} \{tc - \log M(t)\},$$

značí Legendrovu transformaci momentové vytvořující funkce $M(t)$.

Důkaz. Pro každé reálné c a $t > 0$ platí $P[\overline{\xi} \geq c] = P[\mathrm{e}^{t\overline{\xi}} \geq \mathrm{e}^{tc}]$. Použitím Čebyševovy nerovnosti získáme

$$P[\overline{\xi} \geq c] \leq \mathrm{e}^{-tc} E_{\xi}[\mathrm{e}^{t\overline{\xi}}] = \mathrm{e}^{-tc} \left[M\left(\frac{t}{N}\right)\right]^N.$$

Jelikož je $c \geq \mu$, můžeme aplikovat přirozený logaritmus na obě strany nerovnice:

$$\log P[\overline{\xi} \geq c] \leq -tc + N \log \left[M\left(\frac{t}{N}\right)\right].$$

Označíme dále $t' = t/N$. Jelikož nerovnost platí pro libovolné t', bude platit i pokud pravou stranu zminimalizujeme přes všechna $t' > 0$. Můžeme tedy přepsat nerovnost do tvaru

$$\log P[\overline{\xi} \geq c] \leq -N \sup_{t' \in \mathbb{R}^1, t' > 0} [t'c - \log[M(t')] = -I(c)N.$$

Pokud nyní aplikujeme na předešlou nerovnost exponenciální funkci, je důkaz hotov. Pokud bychom zvolili $c \leq \mu$, dostaneme nerovnici

$$P[\overline{\xi} \leq c] \leq \mathrm{e}^{-I(c)N},$$

jejíž důkaz bychom provedli obdobně. $\square$

Příklad 8. Mějme veličinu ξ s normálním rozdělením $N(\mu,\sigma^2)$, tedy s momentovou vytvořující funkcí

$$M(t) = \exp(t\mu + \frac{1}{2}\sigma^2 t^2).$$

Snadno vypočítáme její Legendrovu transformaci:

$$\begin{aligned} I(c) &= \sup_{t \in \mathbb{R}^1} \{tc - \log M(t)\} = \sup_{t \in \mathbb{R}^1} \left\{ tc - \log \exp\left(t\mu + \frac{1}{2}\sigma^2 t^2\right)\right\} = \\ &= \sup_{t \in \mathbb{R}^1} \left\{ t(c - \mu) - \frac{1}{2}\sigma^2 t^2 \right\}. \end{aligned}$$

Výraz v posledním supremu je kvadratická, konkávní funkce, můžeme tedy její derivaci položit rovnou nule a dostaneme tak řešení

$$t = \frac{c - \mu}{\sigma^2}, \quad I(c) = \frac{1}{2}\frac{(c-\mu)^2}{\sigma^2}.$$

Nerovnost (1.7) tedy budeme mít pro normální rozdělení $N(\mu,\sigma^2)$ a $c \geq \mu$ ve tvaru

$$P[\overline{\xi} \geq c] \leq \exp\left\{-N\frac{1}{2}\frac{(c-\mu)^2}{\sigma^2}\right\}.$$

□

Ukažme si nyní podobné omezení, které je jednodušší na výpočet, ale platí pouze pro veličiny, které nabývají hodnot z intervalu $[a,b]$, tedy například rovnoměrné rozdělení $R[0,1]$.

Věta 12 (Shapiro a kol. (2009), P. 7.63., str. 389)**.** Nechť ξ je náhodná veličina nabývající hodnot pouze z intervalu $[a,b]$, kde $a,b \in \mathbb{R}^1$ a $E_\xi[\xi] = 0$. Potom pro vytvořující funkci veličiny ξ a všechna $t > 0$ platí nerovnost

$$M(t) \leq \mathrm{e}^{t^2(b-a)^2/8}.$$

Důkaz. Je založen na známé Madanského nerovnosti a čtenář ho může nalézt v Shapiro a kol. (2009). □

Uvažujeme-li nyní veličinu ξ se střední hodnotou μ, můžeme za použití předchozí věty odhadnout její Legendrovu transformaci

$$I(c) \geq \sup_{t\in\mathbb{R}^1,\, t>0} \{tc - t^2(b-a)^2/8\} = \frac{2c^2}{(b-a)^2}.$$

Pro náhodný výběr $\{\xi_i, i = 1,\ldots,N\}$ z veličiny ξ tedy použitím nerovnosti (1.7) a předešlé nakonec získáme pro $\tau > 0$ nerovnost

$$P[\overline{\xi} - \mu \geq \tau] = P[\overline{\xi} \geq \mu + \tau] \leq \exp\left\{-\frac{2N\tau^2}{(b-a)^2}\right\},$$

které se často říká Hoeffdingova. Pro rovnoměrné rozdělení $R[0,1]$ tak například získáme Hoeffdingovu nerovnost tvaru

$$P\left[\overline{\xi} \geq \frac{1}{2} + \tau\right] \leq \mathrm{e}^{-2N\tau^2}.$$

Získali jsem omezení pravděpodobnosti, že se průměr náhodného výběru neodchýlí od nějaké dané hodnoty. Jak dojít k přesnějšímu odhadu, alespoň z asymptotického hlediska, nám ukazuje následující věta.

Věta 13 (Dai a kol. (2000), T. 3.4., str. 498)**.** Nechť $\{\xi_i, i = 1,\ldots,N\}$ je posloupnost nezávislých, stejně rozdělených veličin s $E_\xi[\xi_i] = 0, i = 1,\ldots,N$ a konečnými momentovými vytvořujícími funkcemi $M(t) = E_\xi[\mathrm{e}^{t\xi_i}]$ pro všechna $t \in \mathbb{R}^1$. Předpokládejme, že $P[\xi_i > c] > 0$, $c > 0$. Potom

$$P\left[\overline{\xi} \geq c\right] \;\sim\; \frac{\gamma}{s\sqrt{2\pi N}}\mathrm{e}^{-I(c)N} \tag{1.8}$$

pro $N \to \infty$, kde $I(c)$ je Legendrova transformace momentové vytvořující funkce $M(t)$, s je dáno vzorcem

$$s = \sqrt{(M''(\theta)/M(\theta)) - c^2}$$

pro θ, které je řešením rovnice

$$I(c) = \theta c - \log M(\theta), \tag{1.9}$$

$$\gamma = \begin{cases} \dfrac{d}{1-\mathrm{e}^{-\theta d}} & \xi_i \text{ nabývá hodnot pouze z množiny} \\ & \{c + kd, k = 0, \pm 1, \pm 2, \dots\} \text{ pro nějaké } c \\ \theta^{-1} & \text{jinak.} \end{cases}$$

Příklad 9. Ukažme si nyní použití věty 13 na normální rozdělení, kde $\xi \sim N(\mu,\sigma^2)$. Abychom vyhověli předpokladu věty, uvažujme veličinu $\tilde{\xi} = \xi - \mu \sim N(0,\sigma^2)$, pro kterou budeme počítat výraz ve vztahu (1.8).

Z příkladu 8 už známe hodnotu $I(c) = \frac{1}{2}\frac{(c-\mu)^2}{\sigma^2}$ pro $\xi \sim N(\mu,\sigma^2)$. Zbavíme se tedy střední hodnoty a dosadíme do rovnice (1.9) pro výpočet parametru θ:

$$\begin{aligned} \frac{1}{2}\frac{c^2}{\sigma^2} &= \theta c - \log\exp\left(\frac{1}{2}\sigma^2\theta^2\right) \\ &= \theta c - \frac{1}{2}\sigma^2\theta^2, \end{aligned}$$

což nám dává

$$\theta^2\frac{1}{2}\sigma^2 - \theta c + \frac{1}{2}\frac{c^2}{\sigma^2} = 0$$

s řešením $\theta = c/\sigma^2$. Získáme tak zároveň i parametr $\gamma = \sigma^2/c$. Pokud dále dvakrát zderivujeme $M(\theta)$ podle θ, dostaneme

$$M''(\theta) = \left(\exp\left\{\frac{1}{2}\sigma^2\theta^2\right\}\right)'' = \exp\left\{\frac{1}{2}\sigma^2\theta^2\right\}(\sigma^4\theta^2 + \sigma^2).$$

Následně dosadíme vypočtenou hodnotu θ do rovnice pro parametr s a dostaneme

$$\begin{aligned} s &= \sqrt{(M''(\theta)/M(\theta)) - c^2} = \sqrt{\exp\left\{\frac{c^2}{2\sigma^2}\right\}(c^2+\sigma^2)/\exp\left\{\frac{c^2}{2\sigma^2}\right\} - c^2} = \\ &= \sqrt{(c^2+\sigma^2) - c^2} = \sigma. \end{aligned}$$

Dosazením do vztahu (1.8) tak pro normální rozdělení $N(\mu,\sigma^2)$ získáme

$$P\left[\bar{\tilde{\xi}} \geq c\right] = P\left[\bar{\xi} - \mu \geq c\right] \;\sim\; \frac{\sigma}{c\sqrt{2\pi N}}\exp\left\{-\frac{1}{2}\frac{c^2}{\sigma^2}N\right\}$$

pro $N \to \infty$. $\square$

Za pomoci věty 13 lze získat obdobu Hoeffdingovy nerovnosti i pro optimální hodnotu úlohy (1.2).

Věta 14 (Dai a kol. (2000), T. 3.5., str. 499)**.** Nechť

$$M(x,t) = E_\xi[\mathrm{e}^{t(h(x,\xi)-H(x))}]$$

je momentová vytvořující funkce $h(x,\xi) - H(x)$ a $I(x,c)$ je její Legendrova transformace ve tvaru

$$I(x,c) = \sup_{t\in\mathbb{R}^1}\{tc - \log M(x,t)\}$$

Nechť dále platí (A1) a $P[h(x,\xi) = H(x)] < 1$ pro každé $x \in U$. Pokud $M(x,t)$ je konečná pro všechna $t \in \mathbb{R}^1$ a všechna $x \in U$, potom

$$P[H(x_N^*)-H(x^*) \geq \epsilon] \leq \min_{x\in U}\left[\left(\frac{\gamma^+}{s^+\sqrt{2\pi N}}\right)^{-I^+(x,\epsilon/2)N} + \left(\frac{\gamma^-}{s^-\sqrt{2\pi N}}\right)^{-I^-(x,\epsilon/2)N}\right]$$

platí pro $N \to \infty$, kde I^+, γ^+, s^+ jsou stejné jako v (1.8) s tím, že $M(t)$ a c, jsou nahrazeny $M(x,t)$ a $\epsilon/2$. Veličiny I^-, γ^-, s^- získáme nahrazením $M(t)$ za $M(x,-t)$ a c za $\epsilon/2$.

Věta 15 (Dai a kol. (2000), T. 3.6., str. 500)**.** Předpokládejme, že všechny předpoklady věty 13 platí. Dále nechť $H(x)$ je dvakrát diferencovatelná a existuje konstanta k, $0 < k < \infty$, že $\nabla^2 H(x) \geq kI$ pro všechna x taková, že $||x - x^*||_2 \leq \delta_0, \delta_0 > 0$, kde I je jednotková matice a $A \leq B$ znamená, že $B - A$ je pozitivně semidefinitní.

Potom

$$P[||x_N^* - x^*||_2 \geq \epsilon] \leq \min_{x\in U}\left\{\left(\frac{\gamma^+}{s^+\sqrt{2\pi N}}\right)^{-I^+(x,\delta_\epsilon)N} + \left(\frac{\gamma^-}{s^-\sqrt{2\pi N}}\right)^{-I^-(x,\delta_\epsilon)N}\right\}$$

platí pro $N \to \infty$, kde $\delta_\epsilon = k\epsilon^2/4$ a ostatní proměnné jsou jako ve větě 13 kromě c, které je nahrazeno δ_ϵ.

Kapitola 2

Závislé posloupnosti

Dosud jsme se zabývali nezávislými výběry, resp. posloupnostmi nezávislých náhodných veličin. V praxi lze zpravidla na takovéto posloupnosti narazit v různých fyzikálních postupech nebo v laboratorních podmínkách.

Ve finanční a ekonomické praxi se ale obvykle zabýváme nejrůznějšími časovými řadami, které zpravidla považujeme za závislé v čase. Klasickým případem jsou jakákoli finanční data (cena akcií, dluhopisů, HDP, úrokové sazby atd.). Závislost ale můžeme najít např. i v demografii (předpokládaný věk dožití), v informačních technologiích (vytížení sítí) nebo i třeba v preferencích politických stran.

