Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Vladimír Vachálek

Čebyševova nerovnost a její varianty

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	doc. RNDr. Daniel Hlubinka, Ph.D.
Studijní program:	Matematika
Studijní obor:	Finanční matematika

Praha 2023

Na tomto mieste by som sa chcel poďakovať pánovi doc. RNDr. Danielovi Hlubinkovi, Ph.D. ako za užitočné rady, tak aj za priateľský a ochotný prístup počas vypracovania tejto práce. Nakoniec by som rád poďakoval svojim rodičom za podporu počas štúdia.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) 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 .... dne ....   Podpis autora

Název práce: Čebyševova nerovnost a její varianty

Autor: Vladimír Vachálek

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Daniel Hlubinka, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V předložené práci se zabýváme variantami Čebyševovy nerovnosti pro omezené náhodné veličiny. V první kapitole uvedeme a dokážeme Hoeffdingovu, Bennettovu a Bernsteinovu nerovnost a vysvětlíme některé souvislosti. V druhé kapitole ukážeme, jak dobré odhady jednotlivé nerovnosti nabízejí ve srovnání jak se skutečnou pravděpodobností, tak i s odhady sestrojenými pomocí centrální limitní věty na čtyřech příkladech rozdělení s grafickým zpracováním výsledků.

Klíčová slova: Čebyševova nerovnost, Hoeffdingova nerovnost, Bennettova nerovnost, Bernsteinova nerovnost

Title: Chebyschev type inequalities

Author: Vladimír Vachálek

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Daniel Hlubinka, Ph.D., Department of Probability and Mathematical Statistics

Abstract: In the presented thesis we deal with Chebyshev type inequalities for bounded random variables. In the first chapter we introduce and prove Hoeffding, Bennett and Bernstein inequalities and explain some relationships. In the second chapter we show how tight are the estimates given by each inequality compared to true probability and to the estimate given by central limit theorem on four distributions with graphical processing of results.

Keywords: Chebyshev inequality, Hoeffding inequality, Bennett inequality, Bernstein inequality

Názov práce: Čebyševova nerovnosť a jej varianty

Autor: Vladimír Vachálek

Katedra: Katedra pravdepodobnosti a matematickej štatistiky

Vedúci bakalárskej práce: doc. RNDr. Daniel Hlubinka, Ph.D., Katedra pravdepodobnosti a matematickej štatistiky

Abstrakt: V predloženej práci sa zaoberáme variantami Čebyševovej nerovnosti pre ohraničené náhodné veličiny. V prvej kapitole uvedieme a dokážeme Hoeffdingovu, Bennettovu a Bernsteinovu nerovnosť a vysvetlíme niektoré súvislosti. V druhej kapitole ukážeme, aké dobré odhady jednotlivé nerovnosti ponúkajú v porovnaní so skutočnou pravdepodobnosťou, ako aj s odhadmi zostrojenými pomocou centrálnej limitnej vety na štyroch príkladoch rozdelení s grafickým spracovaním výsledkov.

Kľúčové slová: Čebyševova nerovnosť, Hoeffdingova nerovnosť, Bennettova nerovnosť, Bernsteinova nerovnosť

Obsah

Úvod

V živote, tak ako v matematike sa často stretávame s rôznymi druhmi odhadov. Niekedy sa jednoducho k presnej odpovedi nevieme dostať, inokedy nás presná odpoveď ani nezaujíma. Špeciálnym prípadom odhadov sú matematické nerovnosti, do ktorých patria aj pravdepodobnostné nerovnosti.

Nech $(\Omega, \mathcal{A}, \mathsf{P})$ je pravdepodobnostný priestor a X je reálna náhodná veličina s konečnou strednou hodnotou $\mathsf{E}\ X$ a nenulovým rozptylom σ^2. Označme F distribučnú funkciu náhodnej veličiny X. Pre $a > 0$ dostávame

$$\mathsf{E}\ |X|^r = \int_{-\infty}^{\infty} |x|^r\, dF(x) \geq \int_{|x|\geq a} |x|^r\, dF(x) \geq a^r \int_{|x|\geq a} dF(x) = a^r\, \mathsf{P}(|X| \geq a).$$

Podelením pravej strany výrazom a^r dostávame nerovnosť v nasledujúcom tvare

$$\mathsf{P}(|X| \geq a) \leq \frac{\mathsf{E}\ |X|^r}{a^r},$$

ktorá je známa ako zobecnená Markovova nerovnosť. Voľbou $r = 1$ dostávame

$$\mathsf{P}(|X| \geq a) \leq \frac{\mathsf{E}\ |X|}{a}.$$

Tento špeciálny prípad je známy ako klasická Markovova nerovnosť. Ak zvolíme $r = 2$, namiesto X uvažujeme $X - \mathsf{E}\ X$ a pre $k > 0$ uvažujeme $a = k\sigma$, dostávame nerovnosť

$$\mathsf{P}(|X - \mathsf{E}\ X| \geq k\sigma) \leq \frac{1}{k^2},$$

známu ako klasická Čebyševova nerovnosť, ktorá prakticky uzatvára základný balík pravdepodobnostných nerovností, ktorými sa bežné kurzy pravdepodobnosti zaoberajú. Ich popularita je spôsobená najmä ich všeobecnosťou - na náhodnú veličinu X kladú iba minimálne nároky. Princíp niečoho za niečo sa uplatňuje aj tu, kde kvôli univerzálnosti trpí sila. Skutočne, odhady získané ako z Markovovej, tak aj z Čebyševovej nerovnosti často silno prestrelujú pravdepodobnosť na ľavej strane. Ak však budeme na náhodnú veličinu X klásť väčšie nároky, bude sa aj daný odhad pravdepodobosti zlepšovať.

Kapitola 1

Teoretická časť

1.1 Chernoffove medze

Chernoffove medze pre náhodnú veličinu X sa získajú aplikovaním Markovovej nerovnosti na e^{tX} pre vhodne zvolené t. Pre $t > 0$ odvodíme nasledujúce nerovnosti:

$$\mathsf{P}(X \geq a) = \mathsf{P}(e^{tX} \geq e^{ta}) \leq \frac{\mathsf{E}(e^{tX})}{e^{ta}}.$$

alebo

$$\mathsf{P}(X \geq a) \leq \min_{t>0} \frac{\mathsf{E}(e^{tX})}{e^{ta}}.$$

Podobne pre $t < 0$,

$$\mathsf{P}(X \leq a) = \mathsf{P}(e^{tX} \geq e^{ta}) \leq \frac{\mathsf{E}(e^{tX})}{e^{ta}}.$$

teda

$$\mathsf{P}(X \leq a) \leq \min_{t<0} \frac{\mathsf{E}(e^{tX})}{e^{ta}}.$$

Pretože tieto nerovnosti dávajú exponenciálne klesajúce medze na chvosty rozdelení, sú silnejšie ako medze zakladajúce sa na prvom či druhom momente. Najvhodnejšia hodnota t je zrejme taká, ktorá minimalizuje $\mathsf{E}(e^{tX})/e^{ta}$ avšak často volíme takú hodnotu t, ktorá dáva nami požadovaný tvar. Medze odvodené týmto spôsobom nesú spoločný názov Chernoffove medze.

1.2 Hoeffdingova nerovnosť

V tejto podkapitole formulujeme a dokážeme Hoeffdingovu nerovnosť. Budeme k tomu potrebovať nasledujúce lemma.

