

**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Matěj Komprda

Zobecněné gama rozdělení

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Šárka Hudecová Ph.D.

Studijní program: Matematika

Studijní obor: Finanční matematika

Praha 2024

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

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 má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Rád bych poděkoval vedoucí své práce RNDr. Šárce Hudecové, Ph.D., za její cenné rady, ochotu a věnovaný čas během psaní bakalářské práce.

Název práce: Zobecněné gama rozdělení

Autor: Matěj Komprda

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Šárka Hudecová Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Zobecněné gama rozdělení je spojité rozdělení se třemi parametry, které nabízí flexibilní nástroj pro modelování nezáporných reálných veličin a které zahrnuje řadu jiných modelů jako speciální případ. Práce se věnuje popisu tohoto rozdělení. Jsou v ní odvozeny jeho základní vlastnosti, jako například momenty či distribuční funkce. V další části jsou odvozeny odhady parametrů zobecněného gama rozdělení metodou maximální věrohodnosti, momentovou metodou a momentovou metodou využívající vhodnou transformaci. Tyto odhady jsou porovnány v simulační studii, kde je posouzena jejich přesnost a konzistence.

Klíčová slova: Zobecněné gama rozdělení, metoda maximální věrohodnosti, momentová metoda

Title: Generalized gamma distribution

Author: Matěj Komprda

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Šárka Hudecová Ph.D., Department of Probability and Mathematical Statistics

Abstract: The generalized gamma distribution is a continuous distribution with three parameters, offering a flexible tool for modeling non-negative real-valued variables and encompassing a variety of other models as special cases. In this work, the description of this distribution is addressed. Its basic properties, such as moments or distribution functions, are derived. Estimates of the generalized gamma distribution parameters are constructed using the maximum likelihood method, the method of moments, and the method of moments employing a suitable transformation. These estimates are compared in a simulation study, where their accuracy and consistency are assessed.

Keywords: Generalized gamma distribution, maximum likelihood estimation, method of moments

Obsah

Úvod	2
1 Základní pojmy	3
1.1 Klasické gama rozdělení	3
1.2 Pojmy z pravděpodobnosti	4
2 Zobecněné gama rozdělení	6
2.1 Vlastnosti zobecněného gama rozdělení	6
2.2 Transformace zobecněného gama rozdělení	10
3 Odhady parametrů zobecněného gama rozdělení	12
3.1 Odhad metodou maximální věrohodnosti	12
3.2 Odhad momentovou metodou	14
3.2.1 Přímý odhad momentovou metodou	14
3.2.2 Odhad momentovou metodou užitím transformace	14
4 Simulační studie	19
4.1 Simulace pro $p > 0$	19
4.2 Simulace pro $p < 0$	22
4.3 Odhady pro klasické gama rozdělení	25
4.4 Shrnutí simulací	26
Závěr	27
Seznam použité literatury	28

Úvod

Čtenář této práce se již jistě setkal s mnoha různými pravděpodobnostními rozděleními, jako je například gama rozdělení, weibullovo rozdělení či exponenciální rozdělení. V této práci se zaměříme na zobecněné gama rozdělení, které výše uvedená rozdělení zahrnuje.

V praxi je často potřeba modelovat rozdělení náhodné veličiny, která nabývá pouze kladných reálných hodnot, jako například při modelování výše pojistných událostí nebo rozdělení příjmů v populaci. Modelem pro taková data může být právě zobecněné gama rozdělení, které nabízí, díky trojici parametrů, velkou flexibilitu.

V práci si nejprve připomeneme klasické gama rozdělení, které má jen dva parametry. Následně si odvodíme, jak transformací náhodné veličiny s tímto rozdělením získáme nový tříparametrický model, který nazveme zobecněným gama rozdělením. Pro zobecněné gama rozdělení uvedeme jeho základní vlastnosti, odvodíme si jeho k -té momenty nebo jeho distribuční funkci. Uvedeme si, jak lze volbou parametrů z tohoto obecného modelu získat známá rozdělení, jako je exponenciální či weibullovo rozdělení. Stěžejní část práce se následně věnuje odhadům parametrů a to metodou maximální věrohodnosti a momentovou metodou, kdy jednou je momentová metoda aplikována přímo a podruhé se využívá transformace náhodné veličiny ze zobecněného gama rozdělení. Na závěr práce jsou odhady porovnány v simulační studii, kde se diskutuje jejich konzistence a přesnost pro různé rozsahy výběrů.

Vlastním přínosem v práci je detailní odvození všech zmíněných vlastností zobecněného gama rozdělení. V práci jsme navíc odvodili jak transformovat klasické gama rozdělení na gama rozdělení zobecněné. Dalším vlastním přínosem je diskuze o existenci maxima hustoty tohoto rozdělení pro různou volbu parametrů. U odhadů parametrů jsme doplnili podrobný postup, jak se odvozují. Závěrečná simulační studie je také vlastním přínosem.

1. Základní pojmy

Na úvod práce si uvedeme několik základních pojmů a tvrzení, které budeme v pozdějších kapitolách využívat. Nejprve si uvedeme klasické gama rozdělení, jak je definováno a jaké má praktické využití. Dále si zmíníme několik pojmů z pravděpodobnosti a s nimi související tvrzení, taková, která nám v následujícím textu budou užitečná.

1.1 Klasické gama rozdělení

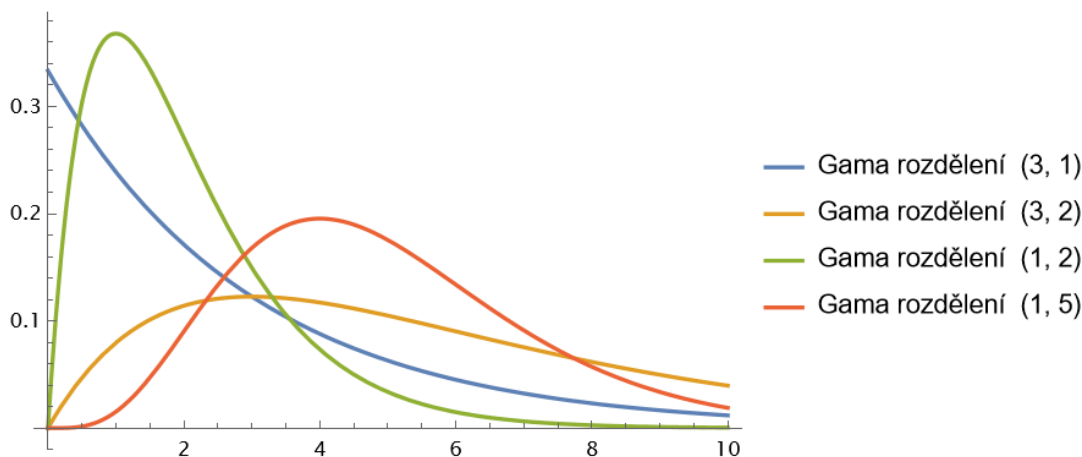
V řadě reálných aplikací je třeba modelovat rozdělení náhodné veličiny, která nabývá pouze kladných hodnot, jako je například rozdělení výše pojistných událostí, výše příjmů nebo doby do další události. Pro modelování těchto situací lze využít mnoho různých modelů. Jedním z často využívaných rozdělení je gama rozdělení s parametry $a, \nu > 0$

Definice 1. Náhodná veličina X má gama rozdělení s parametry $a, \nu > 0$, pokud je její hustota

$$f(x) = \frac{1}{a^\nu \Gamma(\nu)} x^{\nu-1} \exp\left\{-\frac{x}{a}\right\} \mathbb{I}_{(0,\infty)}(x), \quad \text{pro } x \in \mathbb{R}, \quad (1.1)$$

Kde $\Gamma(\nu)$ je hodnota gama funkce v bodě ν . Pokud je navíc parametr $a = 1$, říkáme, že X má gama rozdělení ve standardním tvaru.