Také se nám může stát, že u nějaké posloupnosti nedokážeme s jistotou rozhodnout o její nezávislosti, a proto je vhodné mít k dispozici nástroje, které nám se závislými posloupnostmi dovolí pracovat.

Za určitých podmínek, tedy pro některé typy závislostí, lze upravit důkazy tvrzení pro posloupnosti nezávislé a získat tak podobné vlastnosti. Jedním z typů takových závislostí je tzv. m-závislost. Mluvíme potom o m-posloupnosti nebo také m-závislém procesu.

2.1 m-závislé posloupnosti

Definice 15 (m-závislý proces, Yoshihara (1992), str. 25)**.** Nechť $(\Omega, \mathbf{F}, P)$ je pravděpodobnostní prostor, $(X, \mathbf{B})$ měřitelný prostor a $\{\xi_t, t \in \mathbb{R}^1\}$ náhodný proces s diskrétním nebo spojitým časem, kde ξ_t jsou stejně rozdělené náhodné veličiny.

Nechť M_a^b je σ-algebra generovaná událostmi ve formě

$$\xi_{t_1} \in A_{t_1}, \ldots, \xi_{t_N} \in A_{t_N}, \quad a \leq t_1 \leq \cdots \leq t_N \leq b$$

kde $t_1, \ldots, t_N$ a N jsou libovolné a $A_{t_1}, \ldots, A_{t_N}$ jsou $\mathcal{B}$-měřitelné množiny. Říkáme, že proces $\{\xi_t\}$ je m-závislý pokud $M_{-\infty}^a$ a M_b^∞ jsou nezávislé pro $b-a > m$.

Nyní si odvodíme podobnou nerovnost jako ve větě 9.

Věta 16. Předpokládejme, že platí (A1) a nechť posloupnost $\{\xi_i, i = 1, \ldots, N\}$ nezávislých, stejně rozdělených náhodných veličin je m-závislá.

Potom pro jakékoli $\epsilon > 0$ existují $\alpha > 0$ a $\beta > 0$ takové, že

$$P[H(x_N^*) - H(x^*) \geq \epsilon] \leq \alpha r \mathrm{e}^{-\beta(k+1)} + \alpha(m-r)\mathrm{e}^{-\beta k}$$

pro všechna $N > 0$, kde k,r splňují $N = km+r$ pro k přirozené a $r \in \{0, 1, \ldots, m-1\}$.

Důkaz. Pozorný čtenář si jistě povšiml, že se věta 16 velmi podobá větě 9, ve které předpokládáme posloupnost $\{\xi_i, i = 1, \ldots, N\}$ nezávislých náhodných veličin. Vypůjčíme si tedy i postup důkazu věty 9 a upravíme ho tak, aby platil pro posloupnost $\{\xi_i, i = 1, \ldots, N\}$, která je m-závislá.

Jelikož lze každé N přirozené rozložit na $N = km + r$, kde k je přirozené a $r \in \{0, 1, \ldots, m-1\}$, můžeme si vytvořit následující posloupnosti:

$$\begin{array}{lllll}
 & \xi_m, & \xi_{2m}, & \cdots & \xi_{km} \\
\xi_1, & \xi_{m+1}, & \xi_{2m+1}, & \cdots & \xi_{km+1} \\
\xi_2, & \xi_{m+2}, & \xi_{2m+2}, & \cdots & \xi_{km+2} \\
\vdots & & & & \\
\xi_r, & \xi_{m+r}, & \xi_{2m+r}, & \cdots & \xi_{km+r} \\
\xi_{r+1}, & \xi_{m+r+1}, & \xi_{2m+r+1}, & \cdots & \xi_{(k-1)m+r+1} \\
\vdots & & & & \\
\xi_{m-1}, & \xi_{2m-1}, & \xi_{3m-1}, & \cdots & \xi_{km-1}
\end{array} \tag{2.1}$$

Máme tedy r posloupností o $k+1$ prvcích a $m-r$ posloupností o k prvcích, kdy každá z nich je posloupnost nezávislých náhodných veličin, díky čemuž můžeme využít postupu z důkazu věty 9.

Nejdříve ale budeme muset upravit následující výraz tak, abychom získali výraz pro nezávislé posloupnosti, na které se dá zmíněný postup použít:

$$P[|H_N(x) - H(x)| \geq \delta_\epsilon] = P\left[\left|\frac{1}{N}\sum_{i=1}^{N} h(x,\xi_i) - H(x)\right| \geq \delta_\epsilon\right],$$

kde $\delta_\epsilon = \epsilon/2$. Nyní využijeme rozdělení posloupnosti $\xi_i, i = 1, \ldots, N$ na podposloupnosti (2.1):

$$P\left[\left|\frac{1}{N}\sum_{i=1}^{N}(h(x,\xi_i)-H(x))\right|\geq\delta_\epsilon\right]\leq \tag{2.2}$$

$$\leq\sum_{j=1}^{r}P\left[\left|\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\right|\geq\frac{\delta_\epsilon N}{m}\right]+ \tag{2.3}$$

$$+\sum_{j=r+1}^{m}P\left[\left|\sum_{i=0}^{k-1}(h(x,\xi_{mi+j})-H(x))\right|\geq\frac{\delta_\epsilon N}{m}\right]. \tag{2.4}$$

Zaměříme se nyní na výraz (2.3), tedy na r posloupností o $k+1$ prvcích. Rozložíme si ho opět na dva výrazy a zbavíme se tak absolutní hodnoty:

$$\sum_{j=1}^{r}P\left[\left|\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\right|\geq\frac{\delta_\epsilon N}{m}\right]= \tag{2.5}$$

$$=\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\geq\frac{\delta_\epsilon N}{m}\right]+ \tag{2.6}$$

$$+\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\leq-\frac{\delta_\epsilon N}{m}\right]. \tag{2.7}$$

Opět se zaměříme pouze na jeden výraz, a to na (2.6). Využijeme toho, že $N=km+r$ a tedy $\delta_\epsilon\frac{N}{m}=\delta_\epsilon\frac{km+r}{m}=\delta_\epsilon(k+\frac{r}{m})<\delta_\epsilon(k+1)$. Díky tomu můžeme psát:

$$\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\geq\frac{\delta_\epsilon N}{m}\right]\leq$$

$$\leq\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\geq\delta_\epsilon(k+1)\right]=$$

$$=\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x)-\delta_\epsilon)\geq 0\right].$$

V tuto chvíli jsme dostali výraz, který nám dovoluje použít postupu z důkazu věty 9. Jedná se totiž o sumu pravděpodobností, které lze pomocí Chernoffovy meze (věta 8) omezit exponenciální funkcí ve tvaru $\mathrm{e}^{-\beta(k+1)}$. Můžeme tedy psát

$$\sum_{j=1}^{r}P\left[\sum_{i=0}^{k}(h(x,\xi_{mi+j})-H(x))\geq\frac{\delta_\epsilon N}{m}\right]\leq r\mathrm{e}^{-\beta(k+1)}.$$

Podobně můřeme ukázat i stejnou nerovnost pro (2.7). Ve výsledku bychom tedy mohli omezit výraz (2.5) funkcí $2r\mathrm{e}^{-\beta(k+1)}$.

Podobným postupem bychom mohli získat i nerovnost pro výraz (2.4), tedy $m-r$ posloupností o k prvcích, který bychom mohli omezit funkcí $2(m-r)\mathrm{e}^{-\beta k}$. To nám ve výsledku dává:

$$P[|H_N(x) - H(x)| \geq \delta_\epsilon] \leq 2r\mathrm{e}^{-\beta(k+1)} + 2(m-r)\mathrm{e}^{-\beta k}. \tag{2.8}$$

Nyní využijeme vztahu z důkazu věty 9 – viz Dai a kol. (2000), T. 3.1., str. 496, rovnice (18), což nám v kombinaci s (2.8) dává konečně

$$P[H(x_N^*) - H(x^*) \geq \epsilon] \leq 4r\mathrm{e}^{-\beta(k+1)} + 4(m-r)\mathrm{e}^{-\beta k}.$$

Tím je důkaz hotov pro $\alpha = 4$ a $\beta > 0$. □

2.2 Další typy závislosti – mixingy

Kromě m-závislosti mohou být posloupnosti závislé i jiným způsobem. Uveďme si nyní přehled podmínek nazývané mixingy, které tuto závislost popisují.

Pro následující definice označme M_1 a M_2 jako dvě libovolné σ-algebry z podmnožin Ω a nechť M_a^b je σ-algebra generovaná $\{\xi_t, t = a, a+1, \ldots, b-1, b\}$, kde $\{\xi_t, t \in \mathbb{Z}\}$ je posloupnost náhodných veličin.

Definice 16 (ψ-mixing, Yoshihara (1992), str. 25)**.** Nechť

$$\psi(M_1,M_2) = \sup_{A\in M_1, B\in M_2} \left| \frac{P(A\cap B)}{P(A)P(B)} - 1 \right|.$$

Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je ψ-mixing, pokud

$$\psi(t) = \sup_s \psi(M_{-\infty}^s, M_{s+t}^\infty) \to 0$$

pro $t \to \infty$. ψ-mixing se občas nazývá též *-mixing.

Tomuto mixingu se podobají i následující dva mixingy:

Definice 17 (ψ^*-mixing, ψ'-mixing, Bradley (2005), str. 108)**.** Nechť

$$\psi^*(M_1,M_2) = \sup \left\{ \frac{P(A\cap B)}{P(A)P(B)} : A \in M_1, B \in M_2, P(A)P(B) > 0 \right\},$$

$$\psi'(M_1,M_2) = \inf \left\{ \frac{P(A\cap B)}{P(A)P(B)} : A \in M_1, B \in M_2, P(A)P(B) > 0 \right\}.$$

Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je ψ^*-mixing (resp. ψ'-mixing), pokud

$$\psi^*(t) = \sup_s \psi^*(M_{-\infty}^s, M_{s+t}^\infty) \to 1$$

$$\Big(\text{resp. } \psi'(t) = \inf_s \psi'(M_{-\infty}^s, M_{s+t}^\infty) \to 1\Big)$$

pro $t \to \infty$.

Definice 18 (ϕ-mixing, Yoshihara (1992), str. 27)**.** Nechť

$$\phi(M_1,M_2) = \sup_{A \in M_1, B \in M_2} \left| \frac{P(A \cap B) - P(A)P(B)}{P(A)P(B)} \right|.$$

Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je ϕ-mixing, pokud

$$\phi(t) = \sup_s \phi(M_{-\infty}^s, M_{s+t}^\infty) \to 0 \text{ pro } t \to \infty.$$

Označme $\sigma(M_1 \cup M_2)$ σ-algebru generovanou sjednocením σ-algeber M_1 a M_2.

Definice 19 (γ_1-, γ_2-mixing, Yoshihara (1992), str. 29)**.** Říkáme, že $\{\xi_t, t \in \mathbb{Z}\}$ splňuje γ_1-mixing (resp. γ_2-mixing) podmínku s koeficientem $\gamma_1(t)$ (resp. $\gamma_2(t)$), pokud existuje nerostoucí funkce $\gamma_1(t)$ (resp. $\gamma_2(t)$) taková, že pro všechna reálná čísla a, b a reálné $t \geq 0$, a $A \in \sigma(M_{-\infty}^a \cup M_{b+t}^\infty)$, $B \in \sigma(M_{a+t}^b)$ platí

$$|P(A \cap B) - P(A)P(B)| \leq \gamma_1(t)P(B).$$

$$(\text{resp } |P(A \cap B) - P(A)P(B)| \leq \gamma_2(t)P(A)P(B)) \,.$$

Definice 20 (absolutně regulární podmínka, Yoshihara (1992), str. 31)**.** Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je absolutně regulární s $\beta(t)$ pokud

$$\beta(t) = \sup_s E_\xi \left(\sup_{B \in M_{s+t}^\infty} |P(B|M_{-\infty}^s) - P(B)| \right) \to 0$$

pro $t \to \infty$. Tato posloupnost se také nazývá β-mixing.