Lemma 1.1. *[2] Nech X je náhodná veličina taká, že $a \leq X \leq b$ skoro iste, t.j. s pravdepodobnosťou 1. Potom pre každé s reálne platí*

$$\mathsf{E}\left(e^{s(X-\mathsf{E}\,X)}\right) \leq \exp\left(\frac{s^2(b-a)^2}{8}\right). \tag{1.1}$$

Dôkaz. Podľa [2]. Bez újmy na obecnosti nahradením X výrazom $X - \mathsf{E}\ X$ môžme predpokladať, že $\mathsf{E}\ X = 0$ tak, aby $a \leq 0 \leq b$. Využitím konvexity e^{sx} ako funkcie premennej x dostaneme, že pre každé $a \leq x \leq b$ platí

$$e^{sx} \leq \left(\frac{b-x}{b-a}e^{sa}\right) + \left(\frac{x-a}{b-a}e^{sb}\right).$$

Ďalej aplikovaním strednej hodnoty dostaneme

$$\begin{aligned}
\mathsf{E}\left(e^{sX}\right) &\leq \left(\frac{b-\mathsf{E}\left(X\right)}{b-a}e^{sa}\right) + \left(\frac{\mathsf{E}\left(X\right)-a}{b-a}e^{sb}\right) \\
&= \left(\frac{b}{b-a}e^{sa}\right) + \left(\frac{-a}{b-a}e^{sb}\right) \\
&= e^{T(s(b-a))}.
\end{aligned}$$

kde prvá nerovnosť plynie z toho, že aplikovaním strednej hodnoty na obe strany nerovnice zachováme nerovnosť a druhú rovnosť sme získali vďaka predpokladu $\mathsf{E}\ X = 0$. Tretiu rovnosť dostaneme ako zjednodušenie výrazu v druhom riadku správnou voľbou funkcie T:

$$T(h) = \left(\frac{ha}{b-a}\right) + \log\left(1 + \left(\frac{a - e^h a}{b-a}\right)\right)$$

Derivovaním funkcie T podľa premennej h dostávame

$$T(0) = T'(0) = 0 \text{ a } T''(h) \leq \frac{1}{4} \text{ pre každé h.}$$

Podľa Taylorovej vety je pre nejaké $0 \leq \lambda \leq 1$

$$T(h) = T(0) + hT'(0) + \frac{1}{2}h^2T''(h\lambda) \leq \frac{1}{8}h^2$$

a teda

$$\mathsf{E}\left(e^{sX}\right) \leq e^{\frac{1}{8}s^2(b-a)^2}.$$

❑

Veta 1.2 (Hoeffdingova nerovnosť)**.** *[2] Nech X_1,...,X_n je postupnosť nezávislých náhodných veličín takých, že $a_i \leq X_i \leq b_i$ skoro iste pre každé $i \leq n$. Označme*

$$S_n = \textstyle\sum_{i=1}^n X_i.$$

Potom pre každé $t > 0$ platí:

$$\mathsf{P}\left((S_n - \mathsf{E}\, S_n) \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

Dôkaz. Podľa [2]. Použijeme lemma (1.1), nezávislosť X_i a Markovovu nerovnosť. Pre $t,s > 0$ platí

$$\begin{aligned}
\mathsf{P}\left((S_n - \mathsf{E}\, S_n) \geq t\right) &= \mathsf{P}\left(\exp\big(s(S_n - \mathsf{E}\, S_n)\big) \geq \exp(st)\right) \\
&\leq \exp(-st)\, \mathsf{E}\left(\exp\big(s(S_n - \mathsf{E}\, S_n)\big)\right) \\
&= \exp(-st) \prod_{i=1}^{n} \mathsf{E}\left(\exp\big(s(X_i - \mathsf{E}\, X_i)\big)\right) \\
&\leq \exp(-st) \prod_{i=1}^{n} \exp\left(\frac{s^2(b_i - a_i)^2}{8}\right) \\
&= \exp\left(-st + \frac{s^2 \sum_{i=i}^{n}(b_i - a_i)^2}{8}\right).
\end{aligned}$$

Minimalizáciou výrazu vo vnútri exponenciály podľa premennej s dostávame

$$s = \frac{4t}{\sum_{i=1}^{n}(b_i - a_i)^2}$$

Aplikovaním tejto hodnoty s dostávame

$$\mathsf{P}\left((S_n - \mathsf{E}\, S_n) \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right).$$

❑

Jednoducho povedané, Hoeffdingova nerovnosť nám dáva hornú medzu pravdepodobnosti, že súčet ohraničených náhodných veličín sa nevzďaluje od svojej strednej hodnoty o viac ako nejaké predom dané množstvo. Využíva však pomerne hrubého faktu, a to

$$Var(X) \leq \frac{(b-a)^2}{4} \quad \text{kedykoľvek} \quad a \leq X \leq b$$

To však vie byť pesimistické ak má náhodná veličina X iba malú pravdepodobnosť toho, že sa jej hodnoty budú nachádzať na okrajoch intervalu $[a,b]$. Vylepšenie ponúka Bennettova nerovnosť, avšak predtým, ako ju zavedieme, priblížime jednu z aplikácií Chernoffových nerovností, na ktorú sa v ďalšej časti budeme odkazovať.

Majme náhodnú veličinu X. Potom z Markovovej nerovnosti pre $s > 0$ platí:

$$\mathsf{P}\left(X > t\right) \leq \frac{\mathsf{E}\,(e^{sX})}{e^{st}}.$$

Definujme funkciu ϕ_X predpisom

$$\phi_X(s) = \log\left(\mathsf{E}\left(e^{(sX)}\right)\right)$$

a po dosadení získame

$$\mathsf{P}\left(X > t\right) \leq e^{-st+\phi_X(s)}$$

ďalej ak položíme

$$\phi_X^*(t) = \sup_{t\geq 0}(st - \phi_X(t))$$

máme výsledok

$$\mathsf{P}\left(X > t\right) \leq e^{-\phi_X^*(t)} \tag{1.2}$$

1.3 Bennettova nerovnosť

Veta 1.3 (Bennettova nerovnosť). *[1][3] Nech X_1,...,X_n je postupnosť nezávislých náhodných veličín s konečným rozptylom. Nech $X_i \leq b$ pre každé $i \leq n$. Označme*

$$S_n = \sum_{i=1}^n X_i$$

a

$$\sigma = \sum_{i=1}^n \mathsf{E}\, X_i^2.$$

Pre $u \in \mathbb{R}$ položme $\psi(u) = e^u - u - 1$ a pre $u \geq -1$ položme $H(u) = (1+u)\log(1+u) - u$. Potom pre ľubovoľné $t > 0$ platí:

$$\mathsf{P}\left((S_n - \mathsf{E}\, S_n) \geq t\right) \leq \exp\left(-\frac{\sigma}{b^2} H\left(\frac{bt}{\sigma}\right)\right). \tag{1.3}$$

Dôkaz. Bez újmy na obecnosti predpokladajme $b = 1$. Ďalej si uvedomme, že funkcia $\dfrac{\psi(u)}{u^2}$ je rastúcou funkciou premennej u a pretože $X_i \leq 1$, pre $s > 0$ dostávame

$$\psi(sX_i) = e^{sX_i} - sX_i - 1 \leq \frac{s^2 X_i^2 \psi(s)}{s^2} = X_i^2(e^s - s - 1)$$

z toho

$$e^{(sX_i)} \leq sX_i + 1 + X_i^2(e^s - s - 1)$$

Spočítaním funkcie $\phi_{(S_n - \mathsf{E}\, S_n)}(s)$ zavedenej vyššie a použitím nerovnosti na predošlom riadku máme

$$\begin{aligned}\phi_{(S_n - \mathsf{E}\ S_n)}(s) &= \sum_{i=1}^{n} \log\left(\mathsf{E}\ \left[e^{s(X_i - \mathsf{E}\ X_i)}\right]\right) \\ &= \sum_{i=1}^{n} \left(\log\left(\mathsf{E}\ \left[e^{(sX_i)}\right]\right) - s\,\mathsf{E}\ (X_i)\right) \\ &\leq \sum_{i=1}^{n} \left(\log\left(1 + s\,\mathsf{E}\ (X_i) + \mathsf{E}\ (X_i^2)(e^s - s - 1)\right) - s\,\mathsf{E}\ (X_i)\right)\end{aligned}$$