Jako takové je zobecněním exponenciálního rozdělení, které z gama rozdělení získáme zvolením parametru $\nu = 1$. Na Obrázku 1.1 můžeme vidět, jakých různých tvarů může hustota gama rozdělení nabývat.



Obrázek 1.1: Hustoty gama rozdělení s parametry (a, ν)

V některých případech, může být model se dvěma parametry nedostačující. V následujícím Tvrzení 1 si uvedeme transformaci náhodné veličiny s gama rozdělením, kterou získáme novou náhodnou veličinu. Ta bude z tří parametrického rozdělení, které následně v Definici 4 nadefinujeme jako zobecněné gama rozdělení.

Tvrzení 1. Necht náhodná veličina X má gama rozdělení s parametry (a, ν) , pak transformací $Y = a^{1-\frac{1}{p}} X^{\frac{1}{p}}$, kde $p \neq 0$, dostaneme náhodnou veličinu s hustotou

$$f(y) = \frac{|p|}{a^{p\nu} \Gamma(\nu)} y^{p\nu-1} \exp\left\{-\left(\frac{y}{a}\right)^p\right\} \mathbb{I}_{(0,\infty)}(y), \quad \text{pro } y \in \mathbb{R}.$$

Důkaz. Větu dokážeme použitím věty o monotónní transformaci.

Mějme funkci t , která transformuje X , tak že $t(x) = a^{1-\frac{1}{p}}x^{\frac{1}{p}} = y$, kde $p \neq 0$. Inverzní zobrazení na $(0, \infty)$ je $t^{-1}(y) = \left[\frac{1}{a^{1-\frac{1}{p}}}y\right]^p = \frac{1}{a^{p-1}}y^p = x$. Derivace zobrazení t^{-1} podle proměnné y je $p\frac{y^{p-1}}{a^{p-1}}$. Hustotu $f_y(y)$ náhodné veličiny Y , odvodíme z hustoty $f_x(x)$ náhodné veličiny X , následovně

$$\begin{aligned} f_y(y) &= f_x\left(\frac{1}{a^{p-1}}y^p\right) \left|(t^{-1})'(y)\right| \\ &= \frac{1}{a^\nu \Gamma(\nu)} \left[\frac{y^p}{a^{p-1}}\right]^{\nu-1} \exp\left\{-\frac{y^p}{a^{p-1}a}\right\} |p| \left[\frac{y}{a}\right]^{p-1} \\ &= \frac{|p|}{a^\nu a^{p\nu-\nu} \Gamma(\nu)} \frac{a^{p-1}}{a^{p-1}} y^{p\nu-p} y^{p-1} \exp\left\{-\left(\frac{y}{a}\right)^p\right\} \\ &= \frac{|p|}{a^{p\nu} \Gamma(\nu)} y^{p\nu-1} \exp\left\{-\left(\frac{y}{a}\right)^p\right\}. \end{aligned}$$

Tím jsme dokázali, že náhodná veličina Y má příslušnou hustotu. □

1.2 Pojmy z pravděpodobnosti

Během naší práce se setkáme s centrálními momenty náhodných veličin definovanými jako

$$\mu_k(X) = \mathbf{E}[X - \mathbf{E}(X)]^k,$$

kde $k \in \mathbb{N}$ a $\mathbf{E}(X)^k$ je konečné. Centrální momenty mají některé užitečné vlastnosti, které si dokážeme v následujícím Tvzení 2.

Tvrzení 2. *Nechť X je náhodná veličina, $c \in \mathbb{R}$, $k \in \mathbb{N}$ a necht' $\mathbf{E}(X)^k$ je konečné, pak pro její k -tý centrální moment platí, že $\mu_k(X+c) = \mu_k(X)$ a $\mu_k(cX) = c^k \mu_k(X)$.*

Důkaz.

$$\begin{aligned} \mu_k(X+c) &= \mathbf{E}[(X+c) - \mathbf{E}(X+c)]^k = \mathbf{E}[(X+c) - \mathbf{E}(X) - c]^k \\ &= \mathbf{E}[X - \mathbf{E}(X)]^k = \mu_k(X), \\ \mu_k(cX) &= \mathbf{E}[cX - \mathbf{E}(cX)]^k = \mathbf{E}[c(X - \mathbf{E}(X))]^k \\ &= c^k \mathbf{E}[X - \mathbf{E}(X)]^k = c^k \mu_k(X). \end{aligned}$$

□

Při odhadování parametrů rozdělení náhodné veličiny se můžeme opírat o odhady právě jejích centrálních momentů. Ty lze obecně odhadovat následovně.

Definice 2. *Nechť $\mathbf{X} = (X_1, \dots, X_n)$ je náhodný výběr a $k \in \mathbb{N}$, pak k -tý výběrový centrální moment je definován jako*

$$\widehat{\mu}_k(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^k,$$

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

Jednou z charakteristik určující rozdělení náhodné veličiny je takzvaná momentová vytvořující funkce.

Definice 3. *Nechť X je náhodná veličina. Funkce $M_X(r) = E(e^{rX})$, se nazývá momentová vytvořující funkce náhodné veličiny X a je definovaná pro $r \in \mathbb{R}$ taková, pro která příslušná střední hodnota existuje.*

Její užitečnou vlastností je to, že z ní můžeme přímo určit k -té momenty příslušné náhodné veličiny.

Tvrzení 3. *Nechť X je náhodná veličina a necht' $M_X(r) = E(e^{rX}) < \infty$, $|r| \leq r_0$, pro nějaké $r_0 > 0$. Pak $E|X|^k < \infty$ pro všechna $k > 0$ a platí*

$$M_X^{(k)}(0) = E X^k, \quad k = 1, 2, \dots$$

Důkaz. Viz věta 2.12 Dupač a Hušková (2013).