Definice 21 (strong mixing podmínka, Yoshihara (1992), str. 36)**.** Nechť

$$\alpha(M_1,M_2) = \sup_{A \in M_1, B \in M_2} |P(A \cap B) - P(A)P(B)|.$$

Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je strong mixing pokud

$$\alpha(t) = \sup_s \alpha(M_{-\infty}^s, M_{s+t}^\infty) \to 0$$

pro $t \to \infty$. Tato posloupnost se také nazývá α-mixing.

Definice 22 (ρ-mixing, Bradley (2005), str. 108)**.** Nechť $\mathcal{L}^2(\mathcal{A})$ značí prostor dvakrát integrovatelných, $\mathcal{A}$-měřitelných reálných náhodných veličin a

$$\rho(M_1,M_2) = \sup_{u \in \mathcal{L}^2(M_1), v \in \mathcal{L}^2(M_2)} |\mathrm{cor}(u,v)|,$$

kde cor značí korelaci. Říkáme, že posloupnost $\{\xi_t, t \in \mathbb{Z}\}$ je ρ-mixing pokud

$$\rho(t) = \sup_s \rho(M_{-\infty}^s, M_{s+t}^\infty) \to 0 \text{ pro } t \to \infty.$$

Obecně mixingy mezi sebou nejsou ekvivalentní, za specifických podmínek ale ekvivalence platí, jak ukazuje následující věta.

Věta 17 (Yoshihara (1992), str. 29)**.** Pro Gaussovský, silně stacionární proces platí, že následující vlastnosti jsou ekvivalentní:

(i) $\{\xi_t, t \in \mathbb{Z}\}$ je m-závislý,

(ii) $\{\xi_t, t \in \mathbb{Z}\}$ je ψ-mixing,

(iii) $\{\xi_t, t \in \mathbb{Z}\}$ je ϕ-mixing.

V obecném případě tomu ale tak není a mezi mixingy existuje „řetězec“ implikací, které jsme sestavili do následujícího přehledu na základě prací Yoshihara (1992) a Bradley (2005).

$$
\begin{array}{cccccccc}
 & \nearrow & \psi^*\text{-mixing} & \searrow & & \nearrow & \rho\text{-mixing} & \searrow \\
m\text{-závislost} \rightarrow \psi\text{-mixing} & \rightarrow & \psi'\text{-mixing} & \rightarrow & \phi\text{-mixing} & & & \\
 & \searrow & \gamma_1\text{-mixing} & \nearrow & & \searrow & \gamma_2\text{-mixing} & \nearrow
\end{array}
$$

$$\rightarrow \text{absolutní regularita} \rightarrow \text{strong mixing.}$$

Pomocí mixingů a m-posloupností jsme ale bezpochyby nepokryli veškeré typy závislostí, takže se nyní podíváme na pojem slabá závislost, díky kterému lze studovat i další velkou skupinu závislostí.

2.3 Slabá závislost

Slabě závislou posloupnost si můžeme představit jako posloupnost, ve které platí asymptotická nezávislost mezi „minulostí“ a „budoucností“. To znamená, že kovariance

$$\mathrm{Cov}(h(\text{„minulost“}), k(\text{„budoucnost“}))$$

je malá pro dostatečně velkou „vzdálenost“ mezi „minulostí“ a „budoucností“, kde h a k jsou libovolné vhodné funkce. Pro pořádek uveďme přesnou definici.

Definice 23 (slabá závislost, Doukhan a Louhichi (1999), D. 1, str. 315)**.** Náhodnou posloupnost $\{\xi_t, t = 1, 2, \dots\}$ nazveme $(\epsilon, \mathcal{F}, \psi)$-slabě závislou, pokud existuje třída $\mathcal{F}$ reálných funkcí, posloupnost $\{\epsilon_t, t = 1, 2, \dots\}$ klesající k nule pro $t \to \infty$, funkce $\psi(h,k,u,v)$ s argumenty $h,k \in \mathcal{F}$, $u,v \in \{1,2,\dots\}$ takovými, že pro všechny vektory $(i_1, \dots, i_u), (j_1, \dots, j_v)$ s $i_1 \leq \dots \leq i_u < i_u + r \leq j_1 \leq \dots \leq j_v$ platí

$$|\mathrm{Cov}(h(\xi_{i_1}, \dots, \xi_{i_u}), k(\xi_{i_1}, \dots, \xi_{i_v}))| \leq \psi(h,k,u,v)\epsilon_r$$

pro všechny funkce $h,k \in \mathcal{F}$, které jsou definovány na $\mathbb{R}^u$ a $\mathbb{R}^v$.

Koeficient r můžeme chápat jako (časovou) „vzdálenost" mezi „minulostí" a „budoucností".

2.3.1 Třídy slabé závislosti

Pokud v definici 23 dosadíme za funkci ψ nějakou konkrétní funkci, získáme (jak je uvedeno v Doukhan (2008), str. 148) různé třídy slabé závislosti:

- κ-slabě závislá posloupnost, pokud $\psi(a,b,i,j) = ijab$,
- κ'-slabě závislá posloupnost, pokud $\psi(a,b,i,j) = jab$,
- η-slabě závislá posloupnost, pokud $\psi(a,b,i,j) = ia + jb$,
- θ-slabě závislá posloupnost, pokud $\psi(a,b,i,j) = jb$,
- λ-slabě závislá posloupnost, pokud $\psi(a,b,i,j) = ijab + ia + jb$.

Pro tyto třídy závislostí můžeme dále v textu značit posloupnost $\{\epsilon_t, t \in \mathbb{Z}\}$ jako $\{\kappa_t, t \in N\}$, $\{\eta_t, t \in \mathbb{Z}\}$ apod.

2.3.2 Příklady slabě závislých posloupností

Uveďme si nyní příklady některých známějších posloupností, které využijeme dále v textu včetně jejich vlastností. V některých případech se může stát, že posloupnost nesplňuje žádnou z mixingových podmínek, ale zároveň závislá je. V takovém případě se může hodit $(\epsilon, \mathcal{F}, \psi)$-slabá závislost, kterou taková posloupnost už může splňovat.

Než si uvedeme některé příklady, budeme potřebovat následující definici:

Definice 24 (Lipschitzovský modulus spojitosti, Doukhan a Louhichi (1999), str. 316)**.** Nechť $h : \mathbb{R}^n \to \mathbb{R}^1$. Potom definujme výraz

$$\mathrm{Lip}(h) = \sup_{x \neq y} \frac{|h(x) - h(y)|}{||x - y||_1}$$

jako Lipschitzovský modulus spojitosti.

Nechť dále $\mathcal{L}$ značí množinu omezených Lipschitzovských funkcí.

Bernoulliho shift

Bernoulliho shift ve tvaru uvedeném níže je právě příkladem posloupnosti, která nesplňuje m-závislost ani další z mixingových podmínek, ale jedná se o slabě závislou posloupnost.

Definice 25 (Bernoulliho shift: Doukhan a Louhichi (1999), str. 324, D. 4)**.** Nechť $\{\xi_i, i \in \mathbb{Z}\}$ je posloupnost nezávislých reálných náhodných veličin a F je měřitelná funkce definovaná na $\mathbb{R}^{\mathbb{Z}}$. Pak Bernoulliho shift je posloupnost $\{U_i, i \in \mathbb{Z}\}$ definovaná vzorcem

$$U_i = F(\xi_{i-j}, j \in \mathbb{Z}).$$

Zatímco klouzavý průměr $F(\xi_t, \xi_{t-1}, \ldots, \xi_{t-m})$ je m-závislá posloupnost, Bernoulliho shift už nesplňuje m-závislost a ani žádnou další mixingovou podmínku, jak uvádí Doukhan a Louhichi (1999) (str. 325). Naopak platí (Doukhan a Louhichi (1999), L. 9, str. 325), že posloupnost $\{(U_t - E_\xi U_t), t \in \mathbb{Z}\}$ je $(\theta, \mathcal{L}, \psi)$-slabě závislá s

$$\psi(h,k,u,v) = 4(u||k||_\infty \mathrm{Lip}(h) + v||h||_\infty \mathrm{Lip}(k)) \quad \text{a} \quad \epsilon_r = \delta_{r/2},$$

kde definujeme δ_k pro každé kladné k takto:

$$\delta_k = \sup_{i \in \mathbb{Z}} E_\xi |F(\xi_{i-j}, j \in \mathbb{Z}) - F(\xi_{i-j}\mathbb{1}_{|j|<k}, j \in \mathbb{Z})|,$$

a $||h||_\infty$ značí normu ve tvaru $\sup_{x\in[0,1]} |h(x)|$.

Konkrétním příkladem, který uvádí Doukhan a Louhichi (1999) (str. 325) je AR(1) proces definovaný takto

$$U_i = aU_{i-1} + \xi_i = \sum_{j \geq 0} a^j \xi_{i-j},$$

kde $a \in (0, 1/2]$ a $\{\xi_i, i \in \mathbb{Z}\}$ je výběr z Bernoulliho rozdělení s parametrem p. Tento proces je slabě závislý s $F(u_i, j \in \mathbb{Z}) = \sum_{i \geq 0} a^i u_i$ a $\delta_k = p \sum_{i>k} a^i$.

Markovovy řetězce

Markovův řetězec už splňuje jak některé z mixingových podmínek, tak i (ϵ, $\mathcal{F}$, ψ)-slabou závislost. Nechť $\{\xi_i, i \in \mathbb{Z}\}$ je posloupnost nezávislých reálných náhodných veličin a F měřitelná funkce. Definujme $\{X_i, i \in N\}$ Markovův řetězec pomocí vztahu

$$X_{t+1} = F(X_t, \xi_{t+1}). \tag{2.9}$$

Jak zmiňuje např. Bradley (2005) (sekce 3.1., str. 118), Markovův řetězec je strong mixing posloupnost s

$$\alpha(t) = \sup_{j \in \mathbb{Z}} \alpha(\sigma(X_j), \sigma(X_{j+t})),$$

kde $\sigma(X_j)$ značí sigma algebru generovanou náhodnou veličinou X_j. Pro striktně stacionární proces se tento výraz zřejmě zjednoduší na

$$\alpha(t) = \alpha(\sigma(X_0),\sigma(X_t)).$$

Pokud se jedná navíc o proces s konečným počtem stavů, irreducibilní a aperiodický, potom splňuje i ψ-mixing. Ten implikuje i ostatní mixingy zmíněné v předchozí kapitole.

Dále, jak je uvedeno v Doukhan a Louhichi (1999) (sekce 3.5.), je Markovův řetězec $(\epsilon,\mathcal{L},\psi)$-slabě závislá posloupnost s

$$\psi(h,k,u,v) = 2\min(u||k||_\infty \mathrm{Lip}(h), v||h||_\infty \mathrm{Lip}(k)) \quad \text{a} \quad \epsilon_r = \alpha^r E|X_0|,$$

pokud předpokládáme F ze vztahu (2.9) takovou, že splňuje

$$E|F(0,\xi_0)|^b < \infty \quad \text{a} \quad E|F(x,\xi_0) - F(y,\xi_0)|^b \leq \beta^b|x-y|^b$$

pro nějaké $b \geq 1$ a $0 \leq \beta < 1$.

ARMA proces

Nechť $\{X_t, t \in \mathbb{Z}\}$ je ARMA(p,q) proces takový, který lze zapsat následující rovnicí

$$X_t = \sum_{k=-\infty}^{\infty} a_k \xi_{t-k} \text{ pro } t \in \mathbb{Z},$$

kde $\{a_k, k \in \mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}}$ a $\{\xi_i, i \in \mathbb{Z}\}$ je posloupnost nezávislých náhodných veličin s nulovou střední hodnotou. Jak je uvedeno v Doukhan (2008) (sekce 3, str 154), $\{X_t, t \in \mathbb{Z}\}$ je η-slabě závislá posloupnost s $\eta_r = O\left(\frac{1}{\mathbb{R}^{\mu-1/2}}\right)$, pokud platí $a_k = O(|k|^{-\mu})$, $\mu > 1/2$. Podotkněme, že za určitých podmínek ARMA proces splňuje i absolutně regulární podmínku, viz Doukhan (1994), str. 99, T. 6.

GARCH proces

Nechť $\{X_t, t \in \mathbb{Z}\}$ je GARCH(p,q) proces definovaný rovnicemi

$$X_t = \rho_t \xi_t \quad \text{s} \quad \rho_t^2 = b_0 + \sum_{k=1}^{\infty} b_k X_{t-k}^2 \text{ pro } k \in \mathbb{Z},$$

kde $\{b_k, k \in \mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}}$ závisí na počátečních parametrech a $\{\xi_i, i \in \mathbb{Z}\}$ je posloupnost nezávislých náhodných veličin.

