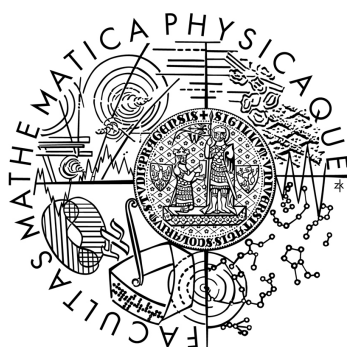


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Katarína Brisudová

Neceločíselné momenty náhodných veličin

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Zbyněk Pawlas, Ph.D.

Studijní program: Matematika

Studijní obor: Finanční matematika

Praha 2014

Moje poďakovanie patrí vedúcemu práce doc. RNDr. Zbyňkovi Pawlasovi, Ph.D. za trpezlivosť, ochotu a cenné pripomienky pri vedení a záverečnom spracovaní bakalárskej práce. Ďalej by som sa rada poďakovala všetkým, ktorí ma počas štúdia neustále podporovali a posúvali vpred.

Prohlašuji, že jsem tuto bakalářskou práci vypracovala 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 4.12.2014

Katarína Brisudová

Název práce: Neceločíselné momenty náhodných veličin

Autor: Katarína Brisudová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: doc. RNDr. Zbyněk Pawlas, Ph.D.

Abstrakt: Cílem práce je zformulovat problematiku neceločíselných momentů náhodných veličin. Pro základní diskrétní a spojitá rozdělení jsou spočteny neceločíselné momenty. Tyto výpočty jsou prováděny analyticky nebo numericky vhodným softwarem v případě, že neexistuje jednoduchý tvar. V práci je zformulovaný i princip momentové metody a její variace s použitím neceločíselných momentů místo celočíselných a také se diskutuje o efektivitě této variace.

Klíčová slova: náhodná veličina, neceločíselný moment, odhad, momentová metoda

Title: Fractional moments of random variables

Author: Katarína Brisudová

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Zbyněk Pawlas, Ph.D.

Abstract: The aim of this thesis is to formulate issues regarding fractional moments of random variables. Fractional moments are calculated for basic discrete and continuous distributions. These calculations are performed analytically or numerically using an appropriate software if a simple form does not exist. The thesis also formulates the principle of method of moments and its variations using fractional moments instead of integers and the effectiveness of this variation is also discussed.

Keywords: random variable, fractional moment, estimate, method of moments

Názov práce: Neceločíselné momenty náhodných veličín

Autor: Katarína Brisudová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: doc. RNDr. Zbyněk Pawlas, Ph.D.

Abstrakt: Cieľom práce je zformulovať problematiku neceločíselných momentov náhodných veličín. Pre základné diskrétne a spojité rozdelenia sú spočítané neceločíselné momenty. Tieto výpočty sú počítané analyticky alebo numericky vhodným softvérom v prípade, že neexistuje jednoduchý tvar. V práci je zformulovaný aj princíp momentovej metódy a jej variácia s použitím neceločíselných momentov namiesto celočíselných a taktiež diskutuje sa o efektivite tejto variácie.

Kľúčové slová: náhodná veličina, neceločíselný moment, odhad, momentová metóda

Obsah

Úvod	2
1 Základné pojmy a definície	3
2 Diskrétné rozdelenia	5
2.1 Alternatívne rozdelenie	5
2.2 Binomické rozdelenie	5
2.3 Poissonovo rozdelenie	8
2.4 Rovnomerné rozdelenie	10
3 Spojité rozdelenia	11
3.1 Normované normálne rozdelenie	11
3.2 Normálne rozdelenie	13
3.3 Gamma rozdelenie	13
3.4 Paretovo rozdelenie	14
3.5 Rovnomerné rozdelenie	16
4 Momentová metóda	17
4.1 Motivácia	17
4.2 Princíp momentovej metódy	17
4.3 Simulačný experiment	19
Záver	22
Zoznam použitej literatúry	23

Úvod

Medzi dôležité charakteristiky náhodných veličín bezpochyby patria ich momenty. V teórii pravdepodobnosti sa najčastejšie počítajú momenty prvého a druhého rádu, a dosť často ešte momenty tretieho a štvrtého rádu. Vyššie momenty sa používajú len málokedy. Sú však prípady, keď momenty prvého, prípadne druhého rádu neexistujú. To je charakteristické hlavne pre rozdelenia s ťažkými chvostmi. Cieľom tejto práce je prehľadne ukázať výpočty neceločíselných momentov vybraných diskretných a spojitých rozdelení. Budeme sa ich snažiť spočítať predovšetkým analyticky a v prípade, keď nebude existovať jednoduchý tvar, použijeme vhodný softvér.

Tiež sa budeme zaoberať otázkou použitia neceločíselných momentov v momentovej metóde a budeme sledovať presnosť odhadov za použitia celočíselných i neceločíselných momentov.

Práca je členená do štyroch kapitol. V prvej kapitole sú uvedené základné definície z oblasti teórie pravdepodobnosti, ktoré slúžia ako teoretický základ pre prácu s náhodnými veličinami a ich momentmi.

Druhá kapitola obsahuje štyri diskretné rozdelenia – alternatívne, binomické, Poissonovo a rovnomerné a opisuje ich základné vlastnosti a momenty.

Tretia kapitola nadväzuje na druhú so spojitými rozdeleniami a takisto opisuje ich základné charakteristiky a momenty.

Vo štvrtej kapitole rozoberáme princíp momentovej metódy ako metódy pre odhad parametrov vybraných rozdelení. Zároveň skúmame využitie neceločíselných momentov a porovnáваме ich efektívnosť v súvislosti s klasickými celočíselnými momentmi v zmysle strednej štvorcovej chyby odhadov.

1. Základné pojmy a definície

Táto kapitola obsahuje zopár základných definícií, ktorých znalosť je potrebná k pochopeniu počítania momentov náhodných veličín. Definície vychádzajú z knižiek [1] a [6], ktoré okrem iného obsahujú aj ďalšie definície a tvrdenia týkajúce sa tejto problematiky. My si však spomenieme iba tie najnevyhnutnejšie.

Definícia 1.1. *Dvojica (Ω, \mathcal{A}) , kde \mathcal{A} je nejaká σ -algebra podmnožín množiny Ω , sa nazýva **merateľný priestor**.*

Definícia 1.2. *Nech (Ω, \mathcal{A}) a $(\mathcal{X}, \mathcal{B})$ sú merateľné priestory. Zobrazenie $X : \Omega \rightarrow \mathcal{X}$ sa nazýva (\mathcal{A} -) **merateľné zobrazenie** práve vtedy, keď platí*

$$B \in \mathcal{B} \Rightarrow X^{-1}(B) \in \mathcal{A}.$$

Zapisujeme $X : (\Omega, \mathcal{A}) \rightarrow (\mathcal{X}, \mathcal{B})$.

Definícia 1.3. ***Pravdepodobnosť** P je definovaná ako miera na \mathcal{A} s vlastnosťou $P(\Omega) = 1$, t.j. P je množinová funkcia na \mathcal{A} s vlastnosťami:*

1. $P(A) \geq 0, A \in \mathcal{A}$;
2. $P(\Omega) = 1, P(\emptyset) = 0$;
3. $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$, ak je $\{A_n\}$ postupnosť po dvoch disjunktných javov.

Trojica (Ω, \mathcal{A}, P) sa nazýva pravdepodobnostný priestor.

Pre nasledujúce tvrdenia bude platiť:

Nech pravdepodobnostný priestor (Ω, \mathcal{A}, P) je pevne daný.

Nech \mathbb{R} je množina reálnych čísel a \mathcal{B} je borelovská σ -algebra jej podmnožín.

Definícia 1.4. *Merateľné zobrazenie $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ sa nazýva **náhodná veličina**.*

Definícia 1.5. *Nech (Ω, \mathcal{A}, P) je pravdepodobnostný priestor a $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B})$ je náhodná veličina.*

1. *Nech existuje postupnosť reálnych čísel x_1, x_2, \dots taká, že platí*

$$\sum_{i=1}^{\infty} P[X = x_i] = 1.$$

*Zoznam hodnôt $\{x_n\}$, ktoré náhodná veličina nadobúda a zoznam pravdepodobností $\{p_n\}$, s ktorými tieto hodnoty nadobúda, udáva **diskrétne rozdelenie pravdepodobnosti**.*

2. Nech $f : \Omega \rightarrow \mathbb{R}$ je lebesgueovsky merateľná nezáporná funkcia taká, že $\int_{\Omega} f(\omega) d\omega = 1$ a $P[X \in B] = \int_B f(x) dx$, potom funkciu f nazývame **hustota pravdepodobnosti** a udáva **spojité rozdelenie pravdepodobnosti**.