□

2. Zobecněné gama rozdělení

Stacy a Mihram (1965) ve svém článku definují zobecněné gama rozdělení a uvádějí některé jeho základní vlastnosti. Mým přínosem v této části bude jejich podrobné odvození a doplnění rozboru vlivu volby parametrů na existenci maxima jeho hustoty.

Klasické gama rozdělení může být v některých případech málo flexibilní a proto by bylo lepší mít nějaký tří parametrický model, který by nabýval pouze kladných hodnot. V Tvzení 1 jsme si takové rozdělení transformací gama rozdělení odvodili. Tento model si nyní zadefinujeme jako zobecněné gama rozdělení.

Definice 4. *Náhodná veličina X má zobecněné gama rozdělení s parametry $a, \nu > 0, p \neq 0$, pokud je její hustota*

$$f(x) = \frac{|p|}{a^{p\nu}\Gamma(\nu)} x^{p\nu-1} \exp\left\{-\left(\frac{x}{a}\right)^p\right\} \mathbb{I}_{(0,\infty)}(x), \quad \text{pro } x \in \mathbb{R}, \quad (2.1)$$

kde $\Gamma(\nu)$ je hodnota gama funkce v bodě ν .

V celém následujícím textu budeme zápis $X \sim \text{G}\Gamma(a, \nu, p)$ chápat jako X je náhodná veličina se zobecněným gama rozdělením s parametry (a, ν, p) .

V definici 4 jsme uvedli, že parametr p může nabývat jak kladných tak záporných hodnot. Takto to ve svém článku definuje i Stacy a Mihram (1965), ovšem většina autorů uvažuje pouze kladné hodnoty parametru p , jelikož se tím mnoho výpočtů usnadňuje a v praxi je rozdělení s kladnou hodnotou parametru p stále užitečné.

2.1 Vlastnosti zobecněného gama rozdělení

Jednou z charakteristik každé náhodné veličiny je její střední hodnota a obecněji její k -tý moment. V následujícím Tvzení 4 si k -tý moment odvodíme pro náhodnou veličinu $X \sim \text{G}\Gamma(a, \nu, p)$.

Tvzení 4. *Uvažujme náhodnou veličinu $X \sim \text{G}\Gamma(a, \nu, p)$, pak její k -tý moment je*

$$E(X^k) = \begin{cases} \frac{a^k \Gamma(\frac{p\nu+k}{p})}{\Gamma(\nu)}, & \frac{k}{p} > -\nu, \\ \infty, & \text{jinak.} \end{cases} \quad (2.2)$$

Důkaz.

$$\begin{aligned} E(X^k) &= \int_{-\infty}^{\infty} x^k f_x(x) dx = \frac{|p|}{a^{p\nu}\Gamma(\nu)} \int_0^{\infty} x^{k+p\nu-1} \exp\left\{-\left(\frac{x}{a}\right)^p\right\} dx \\ &= \left| \begin{array}{l} y = \frac{x}{a} \\ dy = \frac{1}{a} dx \end{array} \right| = \frac{|p|}{a^{p\nu}\Gamma(\nu)} \int_0^{\infty} a^{k+p\nu} y^{k+p\nu-1} \exp(-y^p) dy \\ &= \frac{|p| a^k}{\Gamma(\nu)} \int_0^{\infty} \frac{p y^{p-1}}{p} y^{p(\frac{k+p\nu}{p}-1)} \exp(-y^p) dy = \left| \begin{array}{l} z = y^p \\ dz = p y^{p-1} dy \end{array} \right| \\ &= \frac{a^k}{\Gamma(\nu)} \frac{|p|}{|p|} \int_0^{\infty} z^{\frac{k+p\nu}{p}-1} \exp(-z) dz. \end{aligned}$$

Nyní je nutné rozlišit dvě různé varianty, kdy $\frac{k+p\nu}{p} > 0$ a kdy $\frac{k+p\nu}{p} \leq 0$. Pro první volbu parametrů je tento integrál konvergentní a odpovídá $\Gamma(\frac{k+p\nu}{p})$. Při druhé variantě ovšem integrál u nuly diverguje, neboť se zde chová, jako funkce x^m , kde $m \leq -1$. Odtud tedy získáváme výsledek

$$\mathbb{E}(X^k) = \begin{cases} \frac{a^k \Gamma(\frac{p\nu+k}{p})}{\Gamma(\nu)}, & \frac{k}{p} > -\nu, \\ \infty, & \text{jinak.} \end{cases}$$

□

Ze vzorce (2.2) nám zřejmě vyplývá, že náhodná veličina s tímto rozdělením má vždy konečné momenty, pokud je $p > 0$. Z (2.2) můžeme také odvodit rozptyl náhodné veličiny $X \sim \text{G}\Gamma(a, \nu, p)$. Aby existoval druhý moment budeme uvažovat parametry p a ν takové, že $2 < -p\nu$ pro $p < 0$.

$$\mathbb{E}(X)^2 = \left(\frac{a \Gamma(\frac{p\nu+1}{p})}{\Gamma(\nu)} \right)^2, \quad \mathbb{E}(X^2) = \frac{a^2 \Gamma(\frac{p\nu+2}{p})}{\Gamma(\nu)}$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = a^2 \left[\frac{\Gamma(\frac{p\nu+2}{p})}{\Gamma(\nu)} - \left(\frac{\Gamma(\frac{p\nu+1}{p})}{\Gamma(\nu)} \right)^2 \right]$$

Další významnou charakteristikou každého pravděpodobnostního rozdělení je jeho distribuční funkce.

Tvrzení 5. *Distribuční funkce náhodné veličiny $X \sim \text{G}\Gamma(a, \nu, p)$ je*

$$\begin{aligned} F(y) &= \frac{\gamma(\nu, \left(\frac{y}{a}\right)^p)}{\Gamma(\nu)}, & p > 0, \\ F(y) &= 1 - \frac{\gamma(\nu, \left(\frac{y}{a}\right)^p)}{\Gamma(\nu)}, & p < 0. \end{aligned} \tag{2.3}$$

kde $\gamma(\nu, \left(\frac{y}{a}\right)^p)$ je neúplná gama funkce v bodě ν , definovaná jako

$$\gamma(\nu, \left(\frac{y}{a}\right)^p) = \int_0^{\left(\frac{y}{a}\right)^p} t^{\nu-1} e^{-t} dt.$$

Důkaz.

$$\begin{aligned} F(y) &= \int_0^y f(x) dx = \int_0^y \frac{|p|}{a^{p\nu} \Gamma(\nu)} x^{p\nu-1} \exp\left\{-\left(\frac{x}{a}\right)^p\right\} dx = \left| \begin{array}{l} z = \frac{x}{a} \\ dz = \frac{1}{a} dx \end{array} \right| \\ &= \frac{|p|}{\Gamma(\nu)} \int_0^{\frac{y}{a}} \frac{z^{p\nu}}{z^{1-p+p}} \exp(-z^p) dz = \left| \begin{array}{l} w = z^p \\ dw = pz^{p-1} dy \end{array} \right|. \end{aligned}$$