Z konkavity funkcie $\log(1+u)$ ďalej dostávame

$$\begin{aligned}\phi_{(S_n - \mathsf{E}\ S_n)}(s) &\leq n\left(\log\left(1 + \frac{\sum_{i=i}^{n} \mathsf{E}\ (X_i)}{n} + \frac{\sigma}{n}(e^s - s - 1)\right) - s\frac{\sum_{i=i}^{n} \mathsf{E}\ (X_i)}{n}\right) \\ &\leq \sigma(e^s - s - 1)\end{aligned}$$

Na tomto mieste už vidíme dôvod zavedenia 1.2 pretože

$$\begin{aligned}\mathsf{P}\left((S_n - \mathsf{E}\ S_n) \geq t\right) &\leq e^{-st + \phi_{(S_n - \mathsf{E}\ S_n)}(s)} \\ &\leq e^{-st + \sigma(e^s - s - 1)}\end{aligned}$$

Výraz v poslednom riadku je minimálny pre voľbu $s = \log\left(1 + \frac{t}{\sigma}\right)$. Dosadením dostaneme

$$\begin{aligned}\mathsf{P}\left((S_n - \mathsf{E}\ S_n) \geq t\right) &\leq e^{-\log\left(1+\frac{t}{\sigma}\right)t + \sigma\left(\frac{t}{\sigma} - \log\left(1+\frac{t}{\sigma}\right)\right)} \\ &\leq e^{-\sigma\left[\left(1+\frac{t}{\sigma}\right)\log\left(1+\frac{t}{\sigma}\right) - \frac{t}{\sigma}\right]}\end{aligned}$$

Čím je nerovnosť dokázaná.

❏

Ako sme už spomínali vyššie, veta (1.3) je v niektorých prípadoch vylepšením vety (1.1), no kladie navyše podmienku konečnosti rozptylu danej náhodnej veličiny. Naviac vyjadrenie pravej strany obsahuje pomerne neintuitívnu funkciu H, ktorej voľba je ozrejmená až po dôkladnejšom prezretí dôkazu.

1.4 Bernsteinova nerovnosť

Veta 1.4 (Bernsteinova nerovnosť)**.** *Za platnosti predpokladov vety* (1.3) *platí*

$$P((S_n - E\,S_n) \geq t) \leq \exp\left(-\frac{t^2}{2\left(\sigma + \frac{bt}{3}\right)}\right). \tag{1.4}$$

Dôkaz. Uvažujme funkciu

$$G(u) = H(u) - \frac{u^2}{2\left(1 + \dfrac{u}{3}\right)} \qquad \text{pre } u > 0,$$

kde $H(u) = (1+u)\log(1+u) - u$ z nerovnosti 1.3. Použitím niektorých znalostí matematickej analýzy ukážeme, že

$$H(u) \geq \frac{u^2}{2\left(1 + \dfrac{u}{3}\right)} \qquad \text{pre } u > 0$$

Funkcia G je spojitá na svojom definičnom obore. Predtým, ako začneme počítať derivácie, upravíme funkciu G do nasledujúceho tvaru

$$G(u) = (1+u)\log(1+u) - \frac{5}{2}u + \frac{9}{2}\frac{u}{u+3}.$$

Potom

$$G'(u) = \log(1+u) - \frac{3}{2} + \frac{27}{2(u+3)^2}$$

$$G''(u) = \frac{u^2(u+9)}{(u+1)(u+3)^3} \geq 0$$

Pretože G je rastúca funkcia premennej $u > 0$ a $\lim_{u->0+} G'(u) = 0$, je $G' > 0$ a teda G je rastúca. Zároveň je tiež $\lim_{u->0+} G(u) = 0$, takže $G > 0$ pre $u > 0$.

Výsledný vzťah 1.4 potom dostaneme dosadením $H(u) = \frac{u^2}{2\left(1+\frac{u}{3}\right)}$ do vzťahu 1.3.

❑

Bernsteinove nerovnosti tvoria celý súbor podobných tvrdení, my sme si vybrali jeden konkrétny tvar. Poznamenajme, že nerovnosti vznikali opačným smerom, teda (1.3) vznikla postupnosťou vylepšení vety (1.4), ktoré možno dohľadať v [1]. Je zrejmé, že odhad získaný pomocou vety (1.3) je lepší ako odhad z (1.4). Prečo by sme potom takýto odhad použili? Prirodzenou odpoveďou je, že tvar pravej strany (1.4) je jednoduchší a v niektorých prípadoch nám môže stačiť. Zaujímavou odpoveďou je, že podobný vzťah ako ten z vety (1.4) platí aj za slabších predpokladov.

Poznámka. [4] Ak by sme uvažovali postupnosť nezávislých náhodných veličín X_1,...,X_n s označením rovnakým ako v (1.3) takých, že existujú čísla σ,$b > 0$ pre ktoré platí

$$\textstyle\sum_{i=1}^n \mathsf{E}\,(X_i^2) \leq b$$

a pre každé $k \geq 3$ prirodzené

$$\textstyle\sum_{i=1}^{n} \mathsf{E}\left[(X_i)_+^k\right] \leq \frac{k!}{2}\sigma b^{k-2}$$

potom pre ľubovoľné $t > 0$ platí

$$\mathsf{P}\left((S_n - \mathsf{E}\left(S_n\right)) \geq t\right) \leq \exp\left(-\frac{t^2}{2(\sigma + bt)}\right).$$

Kapitola 2

Porovnanie nerovností

V tejto kapitole porovnáme vyššie spomínané nerovnosti ako medzi sebou, tak aj s odhadom zostrojeným použitím centrálnej limitnej vety. Našou úlohou bude nájsť rozdelenia splňujúce predpoklady viet, kde najväčším obmedzením bude ohraničenosť náhodnej veličiny.

2.1 Rovnomerné rozdelenie

Hovoríme, že náhodná veličina X má *Rovnomerné rozdelenie* na intervale (a,b), $-\infty < a < b < \infty$, značíme $X \sim R(a,b)$, ak jej distribučná funkcia F_X splňuje

$$F_X(t) = \begin{cases} 0\ , & t \leq a \\ \dfrac{t-a}{b-a}, & a < t < b \\ 1\ , & t \geq b. \end{cases}$$

Ak $X \sim R\left(-\frac{1}{2},\frac{1}{2}\right)$, potom

$$\mathsf{E}\,(X) = 0, \qquad\qquad Var(X) = \mathsf{E}\,(X^2) = \frac{1}{12}.$$

Nech X_1,...,X_n je postupnosť nezávislých náhodných veličín takých, že $X_i \sim R\left(-\frac{1}{2},\frac{1}{2}\right)$. Počítajme presné hodnoty $\mathsf{P}\,((S_n - \mathsf{E}\ S_n) \geq t)$ pre $t > 0$, kde stále $S_n = \sum_{i=1}^n X_i$.

Pre $n = 1$ prebieha výpočet dosadením:

$$\mathsf{P}\,(X_1 \geq t) = 1 - F_{X_1}(t) = 1 - \left(t + \frac{1}{2}\right) = \frac{1}{2} - t,\ t \in \left[-\frac{1}{2},\frac{1}{2}\right]$$

Podobný postup budeme chcieť použiť pre $S_2 = X_1 + X_2$. Otázkou však ostáva rozdelenie náhodnej veličiny S_2.