Definícia 1.6. Nech X je náhodná veličina definovaná na pravdepodobnostnom priestore (Ω, \mathcal{A}, P) ; jej **stredná hodnota** $\mathbb{E}X$ je definovaná ako integrál z X vzhľadom na mieru P , tj.

$$\mathbb{E}X = \int_{\Omega} X dP = \int_{\Omega} X(\omega) dP(\omega).$$

Tvrdenie 1.1. Nech X je náhodná veličina a nech $\phi : \mathbb{R} \rightarrow \mathbb{R}$.

1. Ak má náhodná veličina X diskrétné rozdelenie $\{x_n, p_n\}_{n \in \mathbb{N}_0}$, potom

$$\mathbb{E} \phi(X) = \sum_{n \in \mathbb{N}_0} \phi(x_n) p_n,$$

pokiaľ jedna zo strán rovnosti existuje.

2. Ak má náhodná veličina X absolútne spojité rozdelenie s hustotou f , potom

$$\mathbb{E} \phi(X) = \int_{-\infty}^{\infty} \phi(x) f(x) dx,$$

pokiaľ integrál existuje.

Definícia 1.7. Nech n je prirodzené číslo. Potom

- **n -ty moment náhodnej veličiny X** je definovaný predpisom $\mathbb{E}X^n$;
- **n -ty absolútny moment náhodnej veličiny X** je definovaný predpisom $\mathbb{E}|X|^n$;
- **n -ty centrálny moment náhodnej veličiny X** je definovaný predpisom $\mathbb{E}(X - \mathbb{E}X)^n$;
- **n -ty absolútny centrálny moment náhodnej veličiny X** je definovaný predpisom $\mathbb{E}|X - \mathbb{E}X|^n$.

Nás budú zaujímať hlavne neceločíselné momenty. Budeme počítať neceločíselné momenty náhodnej veličiny X rádu r a podľa definície 1.7. značiť ako $\mathbb{E}X^r$, $\mathbb{E}|X|^r$, $\mathbb{E}|X - a|^r$, kde $r \geq 0$ a $a \in \mathbb{R}$.

V kapitole 2 priberieme základné diskkrétne rozdelenia a kapitola 3 bude venovaná spojitým rozdeleniam.

2. Diskrétne rozdelenia

2.1 Alternatívne rozdelenie

Alternatívne rozdelenie, tiež nazývané Bernoulliovo alebo nula-jedničkové rozdelenie, je rozdelenie náhodnej veličiny X , ktorá nadobúda iba hodnoty 0 a 1 s pravdepodobnosťami $1 - p$ a p . Číslo p označuje parameter alternatívneho rozdelenia, $0 < p < 1$.

Z pravdepodobnostného rozdelenia náhodnej veličiny $X \sim Alt(p)$, ktoré je dané nasledovne:

$$P[X = j] = p^j(1 - p)^{1-p}, \quad j \in \{0, 1\},$$

ľahko vypočítame strednú hodnotu $\mathbb{E}X = 0 \cdot (1 - p) + 1 \cdot p = p$ a rozptyl zo vzťahu $\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 0^2 \cdot (1 - p) + 1^2 \cdot p - p^2 = p(1 - p)$.

Pozrieme sa, ako bude vyzerat' absolútny moment rádu r veličiny $X \sim Alt(p)$:

$$\mathbb{E}|X|^r = \mathbb{E}X^r = 0^r \cdot (1 - p) + 1^r \cdot p = p, \quad \forall r > 0.$$

Absolútny neceločíselný moment rádu r náhodnej veličiny X , $\mathbb{E}|X|^r$, teda závisí iba na parametri p a rovná sa tomuto parametru pre každé $r > 0$.

2.2 Binomické rozdelenie

Nech $Y_i \sim Alt(p)$ sú nezávislé náhodné veličiny pre $i = 1, \dots, n$. Súčet týchto nezávislých alternatívnych veličín, teda počet úspechov medzi n pokusmi, má binomické rozdelenie: $\sum_{i=1}^n Y_i \sim Bi(n, p)$.

Binomické rozdelenie je teda rozdelenie náhodnej veličiny X , ktorá nadobúda hodnoty $k = 0, 1, \dots, n$ s pravdepodobnosťami $P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$. Toto rozdelenie má dva parametre : $n \in \mathbb{N}$ a $0 < p < 1$.

Strednú hodnotu a rozptyl náhodnej veličiny $X \sim Bi(n, p)$ vypočítame nasledovne:

$$\begin{aligned} \mathbb{E}X &= \sum_{k=0}^n k \binom{n}{k} p^k (1 - p)^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np(p + (1-p))^{n-1} \\ &= np. \end{aligned}$$

Pomôžeme si tvrdením, že $\mathbb{E}X^2 = \mathbb{E}X(X-1) + \mathbb{E}X$, a preto stačí vypočítat':

$$\begin{aligned}
\mathbb{E}X(X-1) &= \sum_{k=2}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} \\
&= n(n-1)p^2 \sum_{k=2}^n \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} (1-p)^{n-k} \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j (1-p)^{n-2-j} \\
&= n(n-1)p^2.
\end{aligned}$$

Odtiaľ dostaneme

$$\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p).$$

Neceločíselné momenty

V tejto časti počítame r -ty moment náhodnej veličiny $X \sim Bi(n, p)$, kde $r > 0$. Podľa tvrdenia 1.1. je

$$\mathbb{E}|X|^r = \mathbb{E}X^r = \sum_{k=0}^n k^r \binom{n}{k} p^k (1-p)^{n-k}.$$

Pri počítaní momentov náhodnej veličiny $X \sim Bi(n, p)$ sme volili konkrétne parametre n a p , aby sme mohli závislosť $\mathbb{E}|X|^r$ na voľbe momentu r znázorniť graficky, pretože tieto momenty sa všeobecne nedajú vyjadriť v jednoduchom tvare.

Za parameter n sme volili $n = 20$ a v konečnom dôsledku sme numericky počítali

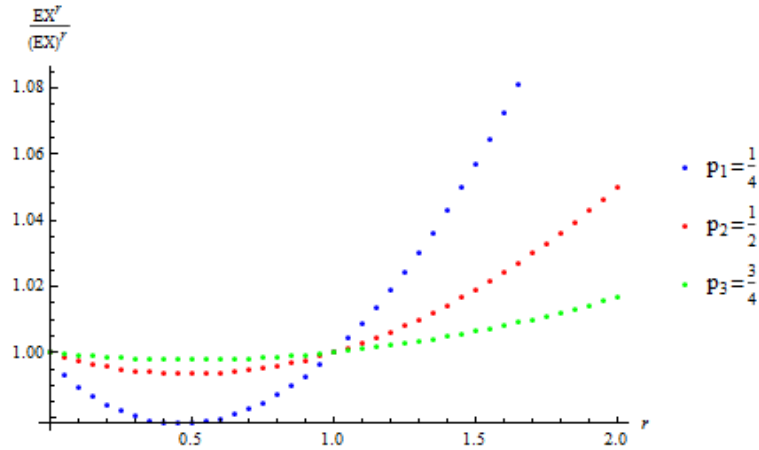
$$\mathbb{E}|X|^r = \mathbb{E}X^r = \sum_{k=0}^{20} k^r \binom{20}{k} p^k (1-p)^{20-k}$$

pre tri rôzne hodnoty parametra p , konkrétne $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{2}$ a $p_3 = \frac{3}{4}$.

Hodnota $\mathbb{E}X^r$ so vzrastajúcim parametrom r rástla pomerne rýchlo, a preto sme ju v grafickom znázornení na obr. 2.1 normovali hodnotou $(\mathbb{E}X)^r$.

Pre $r = 0$ sme dodefinovali $\mathbb{E}X^r = 1$, a tak je v grafe na obr. 2.1 $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r} = 1$ pre $r = 0$ a $r = 1$. Všetky funkcie majú minimum približne v rovnakom mieste, no pozorujeme, že s rastúcou pravdepodobnosťou p sa funkcie čoraz viac približujú ku konštantnej funkcii $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r} = 1$. Môžeme povedať, že so zväčšujúcou sa pravdepodobnosťou p klesá závislosť $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ na r .