Teď musíme rozlišit, jak se tato substituce projeví na mezích integrálu pro kladné a záporné p . Pro $p > 0$ dostaneme

$$\frac{1}{\Gamma(\nu)} \int_0^{\left(\frac{y}{a}\right)^p} w^{\frac{p\nu}{p}-1} \exp(-w) dw = \gamma(\nu, w) / \Gamma(\nu)$$

a pro parametr $p < 0$

$$\frac{1}{\Gamma(\nu)} \int_{\left(\frac{y}{a}\right)^p}^{\infty} w^{\frac{p\nu}{p}-1} \exp(-w) dw = 1 - \gamma(\nu, w)/\Gamma(\nu).$$

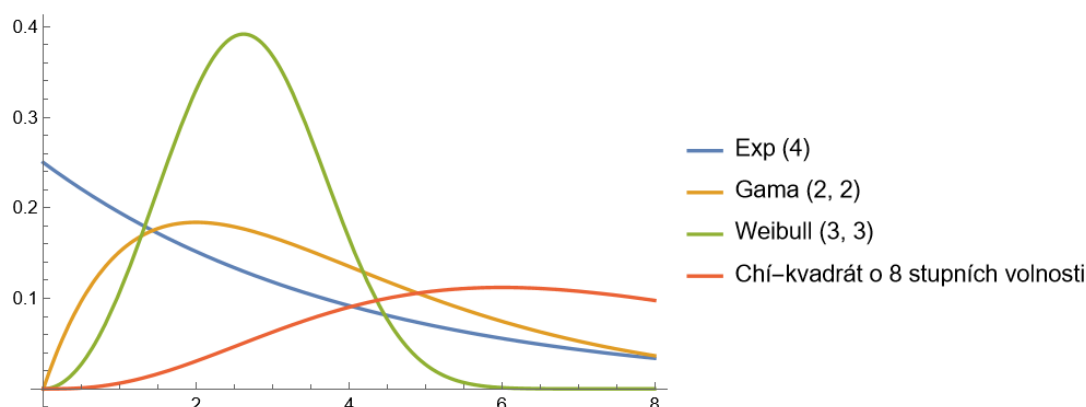
□

Zobecněné gama rozdělení je zobecněním mnoha známých a často používaných pravděpodobnostních rozdělení. Níže jsou v Tabulce 2.1 uvedena některá konkrétní rozdělení a jakou kombinací parametrů je získáme.

Gama	$G\Gamma(a, \nu, 1)$
Exponenciální	$G\Gamma(a, 1, 1)$
Weibullovo	$G\Gamma(a, 1, p)$
Chi-kvadrát o n stupních volnosti	$G\Gamma(2, n/2, 1)$

Tabulka 2.1: Známá rozdělení odvozená ze zobecněného gama rozdělení

Obrázek 2.1 ukazuje, jak mohou vypadat grafy rozdělení odvozených ze zobecněného gama rozdělení.



Obrázek 2.1: Hustoty zobecněného gama rozdělení

Limitní chování hustoty zobecněného gama rozdělení se při různé volbě hodnot parametrů liší. V nekonečnu se hustota samozřejmě vždy blíží k nule. Jak se chová limita pro x blížíící se k nule zprava v závislosti na volbě hodnot parametrů je rozebráno v Tabulce 2.2.

$$\lim_{x \rightarrow 0^+} f(x) = \frac{|p|}{a^{p\nu} \Gamma(\nu)} \lim_{x \rightarrow 0^+} \left[x^{p\nu-1} \exp \left\{ - \left(\frac{x}{a} \right)^p \right\} \right].$$

Parametr $p\nu$	$\lim_{x \rightarrow 0^+} f(x)$
$p\nu > 1$	0
$p\nu = 1$	$p/[\Gamma(\nu) a]$
$0 < p\nu < 1$	∞
$p\nu < 0$	0

Tabulka 2.2: Limity pro různé zvolené parametrizace

Jak lze nahlédnout z Obrázku 2.1 a také z naší Tabulky 2.2 pro různé hodnoty parametrů má zobecněné gama rozdělení jiný tvar a také ne vždy má funkce hustoty maximum, tudíž není nutně unimodální. Z Tabulky 2.2 snadno nahlédneme, že pro parametry p, ν takové, že $0 < p\nu < 1$, maximum funkce neexistuje neboť funkce (2.1) u nuly zprava diverguje.

Pro nalezení argumentu maxima funkce je nutné nejprve najít stacionární body. Ty můžeme určit jako kořeny derivace vyšetřované funkce. Pro hustotu zobecněného gama rozdělení je výpočet kořenů její derivace komplikovaný. Můžeme ovšem využít toho, že pokud bychom hledali stacionární body logaritmu kladné části hustoty, naši bychom stejné body.

Poznámka. V předchozí úvaze využíváme toho, že hustota zobecněného gama rozdělení $f(x)$ je pro $x > 0$ kladná, tudíž funkce $\ln f(x)$ na tomto intervalu existuje a navíc platí

$$\{x > 0 : f'(x) = 0\} = \left\{x > 0 : \frac{f'(x)}{f(x)} = 0\right\} = \left\{x > 0 : \frac{\partial}{\partial x} \ln(f(x)) = 0\right\}.$$

Jelikož je naše hustota pro $x \leq 0$ nulová, nemusíme tyto body při hledání argumentu maxima vůbec uvažovat.

Nejprve tedy vypočteme derivaci logaritmu hustoty (2.1), kterou následně položíme rovnou nule.

$$\begin{aligned} \frac{\partial}{\partial x} \ln(f(x)) &= \frac{\partial}{\partial x} \left[\ln \left(\frac{|p|}{a^{p\nu} \Gamma(\nu)} \right) + (p\nu - 1) \ln(x) - \left(\frac{x}{a} \right)^p \right] = \frac{p\nu - 1}{x} - \frac{p}{a^p} x^{p-1}, \\ \frac{p\nu - 1}{x} - \frac{p}{a^p} x^{p-1} &= 0. \end{aligned} \quad (2.4)$$

Z výše uvedených důvodů budeme uvažovat pouze parametry $p\nu \geq 1 \vee p\nu < 0$. Řešení rovnice (2.4) se liší pro dvě varianty parametrů $p\nu$. Nejprve pro $p\nu = 1$ platí

$$\frac{p\nu - 1}{x} - \frac{p}{a^p} x^{p-1} = -\frac{p}{a^p} x^{p-1} = 0 \Leftrightarrow x = 0.$$

Nula ovšem nespadá mezi přípustná x , tudíž hustota (2.1) nemá na intervalu $(0, \infty)$ žádné stacionární body, a dle limit z Tabulky 2.2 lze vyvodit, že funkce při této volbě parametrů je na daném intervalu ryze klesající a svého maxima nabývá v bodě nula.