Poznámka (Rozdelenie súčtu dvoch spojitých náhodných veličín). Ak X_1,X_2 sú nezávislé a spojité náhodné veličiny s príslušnými hustotami f_{X_1},f_{X_2}, potom náhodná veličina $S_2 = X_1 + X_2$ má hustotu

$$f_{S_2}(t) = \int_{-\infty}^{\infty} f_{X_1}(x)f_{X_2}(t-x)dx.$$

V našom prípade pre $X_1,X_2 \sim R\left(-\frac{1}{2},\frac{1}{2}\right)$ nezávislé a $t \geq 0$ dostávame

$$f_{S_2}(t) = \int_{-\infty}^{\infty} f_{X_1}(x)f_{X_2}(t-x)dx = \int_{t-\frac{1}{2}}^{\frac{1}{2}} 1dx = 1-t, \quad t \in [0,1]$$

S využitím toho, že $\mathsf{P}(S_2 \geq 0) = \frac{1}{2}$ máme

$$F_{S_2}(t) = \int_{-\infty}^{t} f_{S_2}(x)dx = \int_{0}^{t}(1-x)dx = \frac{1}{2} + t - \frac{t^2}{2}, \quad t \in [0,1]\,.$$

Môžme teda počítať pre $n = 2$:

$$\mathsf{P}\left(S_2 \geq t\right) = 1 - F_{S_2}(t) = 1 - \left(\frac{1}{2} + t - \frac{t^2}{2}\right) = \frac{1}{2} - t + \frac{t^2}{2},\ t \in [0,1]$$

Analogicky pre $n = 3$:

$$\begin{aligned}\mathsf{P}\left(S_3 \geq t\right) = 1 - F_{S_3}(t) &= \frac{1}{2} - \frac{3}{4}t + \frac{1}{3}t^3,\ t \in \left[0,\frac{1}{2}\right]\\ &= \frac{1}{6}\left(\frac{3}{2} - t\right)^3,\ t \in \left[\frac{1}{2},\frac{3}{2}\right]\end{aligned}$$

Spôsobov, akými možno porovnať skutočnú pravdepodobnosť s hornými medzami odvodenými z jednotlivých nerovností je niekoľko. Na obrázku 2.1 sa nachádza priamočiare porovnanie skutočných pravdepodobností spočítaných vyššie s hornými medzami jednotlivých pravdepodobnostných nerovností, kde postupnosť nezávislých náhodných veličín $X_1,...,X_n$ má rozdelenie rovnaké ako v celej tejto podkapitole. Všimnime si, že so zväčšujúcim sa t sú už jednotlivé medze blízko skutočnej pravdepodobnosti, najlepšou sa javí M_{benn}.

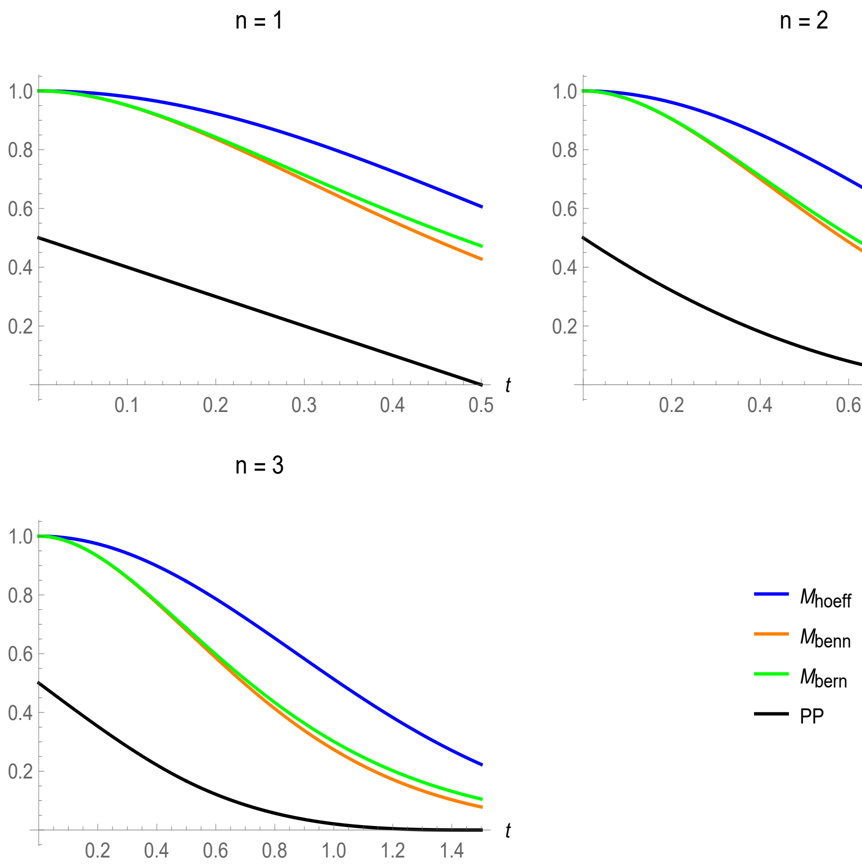
Iným spôsobom by mohlo byť porovnanie rozdielu. Ak by sme uvažovali napríklad $n = 2$, potom z Hoeffdingovej nerovnosti dostaneme odhad

$$\mathsf{P}\left(S_2 \geq t\right) \leq \exp\left(-\frac{2t^2}{2}\right),$$

pravá strana ďalej iba M_{hoeff}, a môžme uvažovať rozdiel

$$M_{hoeff} - \tfrac{1}{2} + t - \tfrac{t^2}{2},\ t \in [0,1]\,.$$

V tomto duchu sa nesie obrázok 2.2 - pre $n \in \{1,2,3\}$ porovnáva rozdiel medzi určenými nerovnosťami a skutočnej pravdepodobnosti javu $S_n \geq t$.



Obr. 2.1: Porovnanie skutočnej pravdepodobnosti označenej PP a odhadov jednotlivých nerovností pre rôzne n

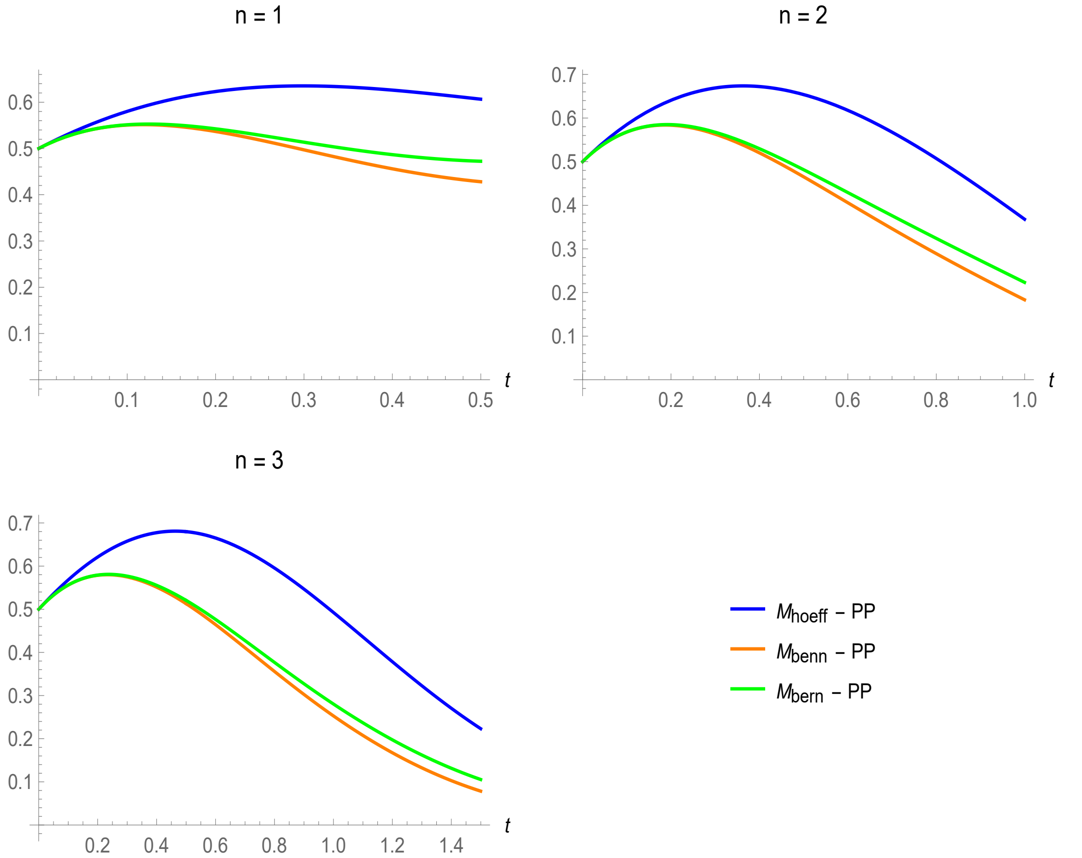
Ak by sme uvažovali n dostatočne veľké, to znamená, aby odhad zostrojený pomocou centrálnej limitnej vety bol platný, dostaneme následujúci odhad