Pri binomickom rozdelení sme ešte skúmali závislosť neceločíselného momentu $\mathbb{E}|X - a|^r$, $a \in \mathbb{R}$, pre štyri rôzne hodnoty parametra r , konkrétne pre $r_1 = \frac{1}{2}$, $r_2 = 1$, $r_3 = \frac{3}{2}$ a $r_4 = 2$. Výpočty vychádzajú z troch rôznych hodnôt parametra p , podobne ako v predchádzajúcom prípade, teda $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{2}$ a $p_3 = \frac{3}{4}$. Výsledné grafy sú znázornené na obr. 2.2.



Obr. 2.1: Graf zobrazujúci priebeh $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ v závislosti na r pre binomické rozdelenie s parametrami $n = 20$ a $p \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$.

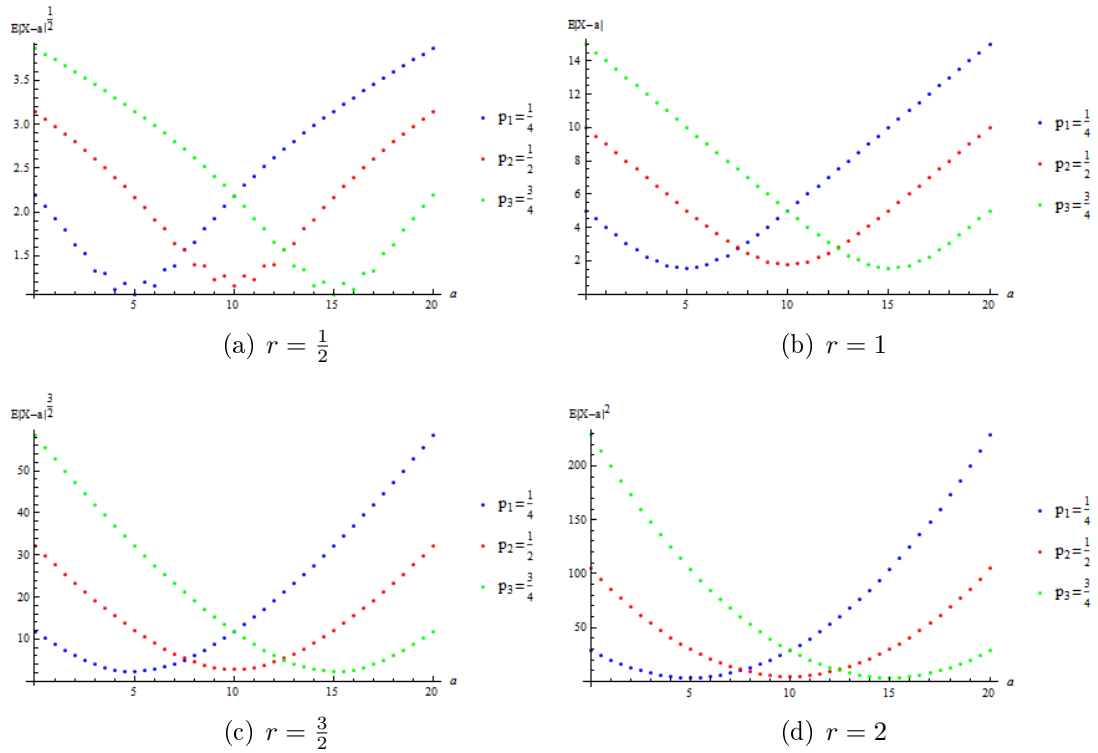
Pre $a \in \mathbb{R}$ sa dá vyjadriť

$$\begin{aligned} \mathbb{E}(X - a)^2 &= \mathbb{E}(X - \mathbb{E}X)^2 + (\mathbb{E}X - a)^2 \\ &= \text{var}X + (\mathbb{E}X - a)^2 \\ &= np(1 - p) + (np - a)^2 \end{aligned}$$

z čoho jasne vidno, že pre $a = \mathbb{E}X$ a $r = 2$ dostaneme druhý centrálny moment náhodnej veličiny $X \sim Bi(n, p)$.

Na obr. 2.2(d) vidíme spomínaný špeciálny prípad, keď $r = 2$. Funkcie majú tvar paraboly a dosahujú minimum v bode $a = np$ s funkčnou hodnotou $\text{var}X = np(1 - p)$.

Graf na obr. 2.2(c) má podobný priebeh ale iný tvar, a teda už nemôžeme hovoriť o parabole. Graf zobrazujúci prvý absolútny moment $\mathbb{E}|X - a|$ na obr. 2.2(b) pripomína graf absolútnej hodnoty s minimom v bodoch $a = np$ a graf na obr. 2.2(a) má tiež minimum v bodoch $a = np$, avšak tieto funkcie už nie sú konvexné. Grafy pre $p_1 = \frac{1}{4}$ a $p_3 = \frac{3}{4}$ sú osovo súmerné.



Obr. 2.2: Graf zobrazujúci priebeh $\mathbb{E}|X - a|^r$ v závislosti na a pre binomické rozdelenie s parametrami $n = 20$ a $p \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$.

2.3 Poissonovo rozdelenie

Poissonovo rozdelenie dostaneme limitným prechodom z binomického rozdelenia pre $n \rightarrow \infty$ a $np \rightarrow \lambda < \infty \Rightarrow Bi(n, p)$ konverguje k $Po(\lambda)$.

Poissonovo rozdelenie je rozdelenie náhodnej veličiny X , ktorá nadobúda hodnoty $k = 0, 1, \dots$ s pravdepodobnosťami $P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$. Číslo $\lambda > 0$ je parameter. Stredná hodnota a rozptyl náhodnej veličiny $X \sim Po(\lambda)$ sa vypočítajú nasledovne:

$$\begin{aligned}
 \mathbb{E}X &= \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
 &= \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\
 &= \lambda.
 \end{aligned}$$

Rozptyl vypočítame rovnakým spôsobom ako pri binomickom rozdelení:

$$\begin{aligned}
\mathbb{E}X(X-1) &= \sum_{k=2}^{\infty} k(k-1)e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \lambda^2 \sum_{k=2}^{\infty} e^{-\lambda} \frac{\lambda^{k-2}}{(k-2)!} \\
&= \lambda^2 e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\
&= \lambda^2.
\end{aligned}$$

A odtiaľ $\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

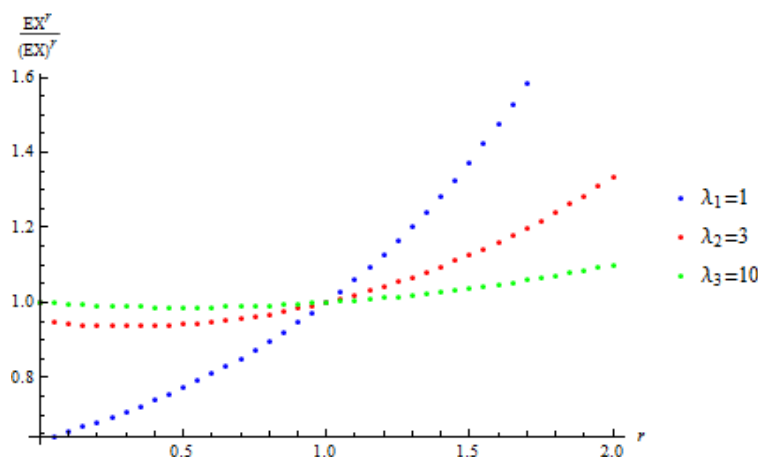
Neceločíselné momenty

Opäť počítame r -ty moment náhodnej veličiny X , tentokrát $X \sim Po(\lambda)$ a $r > 0$ ostáva. Podľa tvrdenia 1.1.

$$\mathbb{E}|X|^r = \mathbb{E}X^r = \sum_{k=0}^{\infty} k^r e^{-\lambda} \frac{\lambda^k}{k!}.$$

Výpočet vyžadoval dosadenie jedného parametra, a to parametra λ . Pre znázornenie sme volili hodnoty $\lambda_1 = 1$, $\lambda_2 = 3$ a $\lambda_3 = 10$. Parameter r sa pohybuje v rozmedzí 0 až 2 s krokom 0,05 rovnako ako v predchádzajúcom prípade.

Výsledné hodnoty $\mathbb{E}|X|^r$ boli opäť pre rýchle stúpanie s rastúcim parametrom r normované hodnotou $(\mathbb{E}X)^r$.