Nyní rovnici (2.4) vyřešíme, pro volbu parametrů $p\nu > 1 \vee p\nu < 0$.

$$\frac{p\nu - 1}{x} - \frac{p}{a^p} x^{p-1} = 0 \Leftrightarrow x = a \left(\nu - \frac{1}{p} \right)^{\frac{1}{p}}.$$

Protože víme, že limity hustoty zobecněného gama rozdělení v nule zprava a v nekonečnu pro parametry $p\nu > 1 \vee p\nu < 0$ jsou nulové, a také platí, že každá hustota nabývá pouze nezáporných hodnot, pak existence právě jednoho stacionárního bodu implikuje, že tento bod je zároveň argumentem maxima této funkce. Tudíž hustota zobecněného gama rozdělení pro parametry splňující $p\nu > 1 \vee p\nu < 0$ nabývá svého maxima právě v bodě $x = a \left(\nu - \frac{1}{p} \right)^{\frac{1}{p}}$.

2.2 Transformace zobecněného gama rozdělení

Při práci s náhodnými veličinami se často využívá nějaké její transformace, která nám zaručí lepší vlastnosti a následně nabízí lepší uplatnění v praxi. V následující větě si odvodíme, jak se základní transformace náhodné veličiny $X \sim \text{GF}(a, \nu, p)$ projeví na parametrech našeho rozdělení.

Věta 6. *Nechť $X \sim \text{GF}(a, \nu, p)$, pak platí následující vztahy pro rozdělení náhodných veličin daných transformací*

$$kX \sim \text{GF}(ka, \nu, p), \quad \text{pro } k > 0, \quad (2.5)$$

$$X^m \sim \text{GF}\left(a^m, \nu, \frac{p}{m}\right), \quad \text{pro } m \neq 0. \quad (2.6)$$

Důkaz. Větu dokážeme použitím věty o monotónní transformaci.

Mějme funkci t , která transformuje náhodnou veličinu X , tak že $t(x) = kx = y$, kde $k > 0$ a $y \geq 0$. Inverzní zobrazení na $(0, \infty)$ je $t^{-1}(y) = \frac{1}{k}y = x$. Derivace zobrazení t^{-1} podle proměnné y je $\frac{1}{k}$. Hustotu $f_y(y)$ náhodné veličiny Y , odvodíme z hustoty $f_x(x)$ náhodné veličiny X , následovně

$$\begin{aligned} f_y(y) &= f_x(y/k) |(t^{-1})'(y)| = \frac{|p|}{a^{\nu} \Gamma(\nu)} (y/k)^{\nu-1} \exp\left\{-\left(\frac{y}{ka}\right)^p\right\} \left|\frac{1}{k}\right| \\ &= \frac{|p|}{(ka)^{\nu} \Gamma(\nu)} y^{\nu-1} \exp\left\{-\left(\frac{y}{ka}\right)^p\right\}. \end{aligned}$$

Tím jsme dokázali, že náhodná veličina $kX \sim \text{GF}(ka, \nu, p)$.

Nyní uvažujme novou funkci t , kterou transformujeme náhodnou veličinu X tak, že $t(x) = x^m = y$, kde $m \neq 0$ a $y \geq 0$. Inverzní zobrazení na $(0, \infty)$ je $t^{-1}(y) = y^{\frac{1}{m}} = x$ pro $m \neq 0$. Derivace zobrazení t^{-1} podle proměnné y je $\frac{1}{m}y^{\frac{1}{m}-1}$. Hustotu $f_y(y)$ náhodné veličiny Y , odvodíme z hustoty $f_x(x)$ náhodné veličiny X , následovně

$$\begin{aligned} f_y(y) &= f_x(y^{\frac{1}{m}}) |(t^{-1})'(y)| = \frac{|p|}{a^{\nu} \Gamma(\nu)} y^{\frac{1}{m}\nu - \frac{1}{m}} \exp\left\{-\left(\frac{y^{\frac{1}{m}}}{a}\right)^p\right\} \left|\frac{1}{m}y^{\frac{1}{m}-1}\right| \\ &= \frac{\left|\frac{p}{m}\right|}{(a)^{\nu} \Gamma(\nu)} y^{\frac{1}{m}\nu-1} \exp\left\{-\left(\frac{y}{a^m}\right)^{\frac{p}{m}}\right\}. \end{aligned}$$

Tedy dostaneme náhodnou veličinu $X^m \sim \text{GF}\left(a^m, \nu, \frac{p}{m}\right)$.

□

Na základě předchozí Věty 6 můžeme snadno odvodit, jaký je vztah mezi zobecněným gama rozdělením a gama rozdělením ve standardním tvaru (tedy s parametrem $a = 1$).

Tvrzení 7. *Nechť $X \sim \text{GF}(a, \nu, p)$, pak náhodná veličina $W = \left(\frac{X}{a}\right)^p$ má gama rozdělení s parametry $(1, \nu)$.*

Důkaz. Tvrzení vyplývá z Věty 6, nejprve dle výrazu (2.5), kde za k dosadíme $\frac{1}{a}$ dostaneme transformací náhodné veličiny X náhodnou veličinu $\frac{X}{a} \sim \text{GF}(1, \nu, p)$. Tu následně transformujeme dle vzorce (2.6), kde položíme $m = p$. Tudíž dostáváme, že

$\left(\frac{X}{a}\right)^p \sim \text{G}\Gamma(1, \nu, 1)$. Hustotu této náhodné veličiny vyjádříme dle Definice 4 jako

$$f(w) = \frac{1}{\Gamma(\nu)} w^{\nu-1} \exp(-w) \mathbb{I}_{(0,\infty)}(w), \quad \text{pro } w \in \mathbb{R}.$$

Tato hustota odpovídá hustotě gama rozdělení s parametry $(1, \nu)$ jako jsme ji popsali v (1.1).

□

3. Odhady parametrů zobecněného gama rozdělení

V této kapitole si odvodíme, jak z náhodného výběru odhadneme parametry zobecněného gama rozdělení. Zaměříme se na odhad metodou maximální věrohodnosti a momentovou metodou. Čerpat v této části budeme z Johnson a kol. (1994) a Stacy a Mihram (1965), kde jsou odhady metodou maximální věrohodnosti a momentovou metodou uvedeny. Vlastním přínosem je pak podrobné odvození těchto odhadů a případné zobecnění pro parametr $p < 0$, doplnění odlišného způsobu odhadu momentovou metodou a opravení nepřesnosti, která se v několika zdrojích objevila.

Pro tuto kapitolu necht' $\mathbf{X} = (X_1, \dots, X_n)$ je náhodný výběr z rozdělení $\Gamma(a, \nu, p)$. Vektor neznámých parametrů $(a, \nu, p)^T$ označíme $\boldsymbol{\theta}$.