Jak je uvedeno v Doukhan (2008) (sekce 3, str 154), $\{X_t, t \in \mathbb{Z}\}$ je θ-slabě závislý proces s $\theta_r = O(\mathrm{e}^{-c\sqrt{r}})$, $c > 0$, pokud platí, že $E(|\xi_0|^m) < \infty$ a $E(\xi_i)$ je stejná pro všechna $i \in \mathbb{Z}$, $||\xi_0||_m^2 \sum_{j=1}^{\infty} |b_j| < 1$, a pokud existuje $C > 0$ a $\mu \in (0,1)$ tak, že $\forall j \in N, 0 \leq b_k \leq C\mu^{-j}$. Podotkněme, že za určitých podmínek je GARCH proces strong mixing, viz Lindner (2009), str. 62.

2.4 Konvergence odhadů pro závislé posloupnosti

Podobně jako jsme se zabývali konvergencí nezávislých posloupností (viz např. věta 9) a m-posloupností (viz věta 16), můžeme za určitých podmínek dokázat i konvergenci posloupností mixingových či slabě závislých. Uveďme si nejdříve větu zaručující konvergenci pro součet strong, resp. ϕ-mixingového výběru.

2.4.1 Mixing posloupnosti

Věta 18 (Doukhan (1994), P. 1, str. 33)**.** Předpokládejme, že $\{\xi_t, t \in \mathbb{Z}\}$ je mixing posloupnost centrovaných náhodných veličin, $|\xi_t| \leq I$. Dále nechť se jedná o strong mixing posloupnost (resp. ϕ-mixing posloupnost) splňující $\alpha(n) \leq uv^n$ (resp. $\phi(n) \leq uv^n$) pro nějaké $u > 0, 0 \leq v < I$. Pokud dále N splňuje $N \sup_t (E_\xi[|\xi_t|^2])^{2/(2+\gamma)} \geq I$ pro $\gamma > 0$ (resp. $N \sup_t E_\xi[|\xi_t|^2] \geq I$), potom existují konstanty $a, b > 0$ takové, že pro libovolné $y \geq 0$ platí

$$P\left(\left|\sum_{t=1}^{N} \xi_t\right| \geq y\sqrt{N}\sigma \log(\sigma^{-I})\right) \leq a\mathrm{e}^{-b\sqrt{y}}$$

$$\left(\text{resp. } P\left(\left|\sum_{t=1}^{N} \xi_t\right| \geq y\sqrt{N}\sigma\right) \leq a\mathrm{e}^{-b\sqrt{y}}\right),$$

kde $\sigma^2 = \sup_t E_\xi[|\xi_t|^{2+\gamma}])^{2/(2+\gamma)}$, (resp. $\sigma^2 = \sup_t E_\xi[|\xi_t|^2]$).

Z hlediska optimalizačních úloh by nás ale podobně jako ve větě 9) a 16 více zajímala konvergence výrazu $P[|H(x_N^*) - H(x^*)| \geq \epsilon]$. Abychom mohli konvergenci tohoto výrazu snadno ukázat, uvažujme méně obecný tvar funkce $h(x,\xi)$. Předpokládejme, že existuje přirozené číslo s a spojité a omezené funkce $h_j^*(x), g_j^*(\xi), j = 1, \ldots, s$ takové, že $h(x,\xi)$ ve tvaru

$$h(x,\xi) = \sum_{j=1}^{s} h_j^*(x) g_j^*(\xi).$$

Pro funkci tohoto tvaru můžeme nyní zformulovat následující větu:

Věta 19. Předpokládejme, že $\{\xi_t, t \in \mathbb{Z}\}$ je strong mixing posloupnost centrovaných náhodných veličin splňující předpoklady věty 18, $h(x,\xi)$ je ve tvaru popsaném výše a $|h(x,\xi_t)| \leq I$. Pokud dále N splňuje $N \sup_t (E_\xi[|h(x,\xi_t)|^2])^{2/(2+\gamma)} \geq I$ pro $\gamma > 0$, potom existují konstanty $a, b > 0$ takové, že pro libovolné $\epsilon \geq 0$ platí

$$P[H(x_N^*) - H(x^*) \geq y\sqrt{N}\sigma \log(\sigma^{-I})] \leq a\mathrm{e}^{-b\sqrt{y}},$$

kde $\sigma^2 = \sup_t E_\xi[|h(x^*,\xi_t)|^{2+\gamma}])^{2/(2+\gamma)}$.

Důkaz. Vzhledem ke tvaru funkce $h(x,\xi)$ můžeme psát

$$H(x) = \sum_{j=1}^{s} h_j^*(x) E_\xi[g_j^*(\xi)] \quad \text{a} \quad H_N(x) = \sum_{j=1}^{s} h_j^*(x) \frac{1}{N} \sum_{i=1}^{N} g_j^*(\xi_i).$$

Rozepíšeme tedy následující výraz:

$$P[|H_N(x) - H(x)| \geq \epsilon] = P\left[\left|\sum_{j=1}^{s} h_j^*(x) \frac{1}{N} \left(\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right)\right| \geq \epsilon\right].$$

Podle předpokladů věty existuje M takové, že můžeme nahradit funkci $h_j^*(x)$ konstantou: $|h_j^*(x)| \leq M, \forall j = 1, \ldots, s$, a pokud zároveň místo N uvedeme $\sqrt{N}$, můžeme dále pro každé x psát

$$P[|H_N(x) - H(x)| \geq \epsilon] \leq P\left[\left|\sum_{j=1}^{s} \frac{M}{\sqrt{N}} \left(\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right)\right| \geq \epsilon\right].$$

V úpravě výrazu dále pokračujeme použitím známého triku, kdy výraz rozložíme na sumu pravděpodobností, čímž dostaneme:

$$\begin{aligned} P\left[\left|\sum_{j=1}^{s} \frac{M}{\sqrt{N}} \left(\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right)\right| \geq \epsilon\right] \leq \\ \leq \sum_{j=1}^{s} P\left[\left|\frac{M}{\sqrt{N}} \left(\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right)\right| \geq \frac{\epsilon}{s}\right] = \\ = \sum_{j=1}^{s} P\left[\left|\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right| \geq \frac{\sqrt{N}}{M}\frac{\epsilon}{s}\right]. \end{aligned}$$

Zvolíme-li nyní vhodně ϵ, tedy $\epsilon = y\sigma \log(\sigma^{-I}) Ms$ dostaneme uvnitř vnější sumy výraz podobný tvrzení z věty 18. Uvnitř absolutní hodnoty je centrovaná náhodná veličina, která splňuje mixingovou podmínku, a můžeme tedy podle věty 18 psát

$$= \sum_{j=1}^{s} P\left[\left|\left(\sum_{i=1}^{N} g_j^*(\xi_i) - E_\xi[g_j^*(\xi)]\right)\right| \geq y\sqrt{N}\sigma \log(\sigma^{-I})\right] \leq \sum_{j=1}^{s} a_j \mathrm{e}^{-b_j\sqrt{y}}.$$

Pokud dále vhodně zvolíme parametry a^*, b tak, aby byl součet exponenciálních funkcí omezen pouze jedinou exponenciální funkcí, lze výraz zjednodušit na tvar

$$P[|H_N(x) - H(x)| \geq y\sqrt{N}\sigma \log(\sigma^{-I})] \leq a^* \mathrm{e}^{-b\sqrt{y}}.$$

Jelikož nás ale zajímá nerovnost pro výraz, ve kterém máme $H(x_N^*) - H(x^*)$, použijeme triku z důkazu věty 16 (viz 9 – viz Dai a kol. (2000), T. 3.1., str. 496, rovnice (18)) a dostaneme tak pro $a = 2a^*$ nerovnost

$$P[H(x_N^*) - H(x^*) \geq y\sqrt{N}\sigma \log(\sigma^{-I})] \leq a\mathrm{e}^{-b\sqrt{y}}.$$

$\square$

Poznámka 1. Podotkněme, že znění věty bylo formulováno pro náhodné veličiny a je relativně slabé. Vzhledem ke tvaru funkce $h(x,\xi)$ by ale bylo snadné větu zobecnit i na náhodné vektory.

2.4.2 Slabě závislé posloupnosti

Podobně jako je zaručena konvergence mixingové posloupnosti ve větě 18, existuje i tvrzení pro slabě závislé posloupnosti, které si zde pro úplnost uvedeme. Nebudeme se ale dále věnovat jeho použití na optimalizační úlohu. Než uvedeme přímo tvrzení, zaveďme jednu potřebnou definici.

Definice 26 (koeficient slabé závislosti, Doukhan a Louhichi (1999), D. 2, str. 318)**.** Nechť $\{\xi_t, t = 1, 2, \dots\}$ je posloupnost centrovaných náhodných veličin. Pro kladné celé číslo r definujme koeficient slabé závislosti jako posloupnost $(C_{r,q})_{q\geq 2}$ takovou, že

$$C_{r,q} := \sup |\mathrm{Cov}(\xi_{t_1}, \dots, \xi_{t_m}, \xi_{t_{m+1}}, \dots, \xi_{t_q})|,$$

kde supremum je přes všechny $\{t_1, \dots, t_q\}$ takové, že $1 \leq t_1 \leq \dots \leq t_q$, m,r splňují rovnici $t_{m+1} - t_m = r$ a symbol $|\mathrm{Cov}(\xi_{t_1}, \dots, \xi_{t_m}, \xi_{t_{m+1}}, \dots, \xi_{t_q})|$ značí $\sup_{i,j} |\mathrm{Cov}(\xi_{t_i}, \xi_{t_j})|$, kde $i \in [1,m]$ a $j \in [m+1,q]$.

Věta 20 (Doukhan a Louhichi (1999), C. 1, str. 320)**.** Nechť $\{\xi_t, t = 1, 2, \dots\}$ je $(\epsilon, \mathcal{F}, \psi)$-slabě závislá posloupnost taková, že pro celá čísla $q \geq 2, p \leq q$ a konstanty M, γ, C platí

$$C_{r,p} \leq C\mathrm{e}^{\gamma p} M^{p-2} \epsilon_r.$$

Nechť dále existuje posloupnost $\{a_t, t = 1, 2, \dots\}$, $a_t \geq 1, t > 1$, taková, že pro jakákoli celá čísla t a $q \geq 2$ a konstantu $\beta > 0$ platí

$$t \sum_{r=0}^{t-1} (r+1)^{q-2} C_{r,q} \leq a_t \frac{q!}{\beta^q},$$

kde q je nezávislé na $\{a_t, t = 1, 2, \dots\}$. Potom existují konstanty $A,B > 0$ takové, že pro libovolné $y > 0$ platí

$$P\left(\left|\sum_{t=1}^{N} \xi_t\right| \geq y\sqrt{a_t}\right) \leq A\mathrm{e}^{-B\sqrt{\beta y}}.$$

Čtenář si může snadno povšimnout podobnosti s větou 18, tedy že pravděpodobnost je omezena exponenciální funkcí. Z hlediska konvergence lze tedy očekávat podobné chování jako u mixingových posloupností, jak bude ukázáno na numerických příkladech dále v textu.

2.4.3 „Rychlost“ konvergence

Podívejme se nyní na to, za jakých okolností je ještě zajištěna konvergence, tedy pro jaké ν, $N \to \infty$ a $\epsilon > 0$ platí

$$P[N^\nu |H(x_N^*) - H(x^*)| \geq \epsilon] \to 0.$$

Připomeňme, že výrazy $P[|H(x_N^*) - H(x^*)| \geq \epsilon]$ a $P[|H_N(x_N^*) - H(x^*)| \geq \epsilon]$ jsou až na konstantu z hlediska konvergence ekvivalentní (viz důkaz věty 9 – Dai a kol. (2000), T. 3.1., str. 496, rovnice (18)), nedopouštíme se tedy nepřesnosti, pokud je některá z vět uvedena v původním textu s jiným výrazem, než uvádíme v této práci.