Obr. 2.3: Graf zobrazujúci priebeh $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ v závislosti na r pre Poissonovo rozdelenie s hodnotami parametrov $\lambda_1 = 1$, $\lambda_2 = 3$, $\lambda_3 = 10$.

V bode $r = 0$ sme opäť dodefinovali $\mathbb{E}X^r = 1$. Funkcia $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ tiež nadobúda hodnotu 1 pre $r = 1$ pri všetkých hodnotách parametra λ . Z grafu na obr. 2.3 vidíme, že funkcia najskôr klesá (rapídnejšie pre menšie λ) a od istého $r \in (0, 1)$ má funkcia rastúci charakter. S rastúcim parametrom λ má krivka tendenciu konvergovať k priamke rovnobežnej s osou r . Môžeme teda poznamenať, že pre zväčšujúci sa parameter sa znižuje závislosť na hodnote r .

2.4 Rovnomerné rozdelenie

Diskrétné rovnomerné rozdelenie $R(T)$, kde $T = \{1, \dots, n\}$ je neprázdna konečná množina, je rozdelenie náhodnej veličiny X , ktoré sa dá popísať takto:

$$P[X = x] = \begin{cases} \frac{1}{n}, & x \in T, \\ 0, & x \notin T; \end{cases}$$

tj. náhodná veličina nadobúda všetky hodnoty z T s rovnakou pravdepodobnosťou. Stredná hodnota a rozptyl náhodnej veličiny $X \sim R(T)$ sú:

$$\mathbb{E}X = \frac{1}{n} \sum_{k=1}^n k = \frac{n+1}{2},$$

$$\text{var}X = \frac{1}{n} \sum_{k=1}^n (k - \mathbb{E}X)^2 = \frac{n(n+1)}{12}.$$

Neceločíselné momenty

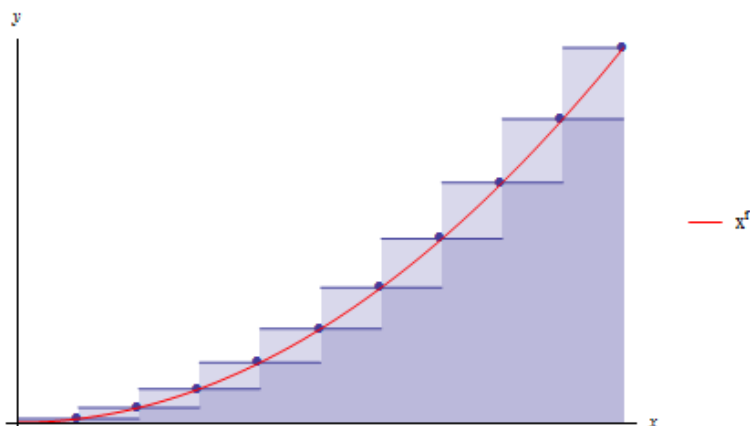
Pre neceločíselný moment rádu r náhodnej veličiny X , ktorá má rovnomerné rozdelenie na množine $T = \{1, \dots, n\}$, platí

$$\mathbb{E}|X|^r = \mathbb{E}X^r = \frac{1}{n} \sum_{k=1}^n k^r = \frac{1}{n} (1 + 2^r + 3^r + \dots + n^r) \approx \frac{n^r}{r+1}.$$

Aproximácia $\frac{n^r}{r+1}$ plynie z nasledujúceho horného a dolného odhadu:

$$\frac{n^r}{r+1} = \frac{1}{n} \int_0^n x^r dx \leq \mathbb{E}X^r \leq \frac{1}{n} \int_1^{n+1} x^r dx = \frac{(n+1)^{(r+1)} - 1}{n(r+1)}.$$

Horným a dolným odhadom, ktorý je napísaný vyššie, vlastne porovnávame plochu pod grafom a grafické zobrazenie vidíme na obr. 2.4.



Obr. 2.4: Ohraničenie $\mathbb{E}X^r$ zhora a zdola.

3. Spojité rozdelenia

3.1 Normované normálne rozdelenie

Náhodná veličina X má normované normálne rozdelenie práve vtedy, keď je definovaná hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty.$$

Skutočnosť, že náhodná veličina má práve normované normálne rozdelenie značíme $X \sim N(0, 1)$.

Stredná hodnota a rozptyl náhodnej veličiny $X \sim N(0, 1)$ sú $\mathbb{E}X = 0$ a $\text{var} X = 1$.

Nulovosť strednej hodnoty plynie z toho, že funkcia $\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}}$ je nepárna a integrovateľná v \mathbb{R} . Táto vlastnosť by sa dala rozšíriť na všetky nepárne momenty, tj. $\mathbb{E}X = \mathbb{E}X^{2k-1} = 0 \quad \forall k = 1, 2, \dots$, pretože funkcia $\frac{1}{\sqrt{2\pi}} x^{2k-1} e^{-\frac{x^2}{2}}$ je nepárna a integrovateľná v \mathbb{R} .

Vzhľadom na to, že stredná hodnota $\mathbb{E}X$ náhodnej veličiny $X \sim N(0, 1)$ sa rovná 0, rozptyl $\text{var}X$ sa bude rovnať druhému momentu $\mathbb{E}X^2$. Ten vypočítame nasledovne:

$$\begin{aligned} \text{var}X = \mathbb{E}X^2 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^2 e^{-\frac{x^2}{2}} dx \\ &= \frac{2}{\sqrt{\pi}} \int_0^{\infty} \sqrt{y} e^{-y} dy \\ &= \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1, \end{aligned}$$

kde sme v druhom riadku využili fakt, že $\frac{2}{\sqrt{2\pi}} x^2 e^{-\frac{x^2}{2}}$ je párna funkcia a následne sme pomocou substitúcie $y = \frac{x^2}{2}$ dostali úlohu do tvaru gamma funkcie.

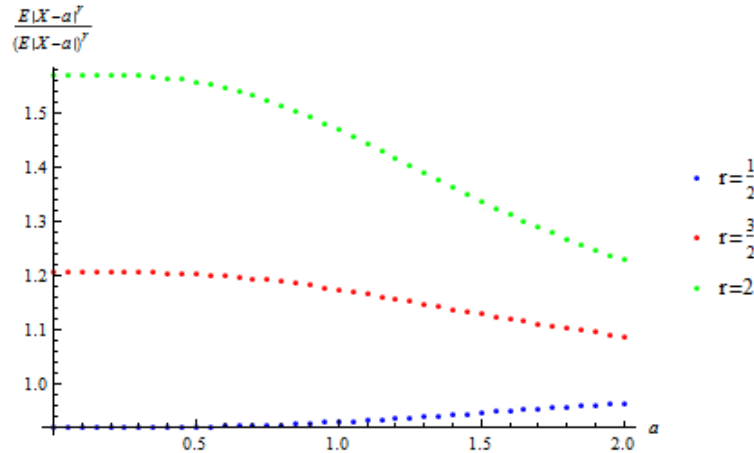
Neceločíselné momenty

Aj tu máme za úlohu vypočítať r -ty absolútny neceločíselný moment náhodnej veličiny $X \sim N(0, 1)$, pre $r \geq 0$. Postupujeme nasledovne:

$$\begin{aligned}
\mathbb{E}|X|^r &= \int_{-\infty}^{\infty} |x|^r \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= 2 \int_0^{\infty} x^r \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\
&= 2 \int_0^{\infty} (\sqrt{2y})^{r-1} \frac{1}{\sqrt{2\pi}} e^{-y} dy \\
&= \frac{(\sqrt{2})^r}{\sqrt{\pi}} \int_0^{\infty} y^{\frac{1}{2}(r+1)-1} e^{-y} dy \\
&= \frac{(\sqrt{2})^r}{\sqrt{\pi}} \Gamma\left(\frac{r+1}{2}\right).
\end{aligned}$$

Opätovným využitím substitúcie $y = \frac{x^2}{2}$ sme výpočet previedli do tvaru pre gamma funkciu.

Pre toto rozdelenie si môžeme ukázať, rovnako ako pri binomickom rozdelení, aj správanie sa neceločíselného momentu $\mathbb{E}|X - a|^r$ v závislosti na parametri $a \in \mathbb{R}$ pre rôzne r . Vo výpočtoch a grafickom znázornení volíme parameter a z intervalu $[0, 2]$. Hodnota $\mathbb{E}|X - a|^r$ bola normovaná hodnotou $(\mathbb{E}|X - a|)^r$. Výsledný graf vidíme na obr. 3.1.