$$\mathsf{P}\left(S_n \geq t\right) = \mathsf{P}\left(\frac{S_n}{\sqrt{\frac{n}{12}}} \geq \frac{t}{\sqrt{\frac{n}{12}}}\right) \approx 1 - \Phi\left(\frac{t}{\sqrt{\frac{n}{12}}}\right) = 1 - \Phi\left(\frac{2\sqrt{3}t}{\sqrt{n}}\right),$$

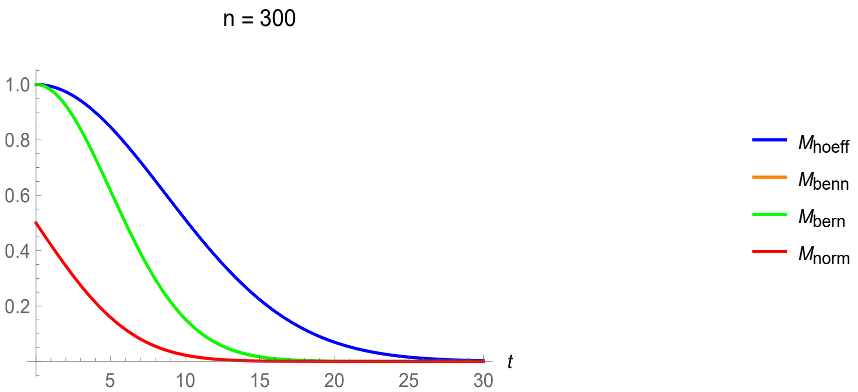
kde Φ označuje distribučnú funkciu normovaného normálneho rozdelenia $N(0,1)$. Vo výpočte sme zároveň využili fakt, že rozptyl náhodnej veličiny S_n je rovný $\frac{n}{12}$. Tento odhad je graficky znázornený a porovnaný na 2.3, kde ukážeme iba jeden graf, pretože pre iné n veľké sú grafy takmer zhodné. V tomto prípade M_{benn} a M_{bern} splývajú.

2.1.1 Diskrétne rovnomerné rozdelenie

Uvažujme postupnosť nezávislých náhodných veličín $X_1,\ldots,X_n$ takých, že X_i má rovnomerné rozdelenie na množine $\{-\frac{1}{2},\frac{1}{2}\}$, teda $\mathsf{P}\left(X_i = -\frac{1}{2}\right) = \frac{1}{2} = \mathsf{P}\left(X_i = \frac{1}{2}\right)$ pre každé $i \in \{1,\ldots,n\}$. V tomto prípade sme teda celú pravdepodobnostnú masu nahrnuli na kraje intervalu $\left(-\frac{1}{2},\frac{1}{2}\right)$.



Obr. 2.2: Porovnanie rozdielu skutočnej pravdepodobnosti a odhadov jednotlivých nerovností pre rôzne n



Obr. 2.3: Porovnanie odhadu zostrojeného pomocou centrálnej limitnej vety a horných odhadov jednotlivých nerovností pre $n = 300$

Náhodná veličina $S_n = \sum_{i=1}^n X_i$ má rozdelenie rovnaké ako $Y - \frac{n}{2}$, kde Y má binomické rozdelenie $Bi(n,\frac{1}{2})$, ktoré je zavedené v 2.2, teda

$$\mathsf{P}\left(S_n = k\right) = \mathsf{P}\left(Y = k + \frac{n}{2}\right) = \binom{n}{k+\frac{n}{2}}\left(\frac{1}{2}\right)^n, \ k \in \left\{-\frac{n}{2}, -\frac{n}{2}+1, \ldots, \frac{n}{2}-1, \frac{n}{2}\right\}$$

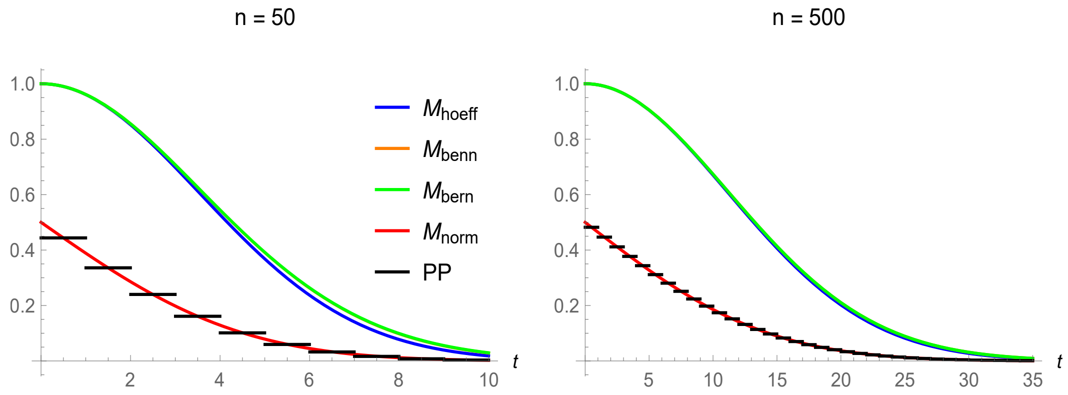
A teda

$$\mathsf{P}\left(S_n \geq t\right) = \sum_{k \geq t} \mathsf{P}\left(S_n = k\right) = \left(\frac{1}{2}\right)^n \sum_{\substack{k \geq t \\ k \in \left\{-\frac{n}{2}, \ldots, \frac{n}{2}\right\}}} \binom{n}{k + \frac{n}{2}} \tag{2.1}$$

V tejto podkapitole sme si zvolili zástupcu diskrétnych rozdelení, takže funkcia na pravej strane výrazu 2.1 je skokovitá. Ďalším dôvodom voľby tohoto príkladu je tvar M_{Hoeff}, ktorý je zhodný ako v prípade spojitého rovnomerného rozdelenia. Dôvodom je fakt, že M_{Hoeff} neberie do úvahy rozptyl ani druhý moment náhodnej veličiny na rozdiel od ostatných nerovností, ako aj následujúceho odhadu použitím centrálnej limitnej vety

$$\mathsf{P}\left(S_n \geq t\right) = \mathsf{P}\left(\frac{2S_n}{\sqrt{n}} \geq \frac{2t}{\sqrt{n}}\right) \approx 1 - \Phi\left(\frac{2t}{\sqrt{n}}\right),$$

kde Φ opäť označuje distribučnú funkciu normovaného normálneho rozdelenia $N\,(0{,}1)$. Na grafe 2.4 môžme vidieť porovnanie ak $X_i \sim R\{-\frac{1}{2},\frac{1}{2}\}$.