Věta 21 (Kaňková (1993), T. 3.1, str. 105)**.** Nechť χ je kompaktní množina a funkci $h(x,\xi)$ lze zapsat ve tvaru $h(x,\xi) = \sum_{j=1}^{s} h_j^*(x) g_j^*(\xi)$, kde $h_i^*(x), g_i^*(z), i = 1, \ldots, s$ jsou reálné, omezené a spojité funkce na $\chi \times \Xi$. Pokud dále platí alespoň jedna z následujících podmínek:

(i) $\{\xi_i\}$ je posloupnost nezávislých vektorů

(ii) $\{\xi_i\}$ je silně stacionární posloupnost splňující ϕ-mixing podmínku pro kterou platí

$$\lim \sup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N-1} (N-k)\phi(k) < \infty,$$

potom pro $N \to \infty, \epsilon > 0, \nu \in (0{,}1/2)$ platí:

$$P[N^\nu |H(x_N^*) - H(x^*)| > \epsilon] \to 0.$$

Tuto vlastnost lze ukázat i pro strong mixing posloupnost. Nejdříve si ukážeme, jak této vlastnosti docílit u věty 18, ze které už plyne zmíněná vlastnost pro optimalizační úlohu.

Věta 22. Za předpokladů věty 18 pro strong mixing posloupnost platí pro $\epsilon > 0$

$$P\left(\frac{N^\nu}{N}\left|\sum_{t=1}^{N} \xi_t\right| \geq \epsilon\right) \to 0$$

pro $N \to \infty$ a $\nu \in (0{,}1/2)$.

Důkaz. Použijeme vztah z věty 18:

$$P\left(\left|\sum_{t=1}^{N} \xi_t\right| \geq y\sqrt{N}\sigma \log(\sigma^{-I})\right) \leq a\mathrm{e}^{-b\sqrt{y}},$$

který upravíme na

$$P\left(\frac{N^\nu}{N}\left|\sum_{t=1}^{N} \xi_t\right| \geq \frac{N^\nu}{N} y\sqrt{N}\sigma \log(\sigma^{-I})\right) \leq a\mathrm{e}^{-b\sqrt{y}}.$$

Pokud položíme $y = \frac{N^{1/2}}{N^\nu}\frac{\epsilon}{\sigma\log(\sigma^{-I})}$, můžeme předchozí výraz přepsat na

$$P\left(\frac{N^\nu}{N}\left|\sum_{t=1}^{N}\xi_t\right| \geq \epsilon\right) \leq a\mathrm{e}^{-cN^{\frac{1}{2}\left(\frac{1}{2}-\nu\right)}},$$

kde jsme veškeré proměnné v y kromě N shrnuli pod konstantu c. Nyní si stačí uvědomit, že z vlastnosti exponenciální funkce je konvergence výrazu na pravé straně nerovnosti zajištěna pro $\nu < 1/2$. □

Poznámka 2. Analogické tvrzení vzhledem k formulaci věty 18 zjevně platí i pro ϕ-mixing posloupnost.

Nyní stačí formulovat tuto vlastnost pro optimalizační úlohu ve tvaru jako je uvedena ve větě 19.

Důsledek 2. Předpokládejme, že platí předpoklady věty 19 a $h(x,\xi) = \sum_{j=1}^{s} h_j^*(x)g_j^*(\xi)$. Potom platí, že

$$P[N^\nu|H(x_N^*) - H(x^*)| \geq \epsilon] \to 0$$

pro $N \to \infty$ a $\nu \in (0,1/2)$.

Nevýhodou uvedených vět je požadavek na tvar účelové funkce, který není splněn např. ani u velmi jednoduchých úloh, jako je ta z příkladu 2. Uveďme si proto větu, která za splnění určitých podmínek zajišťuje konvergenci i pro obecnější tvar účelové funkce a náhodné vektory.

Věta 23 (Kaňková (1993), T. 3.2, str. 106)**.** Nechť

(1) χ je kompaktní množina

(2) $g(x,\xi)$ je spojitá a omezená funkce na $\chi(d) \times \Xi$ pro nějaké $d > 0$ a ξ je náhodný vektor

(3) $g(\cdot,\xi)$ je pro každé $\xi \in \Xi$ Lipschitzovská funkce $x \in X[d]$ s Lipschitzovskou konstantou L nezávislou na $\xi \in \Xi$,

(4) Pokud platí alespoň jedna z následujících podmínek:

   (a) $\{\xi_i, i \in \mathbb{Z}\}$ je posloupnost nezávislých vektorů a $0 < v < \frac{1}{2}$,

   (b) $\{\xi_i, i \in \mathbb{Z}\}$ je m-závislá posloupnost, $m \geq 1$ a $0 < v < \frac{1}{2}$,

   (c) $\{\xi_i, i \in \mathbb{Z}\}$ je silně stacionární posloupnost splňující ϕ-mixing podmínku, pro kterou platí

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N-1}(N-k)\phi(k) < \infty \text{ a } 0 < v < \frac{1}{n+2},$$

   kde n je rozměr prostoru $\mathbb{R}^n$.

Potom pro $N \to \infty$ platí

$$P[N^{\nu}|H(x_N^*) - H(x^*)| > \epsilon] \to 0.$$

Uvedeme si nyní příklad, ve kterém názorně ukážeme, jakým způsobem se dají vypočítat koeficienty mixing posloupností (konkrétně strong mixing) a také jejich použití ve větě 19.

2.4.4 Příklady

Příklad 10. Uvažujme Markovův řetězec $\{X_t, t = 0, 1, \dots\}$ se třemi stavy, ve kterých nabývá hodnot $-1, 0$ a 1. Náhodná veličina X_0 nabude každého stavu se stejnou pravděpodobností 1/3. Tabulka charakterizující matici pravděpodobnosti přechodu je následující:

$$P = \begin{array}{c||c|c|c} & \text{-1} & 0 & 1 \\ \hline\hline \text{-1} & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & 0 & \frac{1}{2} & \frac{1}{2} \end{array},$$

tedy například $P(X_{t+1} = 1|X_t = 1) = 1/2, P(X_{t+1} = 1|X_t = -1) = 0$, atd.

Z kapitoly 2.3.2 víme, že striktně stacionární Markovův řetězec je strong mixing posloupnost s koeficientem

$$\alpha(t) = \alpha(\sigma(X_0), \sigma(X_t)), \text{ kde}$$

$$\alpha(\sigma(X_0), \sigma(X_t)) = \sup_{A \in \sigma(X_0), B \in \sigma(X_t)} |P(A \cap B) - P(A)P(B)|. \tag{2.10}$$

K výpočtu použijeme vztahu (2.10). Jev A (resp. B) znamená, že se Markovův řetězec dostane do určitého stavu v určitém čase. Lze snadno nahlédnout, že tento Markovův řetězec je připraven tak, že $P(A) = P(B) = 1/3$ pro $A \in \sigma(X_t), B \in \sigma(X_k)$, a tedy $P(A)P(B) = 1/9$ pro libovolné $t, k = 0, 1, \dots$.

Ze vztahu (2.10) nás pro výpočet $\alpha(t)$ tedy bude ještě zajímat výraz $P(A \cap B)$. Nejdříve si uvědomme, že tento výraz je pravděpodobnost přechodu z jednoho stavu do druhého po t krocích přenásobená 1/3, tedy pravděpodobností s jakou nabude řetězec výchozího stavu.

Díky tomu výpočet provedeme snadno pomocí mocniny matice přechodu. Využijeme tedy známého pravidla, že z matice P^t získáme pravděpodobnosti přechodu mezi stavy po t krocích.

Přepíšeme-li si matici přechodu jako

$$P = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

lze snadno ukázat, že pro t sudé platí

$$P^t = \left(\frac{1}{2}\right)^t \begin{pmatrix} p_t & q_t & q_t \\ q_t & p_t & q_t \\ q_t & q_t & p_t \end{pmatrix}$$

a pro t liché platí

$$P^t = \left(\frac{1}{2}\right)^t \begin{pmatrix} p_t & p_t & q_t \\ p_t & p_t & p_t \\ q_t & p_t & p_t \end{pmatrix},$$

kde $p_t = 2p_{t-1}$ pro t liché, $p_t = p_{t-1} + q_{t-1}$ pro t sudé, $q_t = p_t - 1$, $p_0 = 1$, $q_0 = 0$. Toto pravidlo tedy platí i pro $t = 0$, kdy P^0 je jednotková matice.

Nyní už tedy máme vše potřebné, abychom mohli napsat, že koeficient $\alpha(t)$ je roven buď $\frac{1}{3}\left|\frac{p_t}{2^t} - \frac{1}{3}\right|$ nebo $\frac{1}{3}\left|\frac{q_t}{2^t} - \frac{1}{3}\right|$. Potřebovali bychom tedy zjistit, který z výrazů je větší. To ale není nutné pro ověření předpokladů věty 18. Spokojíme se tedy s tím, že

$$\frac{1}{3}\left|\frac{p_t}{2^t} - \frac{1}{3}\right| = \frac{1}{9}\frac{1}{2^t}\left|2^t - 3p_t\right|$$

a vzhledem k vlastnostem matice přechodu je výraz v absolutní hodnotě určitě menší nebo roven 2 (pro $t \geq 0$) a můžeme tak napsat

$$\frac{1}{3}\left|\frac{p_t}{2^t} - \frac{1}{3}\right| \leq \frac{2}{9}\frac{1}{2^t}.$$

Pokud místo p_t v předchozí úvaze použijeme q_t, dospějeme ke stejnému výsledku a tak máme splněn předpoklad

$$\alpha(t) \leq uv^t \text{ pro } u = \frac{2}{9}, v = \frac{1}{2},$$

aniž bychom museli nutně vědět, který z výrazů je $\alpha(t)$. Protože $E_\xi[|\xi_t|^2] = 2/3$, snadno ověříme, že platí

$$N \sup_t (E_\xi[|\xi_t|^{2+\gamma}])^{2/(2+\gamma)} \geq I$$

pro $N \geq 2$ a konstantu $\gamma > 0$, zvolíme-li $I \in [1, \frac{4}{3})$ tak, abychom splnili předpoklady věty 18 že $|X_t| \leq I$ a $v \in [0, I)$. Tvrzení věty tedy platí pro výběr o rozsahu 2 a větší.

Jelikož se jedná o řetězec s konečným počtem stavů, ireducibilní a aperiodický, splňuje i ψ-mixing, a proto i ϕ-mixing. Dalo by se tedy ukázat také tvrzení věty 18 věnující se ϕ-mixingu.

Protože nás ale zajímá spíše využití v nějaké optimalizační úloze, vzpomeňme si na příklad 7, kde nás zajímala pravděpodobnost

$$P[|H(x_N^*) - H(x^*)| \geq \epsilon] = P[a_2(x_N^* - a_1\overline{\xi})^2 \geq \epsilon].$$

V příkladu 7 bylo použito centrální limitní věty pro nalezení omezující exponenciální funkce. Nyní ale máme nikoli nezávislou, ale strong mixing posloupnost a využít centrální limitní větu tedy nelze. Díky větě 18 ale můžeme zajistit alespoň existenci této exponenciální funkce.

Připomeňme, že $\overline{\xi} = E_\xi[\xi] = 0$ a $x_N^* = \dfrac{a_1}{N}\sum_{i=1}^{N}\xi_i$, můžeme tedy upravit výraz, se kterým jsme pracovali v příkladu 7 také pro strong mixing posloupnost:

$$\begin{aligned} P[a_2(x_N^* - a_1\overline{\xi})^2 \geq \epsilon] &= P\left[a_2\left(\frac{a_1}{N}\sum_{i=1}^{N}\xi_i\right)^2 \geq \epsilon\right] = \\ &= P\left[\frac{a_2a_1^2}{N^2}\left(\sum_{i=1}^{N}\xi_i\right)^2 \geq \epsilon\right] \leq P\left[\frac{a_2a_1^2}{N}\left(\sum_{i=1}^{N}\xi_i\right)^2 \geq \epsilon\right], \end{aligned}$$

a dále vhodně zvolit ϵ tak, aby $\frac{\sqrt{\epsilon}}{\sqrt{a_2}|a_1|} = y\sigma\log(\sigma^{-I})$. Tím je připraven výraz tak, abychom mohli použít tvrzení věty 18:

$$\begin{aligned} P\left[\frac{a_2a_1^2}{N}\left(\sum_{i=1}^{N}\xi_i\right)^2 \geq \epsilon\right] &= P\left[\left(\sum_{i=1}^{N}\xi_i\right)^2 \geq \frac{\epsilon N}{a_2a_1^2}\right] = \\ &= P\left[\left|\sum_{i=1}^{N}\xi_i\right| \geq y\sqrt{N}\sigma\log(\sigma^{-I})\right] \leq a\mathrm{e}^{-b\sqrt{y}}, \end{aligned}$$

kde $\sigma^2 \cong (2/3)$ podle volby $\gamma > 0$; $I = 1$. Připomeňme, že je $a_2 > 0$ ze zadání příkladu 7 a odmocnění tedy nic nebrání. I přesto, že jsme nenalezli koeficienty exponenciální funkce, máme alespoň zajištěnu její existenci. $\square$

V této práci jsme se z velké části zabývali konvergencí výrazu $P[H(x_N^*) - H(x^*) \geq \epsilon]$, ať už pro nezávislé nebo závislé posloupnosti. Ukázali jsme, že je tento výraz omezen exponenciální funkcí v závislosti na N, ale neporovnávali jsme mezi sebou jednotlivé posloupnosti, s jakou „rychlostí“ pro ně tento výraz konverguje. V numerické simulaci si nyní ukážeme, jak se liší „rychlost“ konvergence pro posloupnosti, jejichž veličiny jsou stejně rozdělené, ale jedna z posloupností je závislá.