Obr. 3.1: Graf zobrazujúci priebeh $\frac{\mathbb{E}|X-a|^r}{(\mathbb{E}|X-a|)^r}$ v závislosti na a pre normálne rozdelenie a rôzne momenty $r_1 = \frac{1}{2}, r_2 = \frac{3}{2}, r_3 = 2$.

Na obr.3.1 vidíme dôsledok Jensenovej nerovnosti:

$$\mathbb{E}|X - a|^r \geq (\mathbb{E}|X - a|)^r \quad \text{pre } r \geq 1,$$

$$\mathbb{E}|X - a|^r \leq (\mathbb{E}|X - a|)^r \quad \text{pre } r \leq 1.$$

Z nasledujúceho navyiac vyplýva, že pre $a \rightarrow \infty$ ide $\frac{\mathbb{E}|X-a|^r}{(\mathbb{E}|X-a|)^r}$ k 1 :

$$\lim_{a \rightarrow \infty} \frac{\mathbb{E}|X - a|^r}{(\mathbb{E}|X - a|)^r} = \lim_{a \rightarrow \infty} \frac{\mathbb{E}\left|\frac{X}{a} - 1\right|^r}{\left(\mathbb{E}\left|\frac{X}{a} - 1\right|\right)^r} = 1, \quad \text{kde } \frac{X}{a} \sim N\left(0, \frac{1}{a^2}\right).$$

3.2 Normálne rozdelenie

Náhodná veličina X má normálne rozdelenie s parametrami μ, σ^2 , $X \sim N(\mu, \sigma^2)$, práve vtedy, keď je definovaná hustotou

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma^2 > 0.$$

Stredná hodnota a rozptyl náhodnej veličiny $X \sim N(\mu, \sigma^2)$ sa rovnako ako pri normovanom normálnom rozdelení rovnajú parametrom normálneho rozdelenia, teda $\mathbb{E}X = \mu$ a $\text{var}X = \sigma^2$. Ukáže sa to nasledujúcim výpočtom:

$$\begin{aligned} \mathbb{E}X &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma y) e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu e^{-\frac{y^2}{2}} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma y e^{-\frac{y^2}{2}} dy \\ &= \mu. \end{aligned}$$

V druhom kroku sme využili substitúciu $y = \frac{x-\mu}{\sigma}$, čím sme dostali hustoty normovaného normálneho rozdelenia a tým pádom sme v poslednom kroku mohli využiť fakt, že $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}} dy = 0$. Okrem toho sme využili rovnosť $\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi}$. K vyjadreniu rozptylu vypočítame druhý moment náhodnej veličiny X ako

$$\begin{aligned} \mathbb{E}X^2 &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma y)^2 e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu^2 e^{-\frac{y^2}{2}} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 2\sigma y e^{-\frac{y^2}{2}} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 y^2 e^{-\frac{y^2}{2}} dy \\ &= \sigma^2 + \mu^2. \end{aligned}$$

Odtiaľ $\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \sigma^2$.

Neceločíselné momenty

Neceločíselné momenty náhodnej veličiny X z rozdelenia $N(\mu, \sigma^2)$ sú špeciálnym prípadom momentov normovaného normálneho rozdelenia. Nech $Y \sim N(0, 1)$. Potom $X = \sigma Y + \mu \sim N(\mu, \sigma^2)$ a $\mathbb{E}|X|^r = \mathbb{E}|\sigma Y + \mu|^r = \sigma^r \mathbb{E}|Y + \frac{\mu}{\sigma}|^r$.

3.3 Gamma rozdelenie

Nech X je náhodná veličina a nech $\alpha > 0$ a $\beta > 0$. Potom hovoríme, že náhodná veličina X má gamma rozdelenie, $X \sim \Gamma(\alpha, \beta)$, práve vtedy, keď je definovaná hustotou

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{I}_{(0, \infty)}(x), \quad x \in \mathbb{R}.$$

Špeciálnym prípadom gamma rozdelenia je exponenciálne rozdelenie. To dostaneme pri voľbe parametra $\alpha = 1$, a teda je definované hustotou

$$f(x) = \beta e^{-\beta x} \mathbb{I}_{(0, \infty)}(x), \quad x \in \mathbb{R}.$$

Stredná hodnota náhodnej veličiny $X \sim \Gamma(\alpha, \beta)$ je rovná $\frac{\alpha}{\beta}$ a rozptyl je rovný $\frac{\alpha}{\beta^2}$. Kontrolu prevedieme nasledujúcim výpočtom:

$$\begin{aligned} \mathbb{E}X &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\beta^{\alpha+1} \Gamma(\alpha)} \int_0^\infty y^\alpha e^{-y} dy \\ &= \frac{1}{\beta \Gamma(\alpha)} \Gamma(\alpha + 1) \\ &= \frac{\alpha}{\beta}. \end{aligned}$$

Rovnakým postupom sa dopracujeme k výsledku $\mathbb{E}X^2 = \frac{\alpha(\alpha+1)}{\beta^2}$. Odtiaľ potom dostaneme:

$$\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{\alpha(\alpha+1)}{\beta^2} - \left(\frac{\alpha}{\beta}\right)^2 = \frac{\alpha}{\beta^2}.$$

Neceločíselné momenty

Výpočet $\mathbb{E}|X|^r$, $r \geq 0$ je podobný výpočtom $\mathbb{E}X$ a $\mathbb{E}X^2$:

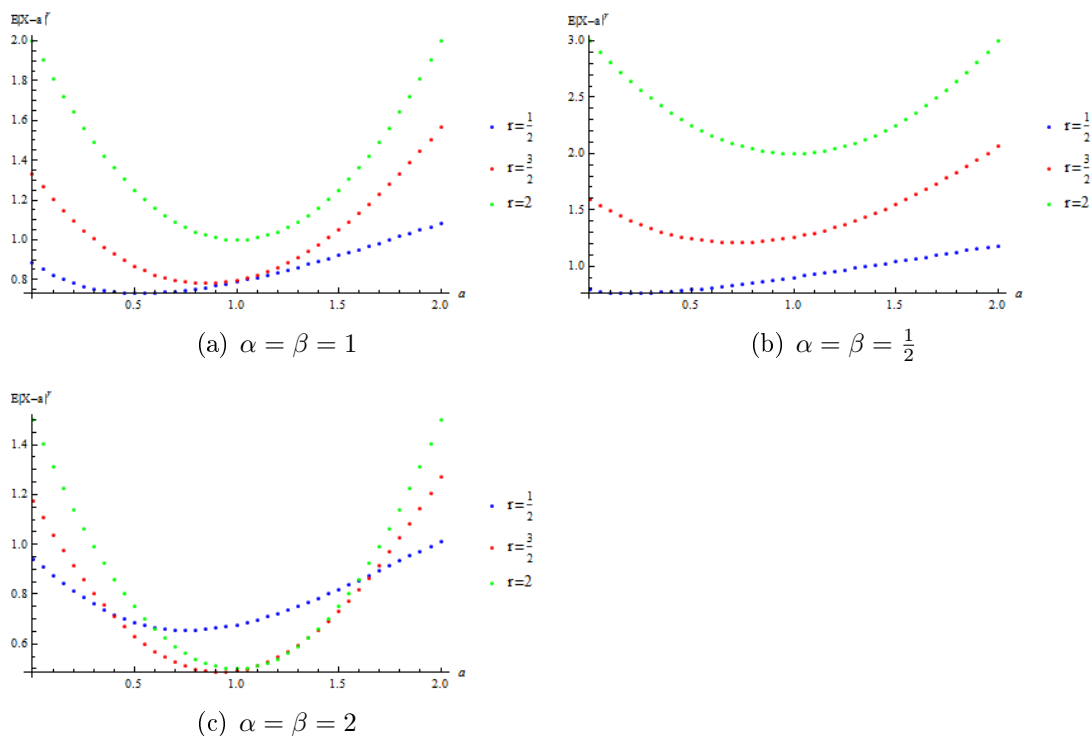
$$\begin{aligned} \mathbb{E}|X|^r &= \mathbb{E}X^r = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{r+\alpha-1} e^{-\beta x} dx \\ &= \frac{\beta^\alpha}{\beta^{r+\alpha} \Gamma(\alpha)} \int_0^\infty y^{r+\alpha-1} e^{-y} dy \\ &= \frac{1}{\beta^r \Gamma(\alpha)} \Gamma(r + \alpha). \end{aligned}$$

Z grafu na obr. 3.2(a) a 3.2(b) môžeme vidieť nárast funkčnej hodnoty $\mathbb{E}|X - a|^r$ pri zvyšovaní momentu r , na grafe na obr. 3.2(c) to už neplatí. Pre $r = 2$ má funkcia tvar paraboly s vrcholom v $\mathbb{E}X = \frac{\alpha}{\beta} = 1$ a funkčnou hodnotou v tomto bode rovnou $\text{var}X = \frac{\alpha}{\beta^2}$.