Obr. 2.4: Porovnanie skutočnej pravdepodobnosti označenej PP, odhadu zostrojeného pomocou centrálnej limitnej vety a horných odhadov jednotlivých nerovností pre $n = 50$ a $n = 500$

Taktiež si všimnime niekoľko skutočností. Prvou je tá, že M_{Hoeff} sa v tomto prípade javí ako najlepší z horných odhadov. Dôvodom je už spomínaný fakt, že neberie do úvahy informáciu o druhom momente. Druhou je sila centrálnej limitnej vety. Ak máme dostatočne veľké n, jej presnosť v našich horných medziach nemá konkurenciu. Nezabúdajme však, že sa nejedná o horný odhad, ale o približnú pravdepodobnosť.

2.2 Binomické rozdelenie

Náhodná veličina X má *Binomické rozdelenie* s parametrami $n \in \mathbb{N}$ a $p \in [0{,}1]$, značíme $X \sim Bi(n{,}p)$, práve vtedy keď

$$\mathsf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \; k = 0,1,...,n. \tag{2.2}$$

Jedná sa o pravdepodobnosť toho, že v n nezávislých realizácií náhodnej veličiny, ktorej výsledok je buď úspech alebo neúspech dostaneme práve k úspechov. Ak takúto náhodnú veličinu označíme X_i, potom $X_i \sim Alt(p)$. Binomické rozdelenie je prvým príkladom necentrovaného rozdelenia.

Pretože $X \sim Bi(np)$, jednoduchým dosadením $\mathsf{E}\, X = np$ obdržíme z Hoeffdingovej nerovnosti

$$\mathsf{P}((S_n - \mathsf{E}\, S_n) \geq t) = \mathsf{P}(S_n \geq np + t) \leq \exp\left(-\frac{2t^2}{n}\right) =: M_{hoeff}$$

Z Bennettovej nerovnosti podobným spôsobom plynie odhad

$$\mathsf{P}((S_n - \mathsf{E}\, S_n) \geq t) = \mathsf{P}(S_n \geq np + t) \leq \exp\left(-npH\left(\frac{t}{np}\right)\right) =: M_{benn}$$

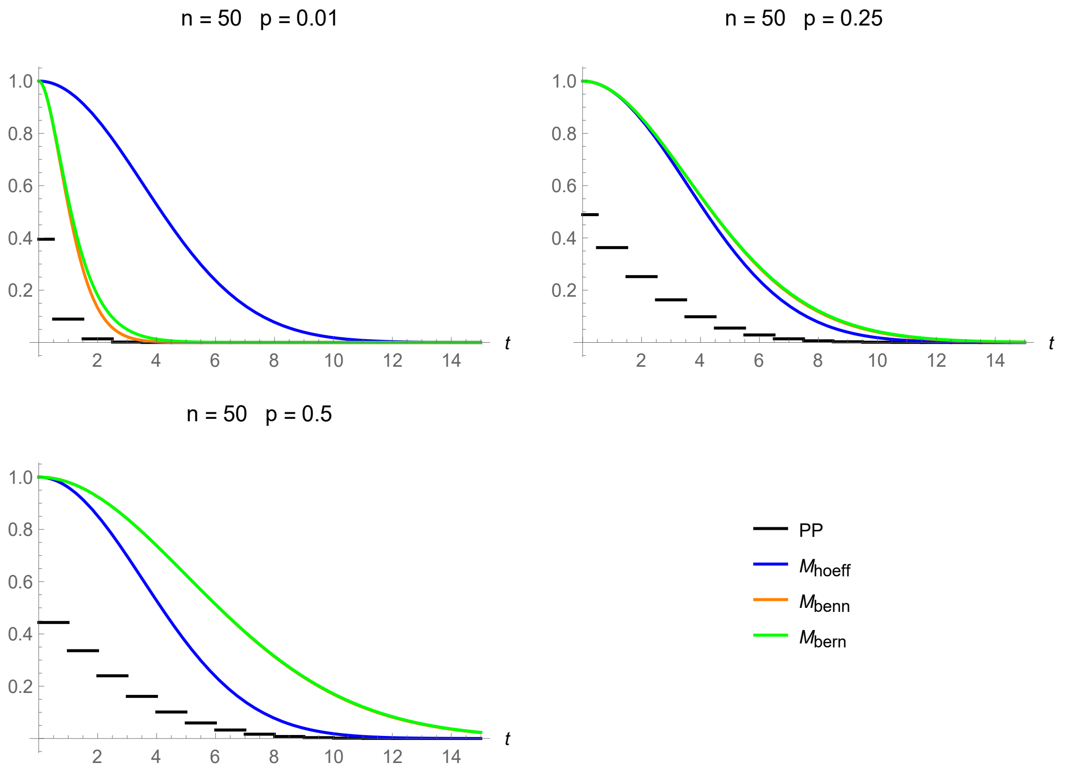
A nakoniec Bernsteinova ponúka

$$\mathsf{P}((S_n - \mathsf{E}\, S_n) \geq t) = \mathsf{P}(S_n \geq np + t) \leq \exp\left(-\frac{t^2}{2\left(np + \frac{t}{3}\right)}\right) =: M_{bern}.$$

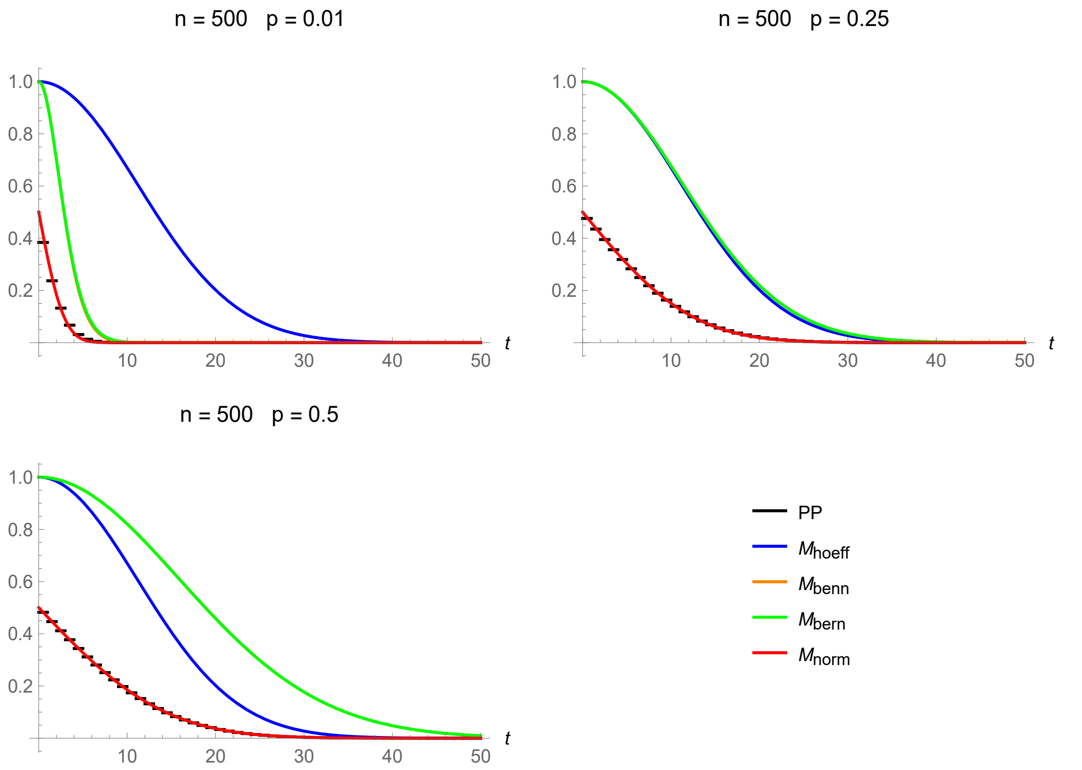
Skúmajme výrazy na pravej strane. Vezmime napríklad M_{hoeff}. Záporná exponenciála je klesajúcou funkciou, a teda čím je hodnota výrazu $\frac{2t^2}{n}$ väčšia, tým je výsledná hodnota M_{hoeff} menšia a nerovnosť je tesnejšia. Pre $p = \frac{1}{4}$ porovnajme výrazy $\frac{2t^2}{n}$ a $\frac{t^2}{2\left(np+\frac{t}{3}\right)}$