3.1 Odhad metodou maximální věrohodnosti

Jedním ze základních způsobů odhadů parametrů rozdělení náhodné veličiny je odhad metodou maximální věrohodnosti, při kterém se snažíme maximalizovat věrohodnostní funkci. Tou je sdružená hustota náhodného výběru. V následující části si tedy postupně odvodíme věrohodnostní funkci $L(\boldsymbol{\theta})$, logaritmickou věrohodnost $\ell(\boldsymbol{\theta})$ a skórovou statistiku $\mathbf{U}_n(\boldsymbol{\theta}|\mathbf{X})$, ze které následně položením rovno nule získáme tři rovnice o třech neznámých, jejichž řešením budou námi hledané odhady.

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^n f(X_i) = \left(\frac{|p|}{a^{\nu p} \Gamma(\nu)} \right)^n \prod_{i=1}^n X_i^{p\nu-1} \exp \left\{ - \left(\frac{X_i}{a} \right)^p \right\}, \\ \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n \log \{ f(X_i) \} \\ &= n \log |p| - n \log [\Gamma(\nu)] - n\nu p \log(a) + \sum_{i=1}^n \left[(p\nu - 1) \log(X_i) - \left(\frac{X_i}{a} \right)^p \right]. \end{aligned}$$

Nyní si vyjádříme skórové statistiky pro jednotlivé parametry.

$$U_n(a|\mathbf{X}) = \frac{\partial}{\partial a} \ell(\boldsymbol{\theta}) = -\frac{n\nu p}{a} + a^{-p-1} \sum_{i=1}^n X_i^p, \quad (3.1)$$

$$U_n(\nu|\mathbf{X}) = \frac{\partial}{\partial \nu} \ell(\boldsymbol{\theta}) = -n \frac{\Gamma'(\nu)}{\Gamma(\nu)} - np \log(a) + \sum_{i=1}^n p \log(X_i), \quad (3.2)$$

$$\begin{aligned} U_n(p|\mathbf{X}) &= \frac{\partial}{\partial p} \ell(\boldsymbol{\theta}) = \operatorname{sgn}(p) \frac{n}{|p|} - n\nu \log(a) + \sum_{i=1}^n \left[\nu \log(X_i) - \left(\frac{X_i}{a} \right)^p \log \left(\frac{X_i}{a} \right) \right] \\ &= \operatorname{sgn}(p) \frac{n}{|p|} + \nu \sum_{i=1}^n \log \left(\frac{X_i}{a} \right) - \sum_{i=1}^n \left(\frac{X_i}{a} \right)^p \log \left(\frac{X_i}{a} \right). \end{aligned} \quad (3.3)$$

Hledané odhady označíme \hat{a} , $\hat{\nu}$ a \hat{p} a najdeme je položením skórových statistik rovno nule. Nejprve si vyjádříme odhad parametru a ze vzorce (3.1).

$$\begin{aligned} 0 &= -\frac{n\hat{\nu}\hat{p}}{\hat{a}} + \hat{p}\hat{a}^{-\hat{p}-1} \sum_{i=1}^n X_i^{\hat{p}}, \\ n\hat{\nu} &= \hat{a}^{-\hat{p}} \sum_{i=1}^n X_i^{\hat{p}}, \\ \hat{a} &= \left[\sum_{i=1}^n \frac{X_i^{\hat{p}}}{n\hat{\nu}} \right]^{\frac{1}{\hat{p}}}. \end{aligned} \quad (3.4)$$

Nyní skórovou statistiku (3.3) položíme rovnou nule a navíc nahradíme parametr a odhadem \hat{a} ze vzorce (3.4). Z této rovnice následně vyjádříme odhad parametru ν .

$$\begin{aligned} -\operatorname{sgn}(\hat{p}) \frac{n}{|\hat{p}|} &= -\frac{n\hat{\nu}}{\hat{p}} \log \left(\sum_{i=1}^n \frac{X_i^{\hat{p}}}{n\hat{\nu}} \right) + \hat{\nu} \sum_{i=1}^n \log(X_i), \\ &\quad - n\hat{\nu} \left(\sum_{i=1}^n X_i^{\hat{p}} \right)^{-1} \sum_{i=1}^n X_i^{\hat{p}} \left[\log(X_i) - \frac{1}{\hat{p}} \log \left(\sum_{i=1}^n \frac{X_i^{\hat{p}}}{n\hat{\nu}} \right) \right], \\ -\operatorname{sgn}(\hat{p}) \frac{n}{|\hat{p}|} &= \hat{\nu} \sum_{i=1}^n \log(X_i) - n\hat{\nu} \frac{\sum_{i=1}^n X_i^{\hat{p}} \log(X_i)}{\sum_{i=1}^n X_i^{\hat{p}}}, \\ \hat{\nu} &= \left\{ \operatorname{sgn}(\hat{p}) |\hat{p}| \left[\frac{\sum_{i=1}^n X_i^{\hat{p}} \log(X_i)}{\sum_{i=1}^n X_i^{\hat{p}}} - \sum_{i=1}^n \frac{\log(X_i)}{n} \right] \right\}^{-1}. \end{aligned} \quad (3.5)$$

Tento odhad se ovšem liší ve znaménku od výsledku uvedeném v disertační práci Hager (1969) a v knize Johnson a kol. (1994). Následná simulace v kapitole 4 ukazuje, že námi odvozený odhad (3.5) je správný.

Na závěr i skórovou statistiku (3.2) položíme rovnou nule a za parametr a dosadíme odhad ze vzorce (3.4), stejně bychom postupovali i s parametrem ν , avšak pro přehlednost nebudeme do rovnice dosazovat celý vzorec (3.5), ale budeme psát pouze $\hat{\nu}$. Dostaneme

$$\begin{aligned} 0 &= -n \frac{\Gamma'(\hat{\nu})}{\Gamma(\hat{\nu})} - n\hat{p} \log(\hat{a}) + \sum_{i=1}^n \hat{p} \log(X_i), \\ 0 &= -n \frac{\Gamma'(\hat{\nu})}{\Gamma(\hat{\nu})} - n \log \left(\sum_{i=1}^n X_i^{\hat{p}} \right) + n \log(n\hat{\nu}) + \hat{p} \sum_{i=1}^n \log(X_i), \\ 0 &= -\frac{\Gamma'(\hat{\nu})}{\Gamma(\hat{\nu})} - \log \left(\sum_{i=1}^n X_i^{\hat{p}} \right) + \log(n\hat{\nu}) + \frac{\hat{p}}{n} \sum_{i=1}^n \log(X_i), \\ H(\hat{p}) &= -\frac{\Gamma'(\hat{\nu})}{\Gamma(\hat{\nu})} - \log \left(\sum_{i=1}^n X_i^{\hat{p}} \right) + \log(n\hat{\nu}) + \frac{\hat{p}}{n} \sum_{i=1}^n \log(X_i). \end{aligned} \quad (3.6)$$

Nalezením kořenů funkce $H(\hat{p})$ najdeme odhad parametru p . V Johnson a kol. (1994) je diskutováno, že funkce $H(\hat{p})$ nemusí mít vždy jen jeden kořen, v některých případech má $H(\hat{p})$ více kořenů a někdy žádné.