3.4 Paretovo rozdelenie

Náhodná veličina X má Paretovo rozdelenie práve vtedy, keď je definovaná hustotou

$$f(x) = \frac{\alpha \lambda^\alpha}{(\lambda + x)^{\alpha+1}}, \quad x \geq 0, \alpha > 0, \lambda > 0.$$



Obr. 3.2: Graf zobrazujúci priebeh $\mathbb{E}|X - a|^r$ v závislosti na a pre gamma rozdelenie a momenty $r_1 = \frac{1}{2}$, $r_2 = \frac{3}{2}$, $r_3 = 2$.

Na výpočet strednej hodnoty $\mathbb{E}X$, druhého momentu $\mathbb{E}X^2$, aj neceločíselných momentov $\mathbb{E}|X|^r$ využijeme substitúciu $y = \frac{x}{\lambda}$ a beta funkciu definovanú predpisom

$$B(p, q) = \int_0^\infty \frac{x^{p-1}}{(1+x)^{p+q}} dx = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

nasledovne

$$\mathbb{E}X = \alpha\lambda^\alpha \int_0^\infty \frac{x}{(\lambda+x)^{\alpha+1}} dx = \alpha\lambda \int_0^\infty \frac{y}{(1+y)^{\alpha+1}} dy = \frac{\lambda}{\alpha-1}$$

$$\mathbb{E}X^2 = \alpha\lambda^\alpha \int_0^\infty \frac{x^2}{(\lambda+x)^{\alpha+1}} dx = \alpha\lambda^2 \int_0^\infty \frac{y^2}{(1+y)^{\alpha+1}} dy = \frac{2\lambda^2}{(\alpha-1)(\alpha-2)}.$$

Odtiaľ dopočítame $\text{var}X$ podobne ako v predchádzajúcich prípadoch.

$$\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{\lambda^3}{(\lambda-1)^2(\lambda-2)}.$$

Neceločíselné momenty

Výpočet $\mathbb{E}|X|^r$, $0 \leq r < \alpha$ je podobný výpočtom $\mathbb{E}X$ a $\mathbb{E}X^2$:

$$\begin{aligned}
\mathbb{E}|X|^r &= \alpha\lambda^\alpha \int_0^\infty \frac{x^r}{(\lambda+x)^{\alpha+1}} dx \\
&= \alpha\lambda^r \int_0^\infty \frac{y^r}{(1+y)^{\alpha+1}} dy \\
&= \alpha\lambda^r B(r+1, \alpha-r).
\end{aligned}$$

3.5 Rovnomerné rozdelenie

Spojité rovnomerné rozdelenie na intervale $[0, b]$ je definované hustotou

$$f(x) = \begin{cases} 0, & x \notin [0, b] \\ \frac{1}{b}, & x \in [0, b]. \end{cases}$$

Stredná hodnota a rozptyl náhodnej veličiny $X \sim R(0, b)$ sú $\mathbb{E}X = \frac{b}{2}$ a $\text{var } X = \frac{b^2}{12}$. Výpočty boli prevedené nasledovne:

$$\begin{aligned}
\mathbb{E}X &= \frac{1}{b} \int_0^b x dx = \frac{1}{b} \cdot \frac{b^2}{2} = \frac{b}{2}, \\
\mathbb{E}X^2 &= \frac{1}{b} \int_0^b x^2 dx = \frac{1}{b} \cdot \frac{b^3}{3} = \frac{b^2}{3}, \\
\text{var } X &= \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{b^2}{12}.
\end{aligned}$$

Neceločíselné momenty

Necentrálny absolútny moment náhodnej veličiny $X \sim R(0, b)$ rádu r vypočítame jednoduchou integráciou

$$\mathbb{E}|X|^r = \frac{1}{b} \int_0^b x^r dx = \frac{1}{b} \cdot \frac{b^{r+1}}{r+1} = \frac{b^r}{r+1}.$$

Absolútny r -ty moment $\mathbb{E}|X - a|^r$ sa dá tiež vypočítať analogicky:

$$\begin{aligned}
\mathbb{E}|X - a|^r &= \frac{1}{b} \int_0^b |x - a|^r dx = \frac{1}{b} \left(\int_0^a (a - x)^r dx + \int_a^b (x - a)^r dx \right) \\
&= \frac{1}{b} \left(\int_0^a y^r dy + \int_0^{b-a} y^r dy \right) = \frac{1}{b} \left(\frac{a^r}{r+1} + \frac{(b-a)^r}{r+1} \right).
\end{aligned}$$

4. Momentová metóda

4.1 Motivácia

Motiváciou pre použitie momentovej metódy v súvislosti s počítaním pomocou neceločíselných momentov náhodných veličín nám boli články [3] a [2]. V oboch článkoch sa rozoberá problematika odhadov parametrov zo zmesi rozdelení práve pomocou neceločíselných momentov.

V článku [3] je rozoberaný prístup k počítaniu odhadov troch parametrov v zmiešanom exponenciálnom rozdelení (jeden z odhadovaných parametrov definuje váhu dvoch exponenciálnych rozdelení, ostatné dva parametre sú parametrami týchto rozdelení) pomocou momentovej metódy. Spomína sa aj teória maximálnej vierohodnosti, ktorú v práci nebudeme potrebovať. V článku sa uvádza, že použitie momentovej metódy vedie k menšiemu množstvu výpočtov ako pri používaní teórie maximálnej vierohodnosti. Momentová metóda v praxi poskytne množstvo odhadov parametrov a navyše aj tam, kde ostatné metódy zlyhávajú alebo vedú často ku komplikovaným výpočtom. V článku je zmienený prístup, ktorý využíva neceločíselné ako aj absolútne neceločíselné momenty, ktorý by mohol zvýšiť presnosť odhadov. Pomocou výpočtov je ukázaný vhodný výber neceločíselných momentov pri snahe o odhad parametrov rozdelenia. Zároveň sú porovnané odhady parametrov pomocou momentovej metódy a metódy maximálnej vierohodnosti. Autori článku uvádzajú aj najvhodnejšie kombinácie momentov pri rôznych hodnotách uvedených parametrov.

V článku [2] sa diskutuje o odhadoch parametrov θ_1, θ_2 pre náhodný výber z parametrickej rodiny $L = \{F(x; \theta) : \theta \in \Theta\}$, kde je vo vopred známom pomere daný počet m náhodných veličín z rozdelenia s parametrom θ_1 a n náhodných veličín z rozdelenia s parametrom θ_2 . Ale nevieme rozoznať, ktoré veličiny sú rozdelené podľa rozdelenia $F(x; \theta_1)$ a ktoré podľa $F(x; \theta_2)$. Článok opäť rozoberá využitie neceločíselných momentov pri odhadoch parametrov θ_1, θ_2 , avšak pri predpoklade parametrickej rodiny rozdelení. V závere sa uvádza, že použitie neceločíselných momentov zvýši efektívnosť momentovej metódy odhadovania parametrov, ak sú odhadované parametre z dvoch nezávislých a zmiešaných rozdelení. Autori sa opäť zamerali na procedúru pre odhad parametrov v exponenciálnej triede rozdelení, kde vhodnou lineárnou kombináciou neceločíselných odhadov dospeli k efektívnej momentovej metóde.

V našom prípade sme nevolili zmes exponenciálnych rozdelení. Namiesto toho sme zvolili náhodný výber z Gamma a Paretovho rozdelenia, ktoré sú definované dvoma parametrami. Obe rozdelenia sú spojité, avšak nie sú úplne totožné, nakoľko Gamma rozdelenie sa radí medzi rozdelenia s ľahkými chvostmi, zatiaľ čo Pareto rozdelenie patrí medzi tie s ťažkými chvostmi.