$$\begin{aligned}
\frac{2t^2}{n} &\overset{?}{\geq} \frac{t^2}{2\left(np + \frac{t}{3}\right)} \\
\frac{4}{n} &\overset{?}{\geq} \frac{1}{np + \frac{t}{3}} \\
\frac{4}{n} &\geq \frac{12}{3n + 4t}.
\end{aligned}$$

Nakoľko uvažujeme v celom texte iba t kladné, je pre $p = \frac{1}{4}$ z predošlého výpočtu už M_{hoeff} vždy tesnejšia ako M_{bern}. Na obrázkoch 2.5 a 2.6 môžme vidieť porovnanie jednotlivých medzí a skutočnej pravdepodobnosti pre rôzne hodnoty parametrov n a p. Všimnime si, ako reagujú Bennettova a Bernsteinova nerovnosť na nerovnosť na zmenu druhého momentu, resp. rozptylu. Zatiaľ čo Hoeffdingova nerovnosť dáva zrejme vždy rovnaký odhad, M_{benn} a M_{bern} sa s rastúcou hodnotou parametru p zhoršujú a tento trend pokračuje aj pre hodnoty p väčšie ako 0.5. Taktiež na niektorých grafoch nemožno vidieť M_{benn}, čo je spôsobené tým, že M_{benn} a M_{bern} sú takmer zhodné a prekrývajú sa.



Obr. 2.5: Porovnanie skutočnej pravdepodobnosti a jednotlivých medzí pre $n = 50$ a rôzne hodnoty p



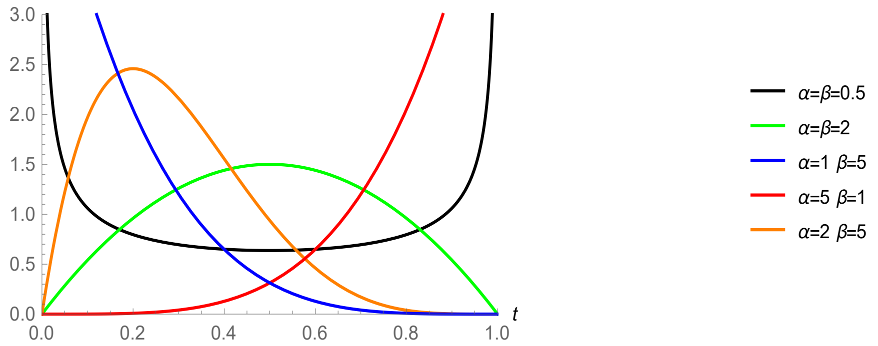
Obr. 2.6: Porovnanie skutočnej pravdepodobnosti, odhadu zostrojeného pomocou CLV a jednotlivých medzí pre $n = 500$ a rôzne hodnoty p

2.3 Beta rozdelenie

Povieme, že náhodná veličina X má *Beta rozdelenie* s parametrami α, β, značíme $X \sim Beta\,(\alpha,\beta)$, ak jej hustota pre $x \in (0,1)$ splňuje

$$f\,(x;\alpha,\beta) = \frac{1}{\mathcal{B}\,(\alpha,\beta)} x^{\alpha-1}\,(1-x)^{\beta-1}\,,$$

Kde $\mathcal{B}\,(\alpha,\beta)$ je Beta funkcia s parametrami $\alpha > 0, \beta > 0$.



Obr. 2.7: Hustoty Beta rozdelenia pre rôzne hodnoty parametrov α a β

Ekvivalentne možno *Beta rozdelenie* zadefinovať pomocou distribučnej funkcie. Náhodná veličina $X \sim Beta\,(\alpha,\beta)$, ak jej distribučná funkcia pre $x \in (0,1)$ splňuje

$$F\,(x;\alpha,\beta) = \frac{\mathcal{B}\,(x;\alpha,\beta)}{\mathcal{B}\,(\alpha,\beta)} = \mathcal{I}_x\,(\alpha,\beta)\,,$$

Kde $\mathcal{I}_x\,(\alpha,\beta)$ je regularizovaná nekompletná Beta funkcia. S týmto tvarom distribučnej funkcie nechceme pracovať, no problém je možno vyriešiť voľbou $\alpha = 1$ alebo $\beta = 1$. Pre $\beta = 1$ je napríklad

$$\mathcal{I}_x\,(\alpha,1) = x^{\alpha}.$$

Väčším problémom však je, že pre α, β obecné nie je triviálne spočítať rozdelenie súčtu nezávislých a rovnako rozdelených náhodných veličín. Pre $n = 2$ označme $S_2 = X_1 + X_2$, kde $X_1 \sim Beta(\alpha,\beta)$ a $X_2 \sim Beta(\alpha,\beta)$. Potom z poznámky 2.1 o konvolúcii môžme vyjadriť hustotu S_2 ako

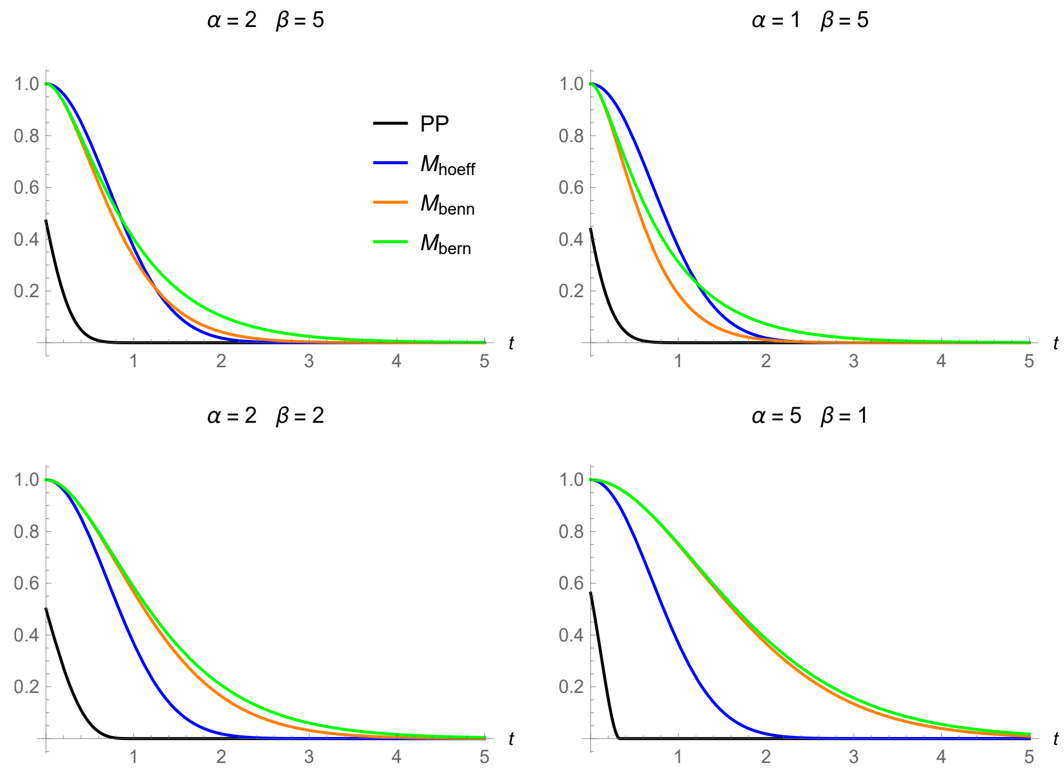
$$\begin{aligned} f_{S_2}(t) &= \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(t-x)dx \\ &= \left(\frac{1}{\mathcal{B}\,(\alpha,\beta)}\right)^2 \int_0^s x^{\alpha-1}\,(1-x)^{\beta-1}\,(t-x)^{\alpha-1}(1-(t-x))^{\beta-1}dx \end{aligned}$$