3.2 Odhad momentovou metodou

Druhou často využívanou metodou odhadu parametrů rozdělení náhodné veličiny je momentová metoda. Ta využívá toho, že máme k dispozici konzistentní odhady momentů, a že momenty rozdělení náhodné veličiny obvykle umíme vyjádřit, jako funkce neznámých parametrů.

3.2.1 Přímý odhad momentovou metodou

Abychom mohli pracovat s prvními třemi momenty, musí náš náhodný výběr být dle Tvzení 4 z rozdělení s parametry, které splňují $3 < -p\nu$ v případě, že $p < 0$. Z Tvzení 4 víme, jak vypadá k -tý moment náhodné veličiny X_i . Dosazením do vzorce (2.2) snadno odvodíme první tři momenty naší náhodné veličiny X_i .

$$\mathbb{E}(X_i) = \frac{a \Gamma(\frac{p\nu+1}{p})}{\Gamma(\nu)}, \quad \mathbb{E}(X_i^2) = \frac{a^2 \Gamma(\frac{p\nu+2}{p})}{\Gamma(\nu)}, \quad \mathbb{E}(X_i^3) = \frac{a^3 \Gamma(\frac{p\nu+3}{p})}{\Gamma(\nu)}.$$

Necentrální momenty umíme odhadnout empiricky z náhodného výběru jako

$$\hat{\mu}'_k = \frac{1}{n} \sum_{i=1}^n X_i^k. \quad (3.7)$$

Dle (3.7) můžeme nyní odhadnout výše uvedené první tři necentrální momenty náhodné veličiny X_i . Jelikož se jedná o odhady je nutné nahradit parametry jejich odhady \tilde{a} , $\tilde{\nu}$, \tilde{p} .

$$\begin{aligned} \frac{\tilde{a} \Gamma(\frac{\tilde{p}\tilde{\nu}+1}{\tilde{p}})}{\Gamma(\tilde{\nu})} &= \frac{1}{n} \sum_{i=1}^n X_i, \\ \frac{\tilde{a}^2 \Gamma(\frac{\tilde{p}\tilde{\nu}+2}{\tilde{p}})}{\Gamma(\tilde{\nu})} &= \frac{1}{n} \sum_{i=1}^n X_i^2, \\ \frac{\tilde{a}^3 \Gamma(\frac{\tilde{p}\tilde{\nu}+3}{\tilde{p}})}{\Gamma(\tilde{\nu})} &= \frac{1}{n} \sum_{i=1}^n X_i^3. \end{aligned} \quad (3.8)$$

Řešením těchto tří rovnic (3.8) o třech neznámých \tilde{a} , $\tilde{\nu}$, \tilde{p} , které je nutné hledat numericky, najdeme odhady našich parametrů. Tyto odhady jsou na základě věty o spojitě transformaci konzistentní.

3.2.2 Odhad momentovou metodou užitím transformace

Stacy a Mihram (1965) popisují výhodnější způsob, jak odhadnout parametry zobecněného gama rozdělení momentovou metodou. Budeme pracovat s logaritmem naší náhodné veličiny X a odvodíme si centrální momenty transformované náhodné veličiny, díky kterým následně odhadneme neznámé parametry.

Mějme náhodnou veličinu $W = \left(\frac{X}{a}\right)^p$, z Tvzení 7 víme, že W má gama rozdělení s parametry $(1, \nu)$ a hustotu

$$f_w(w) = \frac{1}{\Gamma(\nu)} w^{\nu-1} \exp(-w) \mathbb{I}_{(0,\infty)}(w), \quad \text{pro } w \in \mathbb{R}.$$

Tvrzení 8. *Nechť náhodná veličina W má gama rozdělení s parametry $(1, \nu)$. Pak náhodná veličina $Z = \log(W)$ má hustotu*

$$f_z(z) = \frac{\exp\{z\nu - e^z\}}{\Gamma(\nu)} \mathbb{I}_{(0,\infty)}(z), \quad \text{pro } z \in \mathbb{R}.$$

Důkaz. Tvrzení dokážeme užitím věty o monotónní transformaci.

Mějme funkci t , která transformuje náhodnou veličinu W , tak že $t(w) = \log(w) = z$, kde $w \geq 0$. Inverzní zobrazení na $(0, \infty)$ je $t^{-1}: t^{-1}(z) = e^z = w$. Derivace zobrazení t^{-1} podle proměnné z je e^z a interval $(0, \infty)$ se zobrazením t^{-1} transformuje na interval $(-\infty, \infty)$. Hustotu $f_z(z)$ náhodné veličiny Z odvodíme z hustoty náhodné veličiny W následovně:

$$f_z(z) = \frac{1}{\Gamma(\nu)} e^z e^{z(\nu-1)} e^{-e^z} = \frac{\exp\{z\nu - e^z\}}{\Gamma(\nu)}, \quad \text{pro } z \in \mathbb{R}.$$

□

Pro tvorbu odhadu momentovou metodou bude potřeba určit, jak vypadají momenty odvozené náhodné veličiny Z .

Tvrzení 9. *Nechť náhodná veličina Z má hustotu z Tvrzení 8 a nechť $k \in \mathbb{N}$, pak její k -té momenty odpovídají*

$$\mathbf{E}(Z^k) = \frac{\Gamma^{(k)}(\nu)}{\Gamma(\nu)}, \quad (3.9)$$

a pro její druhý a třetí centrální moment platí, že

$$\mu_m(Z) = \frac{\partial^m}{\partial \nu^m} \log(\Gamma(\nu)), \quad \text{kde } m \in \{2, 3\}. \quad (3.10)$$

Důkaz. Pro nalezení necentrálních momentů náhodné veličiny Z využijeme její momentovou vytvořující funkci (Definice 3), pro kterou platí, že její k -tá derivace v bodě nula odpovídá k -tému necentrálnímu momentu příslušné náhodné veličiny viz Tvrzení 3.

$$\begin{aligned} M_Z(r) &= \mathbf{E}(e^{Zr}) = \int_{-\infty}^{\infty} e^{zr} \frac{\exp\{z\nu - e^z\}}{\Gamma(\nu)} dz = \left| \begin{array}{l} t = e^z \\ dt = e^z dz \end{array} \right| \\ &= \frac{1}{\Gamma(\nu)} \int_0^{\infty} t^r t^\nu e^{-t} \frac{1}{t} dt = \frac{1}{\Gamma(\nu)} \int_0^{\infty} t^{r+\nu-1} e^{-t} dt = \frac{\Gamma(\nu+r)}{\Gamma(\nu)}, \quad \text{pro } r > -\nu. \end{aligned}$$