4.2 Princíp momentovej metódy

Momentová metóda je založená výhradne na silnom zákone veľkých čísel (SZVČ), ktorý znie nasledovne:

Tvrdenie 4.1. *Nech X_1, X_2, \dots je postupnosť nezávislých, rovnako rozdelených náhodných veličín s konečnou strednou hodnotou $\mu_X = \mathbb{E}X_1$. Potom*

$$\overline{X}_n \xrightarrow[n \rightarrow \infty]{} \mu_X \quad s.i.,$$

kde $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Princíp odhadovania parametrov pomocou momentovej metódy si ukážeme najskôr za predpokladu rozdelenia s jedným parametrom. Majme nezávislý náhodný výber X_1, X_2, \dots, X_n získaný podľa parametrickej hustoty pravdepodobnosti $f_X(x|\theta)$, kde parameter θ je neznámy.

Nech X_i sú spojité náhodné veličiny a μ_X ich stredná hodnota. Potom označme

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x|\theta) dx = k(\theta).$$

Zo zákona veľkých čísel plynie

$$\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mu_X \quad s.i. \text{ pre } n \rightarrow \infty.$$

Z toho vyplýva, že pre dostatočne veľké n môžeme aproximovať $\overline{X}_n \approx k(\theta)$ a zo vzťahu $\overline{X}_n = k(\hat{\theta})$ vypočítať odhad $\hat{\theta}$.

Ak odhadujeme d -rozmerný vektor parametrov, musíme postupovať všeobecnejšie. Nech máme postupnosť X_1, X_2, \dots, X_n nezávislých náhodných veličín z rozdelenia s hustotou $f_{X_1}(x|\theta_1, \dots, \theta_d)$, ktorá závisí na vektore parametrov $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$.

Označme

$$\mu_i = \mathbb{E}X_1^i = \int_{-\infty}^{\infty} x^i f_{X_1}(x|\boldsymbol{\theta}) dx$$

ako i -ty necentrálny moment náhodnej veličiny X_1 . Predpokladáme, že μ_1, \dots, μ_d sú konečné.

Tento moment vyjadríme pomocou vektora parametrov ako $\mu_i = k_i(\theta_1, \dots, \theta_d)$, $i = 1, \dots, d$:

$$\begin{aligned} \mu_1 &= k_1(\theta_1, \theta_1, \dots, \theta_d), \\ \mu_2 &= k_2(\theta_1, \theta_1, \dots, \theta_d), \\ &\vdots \\ \mu_d &= k_d(\theta_1, \theta_1, \dots, \theta_d). \end{aligned}$$

Na základe dát $\mathbf{X} = (x_1, \dots, x_n)$ vypočítame prvých d výberových momentov:

$$\overline{x^1} = \frac{1}{n} \sum_{i=1}^n x_i, \dots, \overline{x^d} = \frac{1}{n} \sum_{i=1}^n x_i^d.$$

Zo zákona veľkých čísel dostávame $\mu_i \approx \overline{x^i}$, $\forall i = 1, \dots, d$. Riešime sústavu

$$\begin{aligned}\overline{x^1} &= k_1(\theta_1, \theta_1, \dots, \theta_d), \\ \overline{x^2} &= k_2(\theta_1, \theta_1, \dots, \theta_d), \\ &\vdots \\ \overline{x^d} &= k_d(\theta_1, \theta_1, \dots, \theta_d).\end{aligned}$$

Odhad vektora parametrov $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$ označíme ako $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_d)$. My budeme odhadovať d -rozmerný vektor parametrov pomocou neceločíselných momentov. Musíme teda tento postup zovšeobecniť. Označme postupnosť čísel $\mathbf{r} = (r_1, \dots, r_d)$, $r_i \geq 0$, $\forall i = 1, \dots, d$. Nech

$$\mu_{r_i} = \mathbb{E}X_1^{r_i} = \int_{-\infty}^{\infty} x^{r_i} f_{X_1}(x|\theta_1, \theta_1, \dots, \theta_d) dx < \infty$$

je r_i -ty necentrálny a neceločíselný moment náhodnej veličiny X .

Tento moment vyjadríme pomocou vektora parametrov ako $\mu_{r_i} = k_i(\theta_1, \dots, \theta_d)$. Na základe dát $\mathbf{X} = (x_1, \dots, x_n)$ vypočítame d výberových momentov s mocnismi (r_1, \dots, r_d) :

$$\overline{x^{r_1}} = \frac{1}{n} \sum_{i=1}^n x_i^{r_1}, \dots, \overline{x^{r_d}} = \frac{1}{n} \sum_{i=1}^n x_i^{r_d}.$$

Zo zákona veľkých čísel dostávame $\mu_{r_i} \approx \overline{x^{r_i}}$, $\forall i = 1, \dots, d$. Riešime sústavu

$$\begin{aligned}\overline{x^{r_1}} &= k_1(\theta_1, \theta_1, \dots, \theta_d), \\ \overline{x^{r_2}} &= k_2(\theta_1, \theta_1, \dots, \theta_d), \\ &\vdots \\ \overline{x^{r_d}} &= k_d(\theta_1, \theta_1, \dots, \theta_d).\end{aligned}$$

Odhad vektora parametrov $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$ opäť označíme ako $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_d)$.

4.3 Simulačný experiment

Ako už bolo na začiatku spomenuté, budeme odhadovať parametre rozdelení Gamma a Pareto. Všetky numerické výpočty prebiehali v programe Wolfram Mathematica [5].

- Pre každé rozdelenie sme si na začiatku vygenerovali náhodný výber s rozsahom 500 zo známeho rozdelenia so známymi parametrami.
- Z vygenerovaného náhodného výberu sme vytvorili 5 podvýberov nasledovne: 1. náhodný podvýber obsahuje prvých 100 prvkov, 2. náhodný podvýber obsahuje prvých 200 prvkov pôvodného výberu až posledný náhodný podvýber je totožný s výberom vygenerovaným na začiatku. (Tento krok robíme preto, aby sme ukázali, že stredná štvorcová chyba (MSE) klesá so vzrastajúcim počtom pozorovaní.)

- Pre všetkých 5 podvýberov sme napočítali výberové momenty $\overline{x^{r_1}}$ a $\overline{x^{r_2}}$. Tie sme porovnali s teoretickými momentmi μ_{r_1} a μ_{r_2} daných rozdelení a riešili ako sústavu dvoch rovníc o dvoch neznámych parametroch.
- Keďže sme mali podmienku reálnych parametrov a niektoré sústavy boli pre neceločíselné momenty riešiteľné iba v množine komplexných čísel, pretransformovali sme úlohu riešenia sústav na minimalizačnú úlohu, ktorá je numericky lepšia.
- Minimalizovali sme funkciu

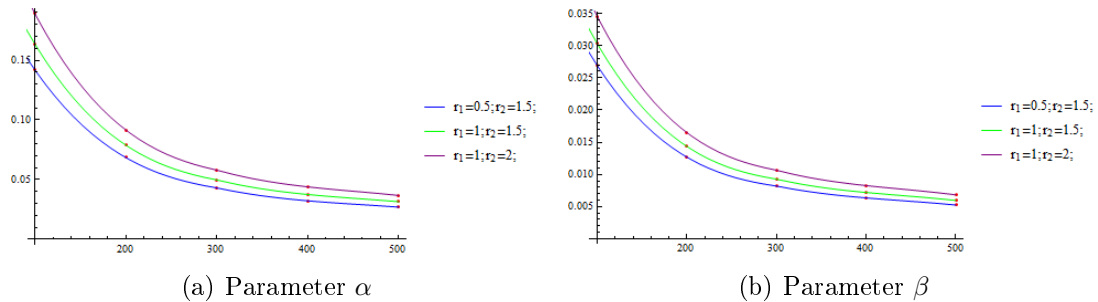
$$(\overline{x^{r_1}} - \mu_{r_1})^2 + (\overline{x^{r_2}} - \mu_{r_2})^2,$$

kde $r_1, r_2 \geq 0$ sú neceločíselné momenty. V prípade, že pôvodná úloha má reálne riešenie, jej minimalizáciou dostaneme to isté riešenie a funkcia nadobudne hodnotu 0. V opačnom prípade minimalizáciou hľadáme najbližšie korene v zmysle čo najmenšej strednej štvorcovej chyby. Navyše, minimalizačná úloha nám poskytla priestor pre prípadné obmedzenia parametrov.