Ak by sme zvolili konkrétne hodnoty parametrov, napríklad $\alpha = \beta = 2$, dostali by sme nasledujúcu hustotu

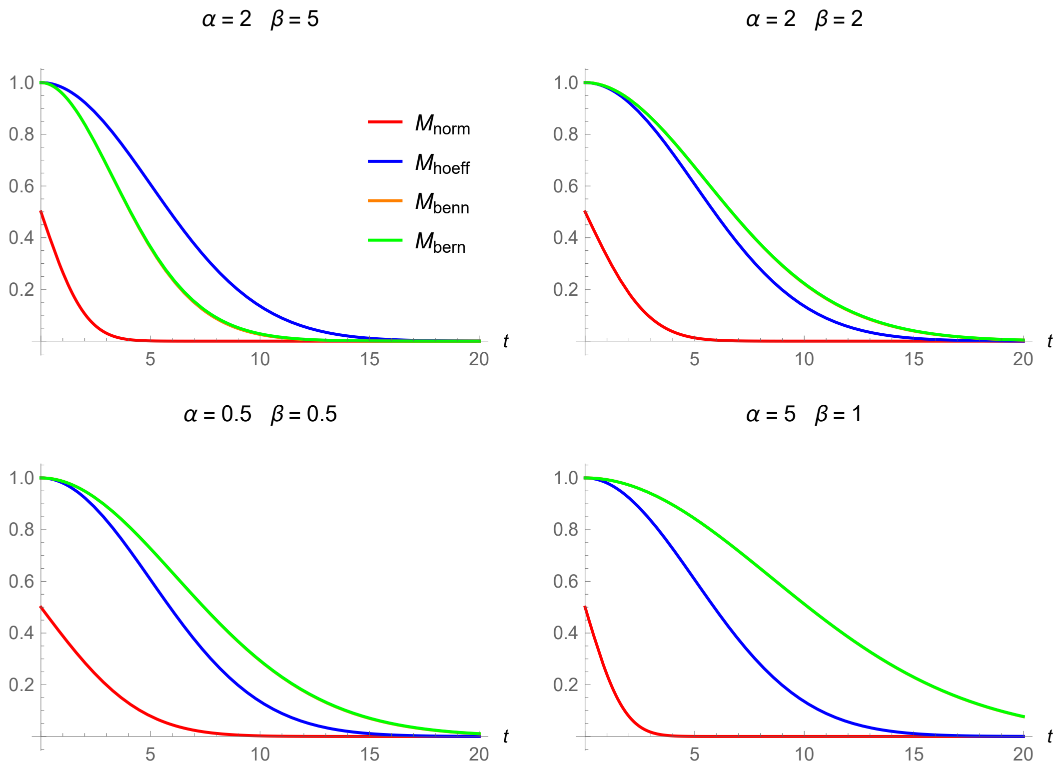
$$f_{S_2}(t) = \begin{cases} \frac{6}{5}\, t^3(t^2 - 5t + 5), & 0 < t \leq 1 \\ -\frac{6}{5}\, (t-2)^3(t^2 + t - 1), & 1 < t < 2 \\ 0\, , & inak. \end{cases}$$

Porovnanie pre $n = 2$ vidíme na obrázku 2.8. Parametre α a β volíme tak, aby zodpovedali rôznym tvarom hustôt beta rozdelenia. Pre n tak malé je ešte možné vidieť rozdiel medzi M_{benn} a M_{bern}, ktoré, ako sme už videli, sa pre n väčšie prekrývajú. Na obrázku 2.7 si všimnime rozloženie pravdepodobnostnej masy pre rôzne parametre α,β. Odhady jednotlivých nerovností v tomto prípade totiž súvisia s odhadmi pri binomickom rozdelení a to tak, že ak koncentrujeme pravdepodobnostnú masu blízko 0, sú M_{benn} a M_{bern} lepšími odhadmi ako v prípade, kde koncentrujeme pravdepodobnosť blízko 1. Pri binomickom rozdelení to odpovedá hodnote parametru p - čím je väčšia, tým väčšia pravdepodobnosť náleží 1.

Niečo podobné môžme vidieť aj na obrázku 2.9, kde namiesto presnej pravdepodobnosti používame odhad zostrojený pomocou CLV.



Obr. 2.8: Porovnanie skutočnej pravdepodobnosti a jednotlivých medzí pre $n = 2$ a rôzne hodnoty parametrov α a β



Obr. 2.9: Porovnanie pravdepodobnosti odhadnutej pomocou CLV a jednotlivých medzí pre $n = 100$ a rôzne hodnoty parametrov α a β

Záver

V kapitole 1 sme sa zaoberali odhadmi pre ohraničené náhodné veličiny. Boli tu formulované a podrobne dokázané Hoeffdingova 1.2, Bennettova 1.3 a Bernsteinova 1.4 nerovnosť a zmienené aj Chernoffove medze.

V kapitole 2 sme potom tieto nerovnosti aplikovali na jednotlivé rozdelenia. Začali sme s rovnomerným rozdelením ako zástupcom jednoduchého spojitého rozdelenia, kde bol odhad daný Bennettovou nerovnosťou najlepší. V ďalšej časti sme uvažovali diskrétne rovnomerné rozdelenie ako zástupcu diskrétneho rozdelenia. V tomto prípade sme nahrnuli pravdepodobnostnú masu na kraje intervalu, pričom sa odhady jednotlivých nerovností zdali približne rovnocenné. Nasledovalo binomické rozdelenie, kde presnosť odhadu závisela na hodnote parametru p a to tak, že s jeho rastom sa odhady dané Bennettovou a Bernsteinovou nerovnosťou zhoršovali. Zástupcom obecnejšieho rozdelenia bolo Beta rozdelenie, pre ktoré sme manipuláciou s parametrami α,β sústredili pravdepodobnostnú masu na rôzne časti intervalu (0,1). Pre väčšie hodnoty parametru β sa javili lepšie Bennettova a Bernsteinova nerovnosť, zatiaľ čo pre väčšie hodnoty parametru α dominovala Hoeffdingova nerovnosť.

Literatúra

[1] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*, **57**(297), 33–45.

[2] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, **58**(301), 13–30.

[3] Klein, T. (2014). Basic concentration inequalities. page 8. URL `https://perso.math.univ-toulouse.fr/klein/files/2014/07/Redice_english.pdf`.

[4] Massart, P. (2007). *Concentration inequalities and model selection: Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003*. Springer.

Zoznam obrázkov