Příklad 11. Snažíme se ukázat rychlost konvergence pro výraz

$$P[H(x_N^*) - H(x^*) \geq \epsilon]$$

postupně pro posloupnost nezávislých veličin a pro m-posloupnost. Pro tento účel použijeme příklad prodavače novin (příklad 2). Abychom mohli porovnat „rychlost“ konvergence obou posloupností, musí mít veličiny v obou posloupnostech stejné rozdělení. Jako m-posloupnost zvolíme posloupnost $\{\xi_i^{(2)}\}$ klouzavých průměrů o délce 2, tedy

$$\xi_i^{(2)} = \frac{1}{2}(\eta_{i-1} + \eta_i), i = 1, 2, \ldots, N,$$

kde $\{\eta_i, i = 0, 1, \ldots N\}$ je výběr z rozdělení $R[0,1]$. Abychom vytvořili posloupnost $\{\xi_i^{(1)}, i = 1, 2, \ldots, N\}$, nezávislých veličin se stejným rozdělením jako $\xi_i^{(2)}$, stačí vzít 2 výběry: $\{\eta_i^{(1)}\}, \{\eta_i^{(2)}\}, i = 0, 1, \ldots, N$ z rozdělení $R[0,1]$, které jsou nezávislé i mezi sebou navzájem a definovat veličinu $\xi_i^{(1)}$ takto:

$$\xi_i^{(1)} = \frac{1}{2}(\eta_i^1 + \eta_i^2), i = 1, 2, \ldots, N.$$

Vidíme tedy, že obě posloupnosti se liší opravdu jen závislostí. Nyní nalezneme numericky řešení úlohy

$$\min_{x \in [0,1]} H(x^*) = \min_{x \in [0,1]} x \cdot p - c \cdot E_D[\min(D, x)],$$

kde veličina D má rozdělení vzniklé zprůměrováním dvou na sobě nezávislých rozdělení $R[0,1]$ a parametry v úloze jsou určeny takto: $p = 0.8, c = 1$. Řešení nám vyjde $x^* = 0.316$ a $H(x^*) = -0.042$.

Pro obě zmíněné posloupnosti dále vypočítáme hodnotu výrazu $P[H(x_N^*) - H(x^*) \geq \epsilon]$, kde rozsah výběru máme postupně $N = 10, 40, 80, 120, \ldots, 600$. Dále volíme $\epsilon = 0.0001$.

Postupujeme tak, že pro každý rozsah N vygenerujeme 2500 různých posloupností, z nichž poté vypočítáme hodnotu x_N^* a dosadíme do výrazu $P[H(x_N^*) - H(x^*) \geq \epsilon]$. Zmíněné hodnoty vložíme do grafu a můžeme porovnat (viz tabulka 2.1 a obrázek 2.1). Zdrojový soubor k příkladu je uveden na přiloženém CD, viz příloha 2.6. □

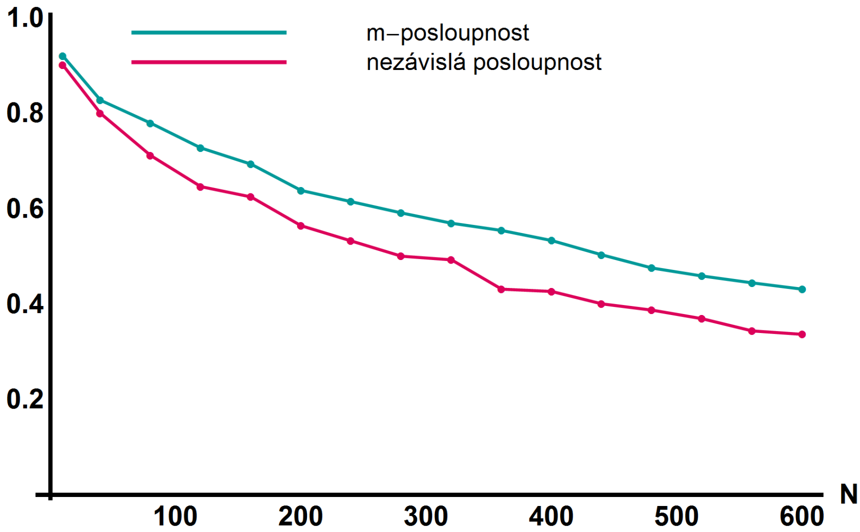
Porovnali jsme si „rychlost“ konvergence nezávislé a závislé posloupnosti. Nyní si ukážeme, jak se liší konvergence u posloupností podobného typu, ale s různou mírou závislosti.

Příklad 12. Představme si farmáře, který pěstuje rýži. Má přístup k říčce s proměnlivým průtokem, ze které může odebírat vodu pro svá pole. Jeho týdenní

N	10	40	80	120	160	200	240	280
m-závislá	0.828	0.780	0.728	0.694	0.639	0.615	0.592	0.570
nezávislá	0.800	0.712	0.647	0.625	0.565	0.533	0.501	0.493
N	360	400	440	480	520	560	600	320
m-závislá	0.921	0.555	0.534	0.504	0.476	0.459	0.445	0.432
nezávislá	0.902	0.432	0.427	0.401	0.388	0.370	0.344	0.337

Tabulka 2.1: Porovnání „rychlosti“ konvergence nezávislé a m-závislé posloupnosti

$P[H(x_N^*) - H(x^*) \geq \epsilon]$

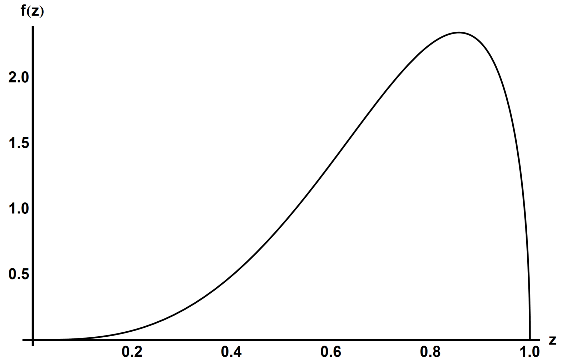


Obrázek 2.1: Porovnání „rychlosti“ konvergence nezávislé a m-závislé posloupnosti

potřeba je 10 jednotek vody. Předpokládejme, že týdenní průtok vody vychází z beta rozdělení s parametry $\alpha = 4$, $\beta = 1.5$, transformovaného na interval [0,10]. Hustotu beta rozdělení je možné si prohlédnout na obrázku 2.2. Jak vidíme, farmář potřebuje zpravidla více vody, než je schopen zajistit z říčky.

Farmář má ale přístup k vodnímu kanálu, ze kterého může po předchozí objednávce odebírat 1 jednotku vody za 1 jednotku peněz. Pokud se rozhodne, že chybějící vodu z kanálu neodebere, vzniká mu ztráta ve výši 2 jednotek peněz za každou chybějící jednotku vody do potřebných 10 kvůli nižší úrodě. Pokud nakoupí více vody, než potřebuje, již z ní žádný užitek nemá. Pokud by farmář znal průtok řeky předem, zjevně by se vyplatilo nakoupit tolik vody, aby měl přesně potřebných 10 jednotek.

Farmář ale průtok předem nezná, zná pouze jeho rozdělení. Snaží se tedy



Obrázek 2.2: Hustota beta rozdělení s parametry $\alpha = 4$, $\beta = 1.5$

docílit přes nějaké dlouhé období co nejnižších nákladů na pěstování rýže a řeší tedy následující úlohu:

$$\min_{x \in [0,10]} x + 2E_\xi \max[0, 10 - x - \xi],$$

kde ξ je veličina značící rozdělení množství vody, které může odebírat, x je množství vody, které odebere. Předpokládejme, že veličině ξ odpovídá posloupnost $\{\xi_i, i = 1, 2, \ldots N\}$ určená následovně:

$$\xi_i = \frac{1}{\sum_{j=0}^{4} a_j} \sum_{j=0}^{4} a_j \eta_{i-j},$$

kde η_i je množství srážek spadlých na horním toku říčky a má právě výše zmíněné beta rozdělení s parametry $\alpha = 4$, $\beta = 1.5$. Jedná se tedy o m-závislou posloupnost, $m = 5$, kde koeficient a_j určuje, jak závisí množství vody v říčce na srážkách před j týdny. Podle toho, jak stanovíme tyto koeficienty, bude závislost na předchozích srážkách buď „slabší" nebo „silnější".

Pro následující numerickou simulaci využijeme různých koeficientů a_j, v tabulce jsou uvedené od „nejsilnější" po „nejslabší" závislost.

a_4	a_3	a_2	a_1	a_0
1	1	1	1	1
0.6	0.7	0.8	0.9	1
0.1	0.3	0.5	0.7	1
0.1	0.1	0.1	0.5	1

Tabulka 2.2: koeficienty a_j pro jednotlivé m-posloupnosti, seřazené od „nejsilnější" po „nejslabší" závislost

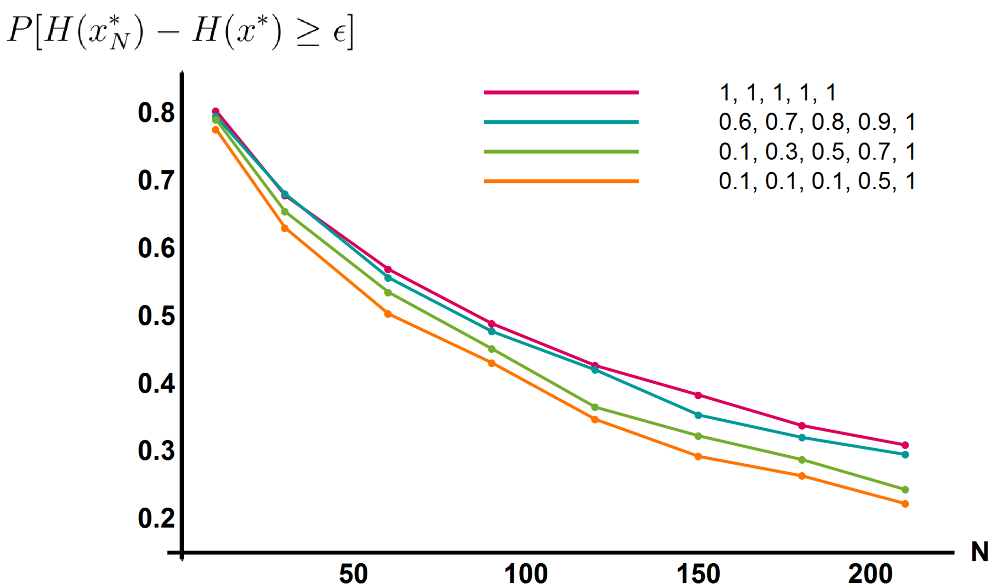
Pro 4 m-posloupnosti určené těmito koeficienty nyní vypočítáme hodnotu výrazu $P[H(x_N^*) - H(x^*) \geq \epsilon]$, kde $\epsilon = 0.01$ tak, že vypočteme celkem 2500 hodnot výrazu $H(x_N^*) - H(x^*)$ pro náhodně generované posloupnosti délky N, z nichž pak určíme výše zmíněnou pravděpodobnost. Vše postupně provádíme pro $N = 10, 30, 60, 90, \ldots, 210$.