Celý simulačný experiment sme zopakovali 1000krát a spočítali strednú štvorcovú chybu odhadu $\hat{\theta}$ parametra θ pre všetkých 5 podvýberov zo vzorca

$$\text{MSE} = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{\theta}_{ij} - \theta_j)^2,$$

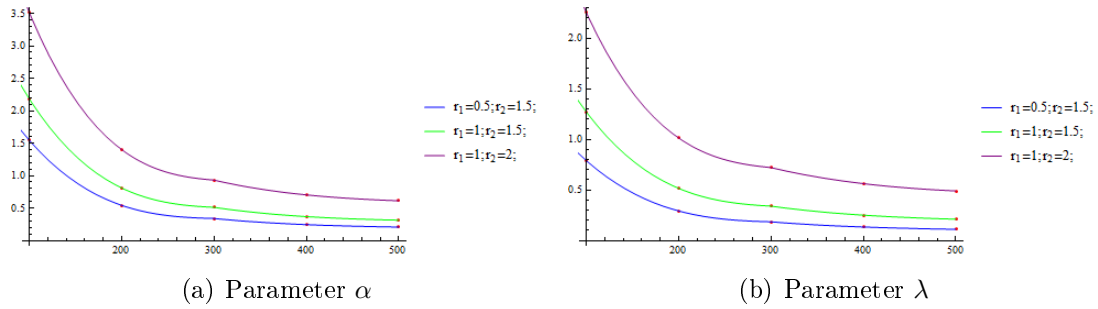
kde θ_j je j -ta zložka θ a $\hat{\theta}_{ij}$ je odhad θ_j v i -tom experimente.



Obr. 4.1: Stredná štvorcová chyba odhadov parametrov α, β Gamma rozdelenia.

Náhodné výbery boli generované z Gamma rozdelenia s parametrami $\alpha = 2$ a $\beta = 1$. V grafe na obr. 4.1(a) a 4.1(b) je znázornená stredná štvorcová chyba odhadov parametrov pre Gamma rozdelenie. Z grafu tiež môžeme vidieť, že vzhľadom na hodnoty strednej štvorcovej chyby sú odhady parametrov Gamma rozdelenia aproximované s chybou menšou ako 0,2 pre parameter $\alpha = 2, 5$ a 0,035 pre parameter $\beta = 1$. So zväčšujúcim sa počtom pozorovaní chyby klesajú. Môžeme teda povedať, že sme aplikovaným postupom dobre aproximovali parametre Gamma rozdelenia.

Náhodné výbery boli generované z Paretoho rozdelenia s parametrami $\alpha = 2$ a $\lambda = 1$. V grafe na obr. 4.2(a) a 4.2(b) je znázornená stredná štvorcová chyba odhadov parametrov pre Paretoho rozdelenie. Z tohto grafu môžeme vidieť, že



Obr. 4.2: Stredná štvorcová chyba odhadov parametrov α, λ Paretovho rozdelenia.

vzhľadom na hodnoty strednej štvorcovej chyby sú odhady parametrov Pareto rozdelenia aproximované s chybou menšou ako 3,5 pre parameter $\alpha = 2$ a 2,5 pre parameter $\lambda = 1$. Tieto chyby sú pomerne väčšie ako u Gamma rozdelenia. Môžeme argumentovať tým, že Pareto rozdelenie je rozdelenie s ťažkými chvostmi a ľahšie nadobudne extrémne hodnoty, čo má veľký vplyv na kvalitu odhadu. So zväčšujúcim sa počtom pozorovaní podobne ako u Gamma rozdelenia chyby klesajú.

Aby bol r -ty moment Paretovho rozdelenia konečný, požadujeme $r < \alpha$. V jednom prípade sme však pri voľbe $\alpha = 2$ skúmali moment $r = 2$. Výberový moment je vždy konečný, ale teoretický moment v tomto zadaní nedosahuje konečné hodnoty. Odhad momentovou metódou sme museli hľadať v množine $\alpha \in (2, \infty)$, a teda vždy došlo k nadhodnoteniu skutočného parametra $\alpha = 2$. V malom množstve prípadov sa dokonca stalo, že numerická optimalizácia dala extrémne veľký odhad parametrov, ktorý sa pohyboval rádovo až v 10^{10} . Aby touto veľkou chybou nebola ovplyvnená stredná štvorcová chyba, pri výpočte MSE sme nebrali do úvahy klasický priemer, ale useknutý o 1% extrémnych hodnôt.

V konečnom dôsledku vidíme, že v obidvoch rozdeleniach a pre obidva parametre dosahovali neceločíselné momenty lepšie výsledky v zmysle menšej kvadratickej chyby. Ich odhady boli teda najpresnejšie. Za nimi nasledovali zmiešané momenty, kde bol jeden z momentov neceločíselný a druhý celočíselný a najvyššiu MSE pozorujeme pri počítaní čisto s celočíselnými momentmi. Neceločíselné momenty sa navyše ukázali veľmi vhodné pri Pareto rozdelení, kde teoreticky nemôžeme použiť druhý moment.

Záver

Pri práci s neceločíselnými momentmi náhodných veličín vybraných diskretných a spojitých rozdelení sme prišli k záveru, že nie vždy existuje jednoduché analytické vyjadrenie a výpočty je nutné robiť numericky. V prípadoch, keď sme museli počítať numericky, sme hodnoty momentov pre jednotlivé parametry zobrazili do grafov, ktoré boli bližšie opísané v príslušných kapitolách.

V poslednej kapitole, zaoberajúcej sa momentovou metódou, sme zistili, že neceločíselné momenty sú použiteľné namiesto celočíselných momentov a dokonca dávajú lepšie odhady v zmysle strednej štvorcovej chyby. Pri rozdeleniach s ťažkými chvostmi sú teda výbornou alternatívou, pretože celočíselné momenty týchto rozdelení nemusia byť konečné dokonca ani pre $r = 1$. V našom prípade sme sa o tom presvedčili pri Paretovom rozdelení.

Zoznam použitej literatúry

- [1] DUPAČ, Václav, HUŠKOVÁ, Marie. *Pravděpodobnost a matematická statistika*. 1. vydání. Nakladatelství Karolinum, Praha, 2005. (ISBN 80-246-0009-9).
- [2] FROM, Steven G., SAXENA, K.M. Lal. *Estimating parameters from mixed samples using sample fractional moments*. Journal of statistical planning and inference 21 (1989) 231-244. North-Holland.
- [3] TALLIS, G.M., LIGHT, R. *The use of fractional moments for estimating the parameters of a mixed exponential*. Technometrics, Vol 10, No 1(Feb., 1968). The John Hopkins University.
- [4] WATKINS, Joseph C. *Introduction to the Science of Statistics*.
http://math.arizona.edu/~jwatkins/M_moments.pdf
2013.
- [5] WOLFRAM RESEARCH, Inc., *Mathematica*, Version 9.0, Champaign, IL (2012).
- [6] ZVÁRA, Karel, JOSEF, Štěpán. *Pravděpodobnost a matematická statistika*. 3. vydání. MATFYZPRESS, Praha, 2002. (ISBN 80-85863-93-6).

Zoznam obrázkov

2.1	Graf zobrazujúci priebeh $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ v závislosti na r pre binomické rozdelenie s parametrami $n = 20$ a $p \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$	7
2.2	Graf zobrazujúci priebeh $\mathbb{E} X - a ^r$ v závislosti na a pre binomické rozdelenie s parametrami $n = 20$ a $p \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$	8
2.3	Graf zobrazujúci priebeh $\frac{\mathbb{E}X^r}{(\mathbb{E}X)^r}$ v závislosti na r pre Poissonovo rozdelenie s hodnotami parametrov $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 10$	9
2.4	Ohraničenie $\mathbb{E}X^r$ zhora a zdola.	10
3.1	Graf zobrazujúci priebeh $\frac{\mathbb{E} X-a ^r}{(\mathbb{E} X-a)^r}$ v závislosti na a pre normálne rozdelenie a rôzne momenty $r_1 = \frac{1}{2}, r_2 = \frac{3}{2}, r_3 = 2$	12
3.2	Graf zobrazujúci priebeh $\mathbb{E} X - a ^r$ v závislosti na a pre gamma rozdelenie a momenty $r_1 = \frac{1}{2}, r_2 = \frac{3}{2}, r_3 = 2$	15
4.1	Stredná štvorcová chyba odhadov parametrov α, β Gamma rozdelenia.	20
4.2	Stredná štvorcová chyba odhadov parametrov α, λ Paretoho rozdelenia.	21