Připomeňme, že hodnota $H(x^*)$ se liší pro jednotlivé posloupnosti – rozdělení veličin v jednotlivých posloupnostech se liší vzhledem k různým hodnotám koeficientů a_j. Hodnota $H(x^*)$ činí $3.349, 3.360, 3.461, 3.610$, opět v pořadí od „nejsilnější“ po „nejslabší“ závislost.

Výsledky si lze prohlédnout v tabulce 2.3 nebo přehledně na obrázku 2.3. Pozorujeme, že pro posloupnosti s menší mírou závislosti je „rychlost“ konvergence vyšší. Zdrojový soubor k příkladu je uveden na přiloženém CD, viz příloha 2.7. □

N	10	30	60	90	120	150	180	210
1, 1, 1, 1, 1	0.803	0.678	0.569	0.489	0.427	0.383	0.338	0.309
0.6, 0.7, 0.8, 0.9, 1	0.795	0.681	0.557	0.477	0.420	0.354	0.320	0.295
0.1, 0.3, 0.5, 0.7, 1	0.790	0.654	0.535	0.452	0.365	0.323	0.288	0.243
0.1, 0.1, 0.1, 0.5, 1	0.776	0.630	0.503	0.431	0.347	0.292	0.264	0.222

Tabulka 2.3: koeficienty a_j pro jednotlivé m-posloupnosti, seřazené od „nejsilnější“ po „nejslabší“ závislost



Obrázek 2.3: Porovnání „rychlosti“ konvergence různě závislých m-posloupností

2.5 Centrální limitní věta pro závislé posloupnosti

V kapitole 1.2.2 jsme si ukázali použití centrální limitní věty pro určení intervalu spolehlivosti pro $H(x)$, to ale platí pouze v případě posloupnosti nezávislých veličin. Jelikož jsme se v této práci zabývali posloupnostmi s různými typy závislosti, bylo by tedy vhodné uvést tvrzení vycházející z centrální limitní věty i pro tyto posloupnosti. Nejdříve uveďme tvrzení pro posloupnosti m-závislé. Používáme zde značení: $S_N = \sum_{i=1}^{N} \xi_i$.

Věta 24 (Shang (2012), T. 1.1, str. 713)**.** Nechť $\{\xi_i, i = 1, 2, \dots\}$, je striktně stacionární, m-závislá posloupnost náhodných veličin taková, že $E_\xi[\xi_1] = \mu$ a $0 < \mathrm{Var}_\xi[\xi_1] = \sigma^2 < \infty$. Potom platí

$$\frac{\sqrt{N}}{\tau}\left(\frac{S_N}{N} - \mu\right) \xrightarrow{D} N(0{,}1)$$

pro $N \to \infty$, kde $\tau^2 = \sigma^2 + 2\sum\limits_{i=2}^{m+1} Cov(\xi_1, \xi_i)$.

Dále se věnujme strongly mixing posloupnostem s parametrem $\alpha(n)$.

Věta 25 (Merlevède a Peligrad (2000), T. 1.2, str. 1337)**.** Nechť $\{\xi_i, i = 1, 2, \dots\}$ je striktně stacionární, strongly mixing posloupnost náhodných veličin taková, že $E_\xi[\xi_1] = 0$ a $E_\xi[\xi_1^2] < \infty$. Pokud platí

$$\sum_{t=1}^{\infty} \int_0^{\alpha(t)} Q^2_{|\xi_1|}(0)\mathrm{d}u < \infty,$$

kde $Q_\xi(u) = \inf\{s \geq 0 : P(\xi > s) \leq u\}$ je kvantilová funkce. Potom $\sigma^2 = E_\xi[\xi_1] + 2\sum\limits_{n=1}^{\infty} E_\xi[\xi_1 \xi_n], \sigma^2 \in [0{,}\infty)$. Pokud navíc $\sigma^2 > 0$, potom platí

$$\frac{S_N}{\sqrt{N}\sigma} \xrightarrow{D} N(0{,}1) \quad \text{pro } N \to \infty.$$

Čtenáře, který by měl zájem o tvrzení týkající se i dalších typů mixingu, konkrétně ϕ- a ρ-mixingu, můžeme odkázat na Peligrad a Utev (1997), T. 2.2., str. 445. Nyní nám zbývá ještě uvést tvrzení pro slabě závislé posloupnosti:

Věta 26 (Bardet a kol. (2008), T. 2, str. 161)**.** Nechť $\{\xi_i, i = 1, 2, \dots\}s$ je striktně stacionární posloupnost náhodných veličin s $E_\xi[\xi_1] = 0$ a $E_\xi[\xi_1^{2+\delta}] < \infty$ pro $\delta > 0$. Dále předpokládejme, že $\{\xi_i, i = 1, 2, \dots\}$ je λ- (nebo θ-) slabě závislá posloupnost splňující $\lambda_r = O(r^{-c})$ (nebo $\theta_r = O(r^{-c})$) pro $r \to \infty, c > 4 + 2/\delta$. Potom existuje $0 < \sigma^2 < \infty$ takové, že

$$\frac{S_N}{\sqrt{N}\sigma} \xrightarrow{D} N(0{,}1) \quad \text{pro } N \to \infty.$$

Přesvědčili jsme se tedy, že lze za určitých okolností použít centrální limitní větu i v případě výše zmíněných závislých posloupností. Uvedené věty by šlo zajisté upravit i pro použití v optimalizačních úlohách jako v kapitole 1.2.2, tedy aby pro výraz

$$\sqrt{N}[H_N(x) - H(x)]$$

byla zajištěna konvergence k normálnímu rozdělení $N(0,\sigma^2(x))$ pro nějaké $\sigma^2(x) > 0$. Připomeňme, že ve zmíněných větách není rozptyl normálního rozdělení, ke kterému výraz konverguje, „klasickým" rozptylem, ale jinak odvozeným výrazem. Ukažme si nyní konvergenci na numerickém příkladu, ve kterém použijeme pro porovnání s normálním rozdělením odhad rozptylu na základě vypočtených hodnot.

Příklad 13. V tomto příkladu postupně provedeme to samé pro 4 různé posloupnosti. První z nich je nezávislý výběr z rozdělení $R[-1,1]$, druhou posloupností je m-závislá posloupnost pro $m = 3$, která je tvaru

$$\xi_i = \frac{1}{3}(\eta_i + \eta_{i-1} + \eta_{i-2}),$$

kde η_i má rozdělení $R[-1,1]$. Třetí posloupností je Markovův řetězec, tedy zástupce strong mixing posloupností. Ten je postaven na podobné myšlence jako v příkladu 10, jeho množina stavů je ale následující: $\{i/100, i = -100, 100\}$, tudíž nabývají hodnot na intervalu [-1,1].

Pravděpodobnosti přechodu mezi stavy jsou určeny tak, aby se pro stavy $\{i/100, i = -50, 50\}$ řetězec posunul do stavu $(i+j)/100, j \in \{-50, 49, \ldots, 49, 50\}$ s pravděpodobností $1/100$.

V případě, že se řetězec nachází ve stavu $\{i/100, i = -100, -51\}$, může přejít do stavu $(i + j)/100, j \in \{-k, -k + 1, \ldots, 49, 50\}$, kde k je takové, aby $i + j \geq -100$. Pravděpodobnost přechodu je potom pro všechny stavy rovna $1/(50 + k)$.

Analogicky zavedeme pravděpodobnosti přechodu pro stavy $\{i/100, i = 51, 100\}$. Řetězec nabude výchozího stavu z množiny $\{i/100, i = -100, 100\}$, každý s pravděpodobností $1/200$.

Zjednodušeně můžeme tento řetězec popsat tak, že se vždy může přesunout o maximálně 50 stavů „nahoru" či „dolů" (v „krajních" stavech samozřejmě méně) a v tomto rozmezí se přesune do každého stavu se stejnou pravděpodobností. Připomeňme, že zavedené pravděpodobnosti nevylučují setrvání řetězce ve stejném stavu.

Čtvrtou posloupností je AR(1) posloupnost uvedená v kapitole 2.3.2. Jedná se tedy o slabě závislou posloupnost, která zároveň není mixing. V tomto příkladu použijeme konkrétně posloupnost s parametrem $a = 1/2$, ve které jsou náhodné veličiny z Bernoulliho rozdělení s parametrem $p = 1/2$.

Pro každou z těchto posloupností ukážeme konvergenci výrazu

$$\sqrt{N}[H_N(x^*) - H(x^*)], \tag{2.11}$$

kde $H_N(x)$ je ve tvaru z příkladu 7 s parametry $a_2 = 2, a_1 = 1, a_0 = 0$. Postupně budeme generovat posloupnosti délky $N = 1, 2, 3, 5, 10, 25, 50, 100, 250, 500, 1000$ a spočteme pro ni výraz 2.11. Pro každé N dále spočteme celkem 200 hodnot, které otestujeme pomocí Shapirova–Wilkova testu a uložíme si p-hodnotu. Takovýchto simulací provedeme celkem 1000 a vypočítáme z nich průměrnou p-hodnotu pro dané N. Výsledky uvádíme v následující tabulce s přesností na 3 desetinná místa.

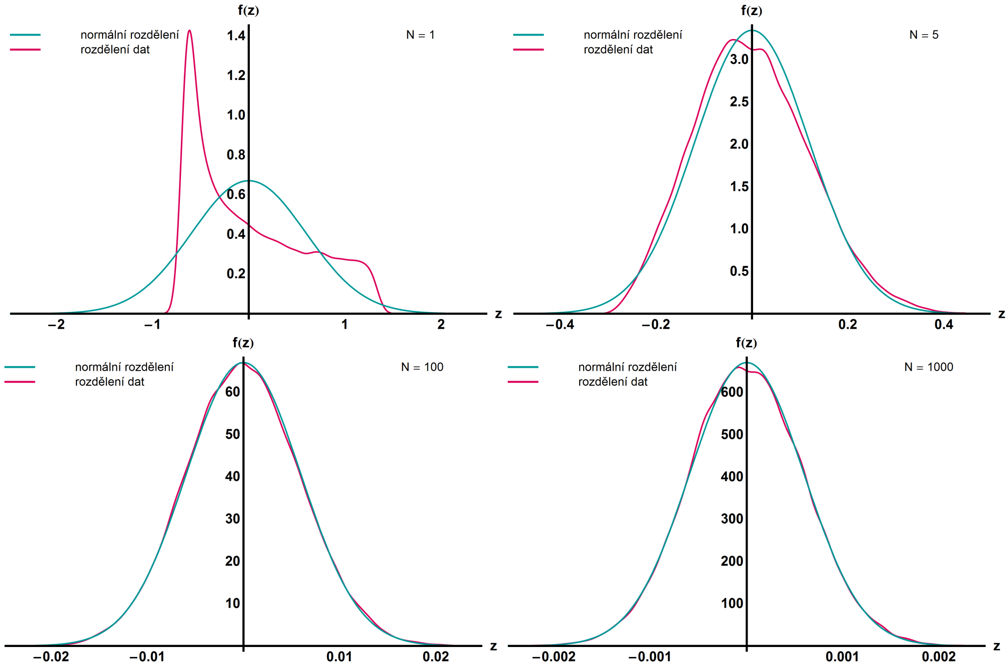
N	5	10	25	50	100	250	500	1000
nezávislá	0.166	0.331	0.453	0.486	0.494	0.487	0.494	0.510
m-závislá	0	0	0.008	0.068	0.202	0.357	0.409	0.453
strong mixing	0	0	0.098	0.298	0.412	0.470	0.478	0.485
slabě závislá	0	0	0.012	0.085	0.192	0.362	0.410	0.441

Tabulka 2.4: Srovnání konvergence k normálnímu rozdělení pro různě závislé posloupnosti

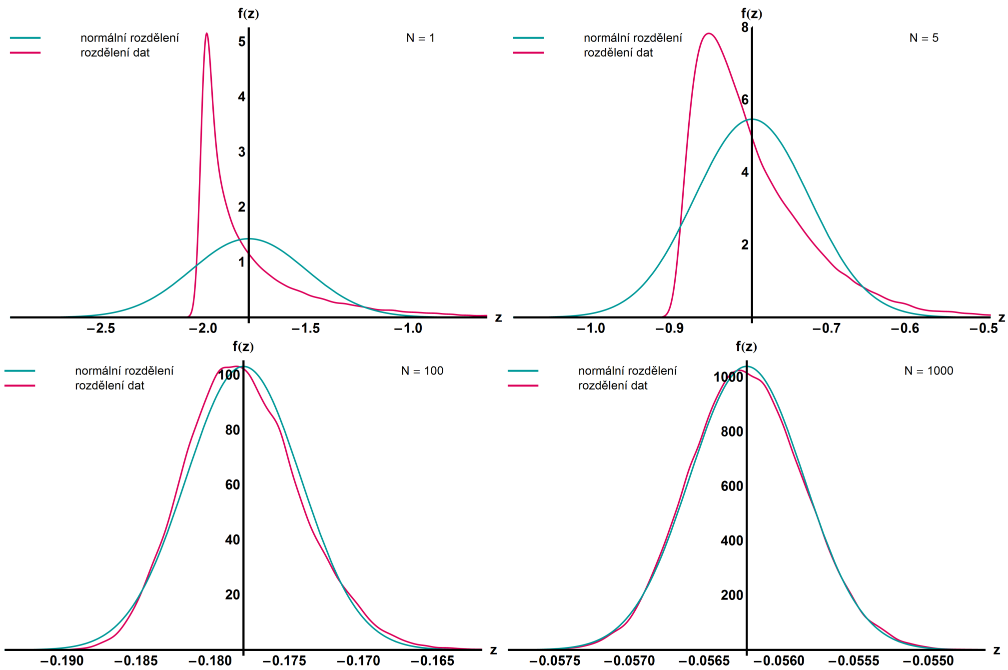
Vidíme, že s rostoucím rozsahem výběru N roste v průměru i p-hodnota, což ukazuje na konvergenci k normálnímu rozdělení. Můžeme předpokládat, že s dalším zvýšením rozsahu N by p-hodnota dále rostla. Připomeňme, že v triviálním případě $N = 1$ nemůžeme mluvit o posloupnosti a výraz 2.11 má rozdělení závisející na výchozím rozdělení generované posloupnosti, tedy např. R[-1,1] v případě nezávislé a m-závislé posloupnosti. Opomeňme též drobnou nepřesnost u m-závislé posloupnosti, kdy ze zjevného důvodu pro $N = 2$ nemůžeme mluvit o m-závislé posloupnosti s $m = 3$.

Uvědomme si, že v tabulce ukazujeme konvergenci k normálnímu rozdělení pro jednotlivé posloupnosti, ale neporovnáváme je mezi sebou. To, že p-hodnota je u některé z posloupností vyšší nevypovídá o tom, že např. m-závislé posloupnosti konvergují k normálnímu rozdělení pomaleji než strong mixing posloupnost, ale spíše o tom, že u konkrétní posloupnosti je míra závislosti vyšší nebo nižší.

Aby bylo snazší si představit, jak konverguje skutečné rozdělení k normálnímu, uvádíme na následujících stránkách obrázky 2.4 – 2.7, na kterých lze rozdělení srovnat pro různé rozsahy výběru a posloupnosti pomocí hustoty normálního rozdělení. V každém obrázku vykreslujeme hustotu veličiny z, která má normální rozdělení s rozptylem a střední hodnotou odpovídající odhadům spočteným z hodnot, ze kterých vykreslujeme hustotu empirického rozdělení. Zdrojový soubor k příkladu je uveden na přiloženém CD, viz příloha 2.8.

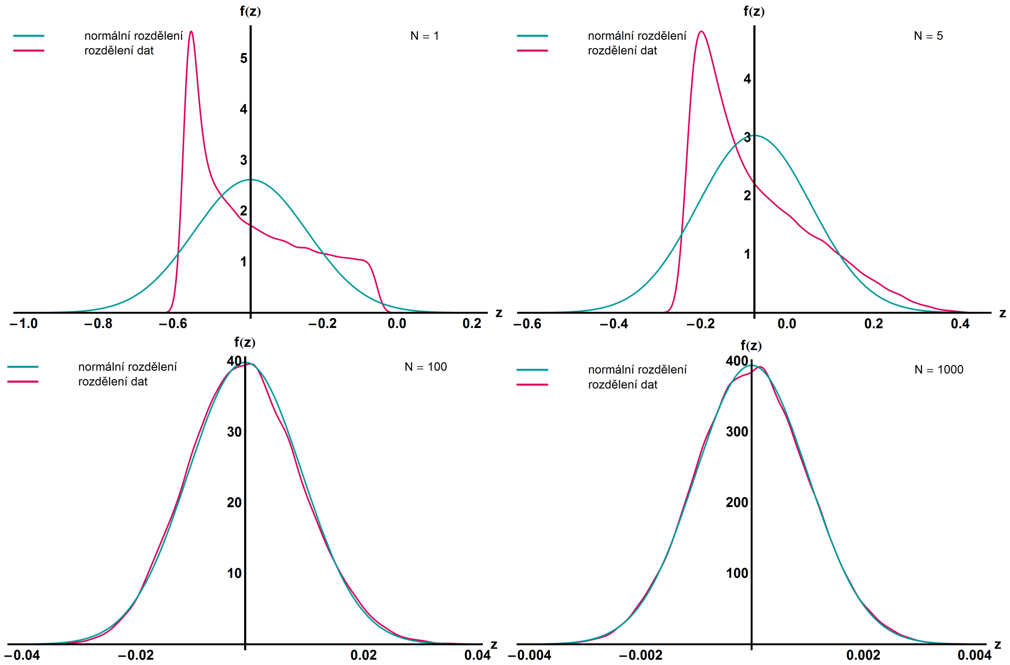


Obrázek 2.4: Porovnání skutečného a normálního rozdělení pro nezávislé posloupnosti

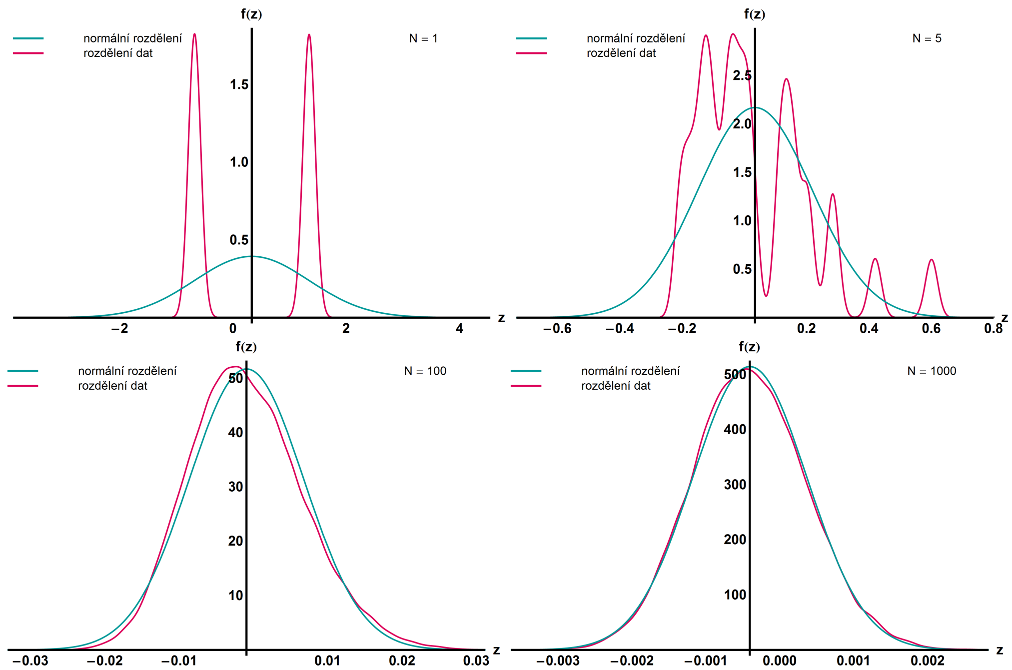


Obrázek 2.5: Porovnání skutečného a normálního rozdělení pro m-závislé posloupnosti

□



Obrázek 2.6: Porovnání skutečného a normálního rozdělení pro strong mixing posloupnosti



Obrázek 2.7: Porovnání skutečného a normálního rozdělení pro slabě závislé posloupnosti

Závěr

V práci byly uvedeny vlastnosti optimalizačních úloh pro nezávislé i závislé výběry. Podstatná část byla také věnována různým koeficientům a druhům závislosti. Ukázalo se, že za určitých podmínek lze i pro závislé výběry dokázat tvrzení podobná tvrzením pro nezávislé výběry a tyto výsledky lze aplikovat na běžné úlohy stochastického programování.

Jak bylo demonstrováno na numerických příkladech, konvergence empirických úloh je „pomalejší“ pro závislé posloupnosti než pro posloupnosti nezávislé a konvergence se „zpomaluje“ podle toho, s jakou se zvyšuje míra závislosti.

Jedná se tedy o velmi příhodný výsledek, protože při použití reálných dat se jen velmi těžko vyhneme závislosti, pokud tedy nepoužijeme hodnoty získané v laboratorních podmínkách. Běžná data využívaná v ekonometrii nebo ekonomické analýza ale určitou závislost zpravidla vykazují.

Literatura

ANDĚL, J. (2011). *Základy matematické statistiky.* matfyzpress.

BARDET, J.-M., DOUKHAN, P., LANG, G. a RAGACHE, N. (2008). Dependent lindeberg central limit theorem and some applications. *Probability and Statistics*, **12**, 154–172.

BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. a survey and some open questions. *Probability Surveys*, **2**, 107–144.

DAI, L., CHEN, C. H. a BIRGE, J. R. (2000). Convergence properties of two-stage stochastic programming. *Journal of Optimization Theory and Applications*, **106**, 489–509.

DOUKHAN, P. (1994). *Mixing: Properties and Examples.* Springer – Verlag.

DOUKHAN, P. (2008). The notion of ψ-weak dependence and its applications to bootstrapping time series. *Probability Surveys*, **17**, 146–148.

DOUKHAN, P. a LOUHICHI, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, **84**, 313–342.

KAŇKOVÁ, V. (1993). A note on estimates in stochastic programming. *Journal of Computational and Applied Mathematics*, **53**, 97–112.

KALL, P. a WALLACE, S. W. (1994). *Stochastic Programming.* John Wiley & Sons.

LINDNER, A. M. (2009). Stationarity, mixing, distributional properties and moments of garch(p, q)–processes. *Handbook of Financial Time Series*, s. 43–69.

MANDL, P. a MAZUROVÁ, L. (1999). *Matematické základy neživotního pojištění.* matfyzpress.

MERLEVÈDE, F. a PELIGRAD, M. (2000). The functional central limit theorem under the strong mixing condition. *The Annals of Probability*, **28**, 1336–1352.

OVCHINNIKOV, A. a RAZ, G. (2011). A newsvendor model with pricing for public interest goods. *SSRN: http://ssrn.com/abstract=1763803*.

PELIGRAD, M. a UTEV, S. (1997). Central limit theorem for linear processes. *The Annals of Probability*, **25**, 443–456.

SEARCÓID, M. . (2006). *Metric spaces*. Springer.

SHANG, Y. (2012). A central limit theorem for randomly indexed m-dependent random variables. *Filomat*, **26**, 713–717.

SHAPIRO, A., DENTCHEVA, D. a RUSZCZYŃSKI, A. (2009). *Lectures on Stochastic Programming. Modeling and Theory*. Society for Industrial and Applied Mathematics.

WOLFRAM RESEARCH, INC. (2012). *Mathematica, Version 9.0*. Wolfram Research, Inc., Champaign, Illinois.

YOSHIHARA, K. (1992). *Weakly dependent stochastic sequences and their applications*. Sanseido.

ZELENÝ, M. (2009). 13. stejnoměrná konvergence. *http://www.karlin.mff.cuni.cz/~zeleny/mff/MA2A/MA2a_Kap_13_tisk.pdf*.

Seznam obrázků

Seznam tabulek

Přílohy

Na přiloženém CD jsou následující soubory, v nichž je uveden zdrojový kód v programu Mathematica (Wolfram Research, Inc. (2012)) pro řešení numerických příkladů. Pro exportování obrázků je potřeba nejdříve spustit soubor knihovna.nb, který se nachází na CD spolu s ostatními soubory.

2.6 srovnaniZavisleNezavisle.nb

Příloha k příkladu 11.

2.7 srovnaniMzavislosti.nb

Příloha k příkladu 12.

2.8 normalita.nb

Příloha k příkladu 13.