Jelikož je momentová vytvořující funkce $M_Z(r)$ definována na intervalu $(-\nu, \infty)$ a $\nu > 0$, je definována i v bodě $r = 0$ a navíc platí

$$\frac{1}{\Gamma(\nu)} \frac{\partial}{\partial r} \Gamma(\nu+r) = \frac{1}{\Gamma(\nu)} \frac{\partial}{\partial r} \int_0^{\infty} x^{(\nu+r)-1} e^{-x} dx = \frac{1}{\Gamma(\nu)} \frac{\partial}{\partial \nu} \Gamma(\nu+r).$$

Tedy každá derivace $\Gamma(\nu+r)$ je v bodě $r = 0$ stejná, ať už je derivace podle proměnné r nebo ν . Proto můžeme psát, že k -tý necentrální moment náhodné veličiny Z je

$$\mathbf{E}(Z^k) = \frac{\Gamma^{(k)}(\nu)}{\Gamma(\nu)}.$$

Dosažením jedničky za k dostaneme střední hodnotu náhodné veličiny Z .

$$\mathbf{E}(Z) = \frac{\Gamma'(\nu)}{\Gamma(\nu)}. \quad (3.11)$$

Což odpovídá derivaci funkce $\log[\Gamma(\nu)]$. Při studování prvních dvou derivací funkce (3.11) zjistíme, že odpovídají druhému a třetímu centrálnímu momentu naší náhodné veličiny Z .

$$\begin{aligned}
\frac{\partial}{\partial \nu} \frac{\Gamma'(\nu)}{\Gamma(\nu)} &= \frac{\Gamma''(\nu)\Gamma(\nu)}{\Gamma^2(\nu)} - \left(\frac{\Gamma'(\nu)}{\Gamma(\nu)} \right)^2 = \mathbf{E}(Z^2) - \mathbf{E}^2(Z) \\
&= \mathbf{E}(Z - \mathbf{E}(Z))^2, \\
\frac{\partial^2}{\partial \nu^2} \frac{\Gamma'(\nu)}{\Gamma(\nu)} &= \left(\frac{\Gamma''(\nu)\Gamma(\nu)}{\Gamma^2(\nu)} - \left(\frac{\Gamma'(\nu)}{\Gamma(\nu)} \right)^2 \right)' \\
&= \frac{\Gamma'''(\nu)}{\Gamma(\nu)} - \frac{\Gamma''(\nu)}{\Gamma(\nu)} \frac{\Gamma'(\nu)}{\Gamma(\nu)} - 2 \frac{\Gamma'(\nu)}{\Gamma(\nu)} \left[\frac{\Gamma''(\nu)}{\Gamma(\nu)} - \left(\frac{\Gamma'(\nu)}{\Gamma(\nu)} \right)^2 \right] \\
&= \mathbf{E}(Z^3) - \mathbf{E}(Z^2)\mathbf{E}(Z) - 2\mathbf{E}(Z)[\mathbf{E}(Z^2) - \mathbf{E}^2(Z)] \\
&= (Z^3) - 3\mathbf{E}(Z^2)\mathbf{E}(Z) - 2\mathbf{E}^3(Z) = \mathbf{E}(Z - \mathbf{E}(Z))^3.
\end{aligned} \tag{3.12}$$

□

Poznámka. Funkce (3.11) se nazývá digama funkce. Pro následující část textu budeme digama funkci v bodě x značit $\psi(x)$

$$\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}, \quad x > 0.$$

Tedy druhý a třetí centrální moment náhodné veličiny Z splňuje, že

$$\mu_2(Z) = \psi'(\nu), \quad \mu_3(Z) = \psi''(\nu).$$

Nadefinujme si náhodnou veličinu Y takovou, že $Y = \log(X)$, pak k -tý centrální moment náhodné veličiny Z lze vyjádřit pomocí k -tého centrálního momentu náhodné veličiny Y užitím Tvzení 2 jako

$$\mu_k(Z) = \mu_k \left[\log \left(\frac{X}{a} \right)^p \right] = \mu_k(pY - p \log a) = p^k \mu_k(Y). \tag{3.13}$$

Nyní můžeme užitím (3.12) a (3.13) konkrétně vyjádřit první tři momenty náhodné veličiny Z , tak abychom pomocí nich mohli odhadovat parametry zobecněného gama rozdělení, ze kterého pochází náš náhodný výběr X_1, \dots, X_n . Ten můžeme zlogaritmováním hodnot transformovat na náhodný výběr Y_1, \dots, Y_n , pocházející z rozdělení určeného náhodnou veličinou Y .

$$\psi(\nu) = p \mathbf{E}[Y_i - \log(a)] = p \mathbf{E}(Y_i) - p \log(a), \tag{3.14}$$

$$\psi'(\nu) = p^2 \mu_2(Y_i), \tag{3.15}$$

$$\psi''(\nu) = p^3 \mu_3(Y_i). \tag{3.16}$$

Z výše uvedených rovnic si postupně vyjádříme jednotlivé parametry. Nejprve z rovnice (3.16) vyjádříme parametr p pomocí proměnné ν .

$$p = \left[\frac{\psi''(\nu)}{\mu_3(Y_i)} \right]^{\frac{1}{3}}. \tag{3.17}$$

Nyní z rovnice (3.14) vyjádříme parametr a opět v závislosti na ν a dosadíme za p výraz (3.17).

$$\begin{aligned}\psi(\nu) &= p \mathbf{E}(Y_i) - p \log(a) \\ \log(a) &= \mathbf{E}(Y_i) - \frac{\psi(\nu)}{p} \\ a &= \exp \left\{ \mathbf{E}(Y_i) - \left[\frac{\mu_3(Y_i)}{\psi''(\nu)} \right]^{\frac{1}{3}} \psi(\nu) \right\}.\end{aligned}\quad (3.18)$$

I do rovnice (3.15) dosadíme za p výraz (3.17) a na jednu stranu rovnice převedeme funkce závislé na parametru ν .

$$\begin{aligned}\psi'(\nu) &= \left[\frac{\psi''(\nu)}{\mu_3(Y_i)} \right]^{\frac{2}{3}} \mu_2(Y_i) \\ \frac{\psi'(\nu)}{[\psi''(\nu)]^{\frac{2}{3}}} &= \frac{\mu_2(Y_i)}{\mu_3(Y_i)^{\frac{2}{3}}}.\end{aligned}\quad (3.19)$$

V rovnicích (3.17), (3.18), (3.19) odhadneme centrální momenty náhodné veličiny Y_i výběrovými centrálními momenty dle Definice 2.

$$\begin{aligned}\widehat{\mu}_1(Y_i) &= \bar{Y}_{in} = \frac{1}{n} \sum_{i=1}^n Y_i, \\ \widehat{\mu}_2(Y_i) &= S_y^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2, \\ \widehat{\mu}_3(Y_i) &= \widehat{\mu}_3 = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^3.\end{aligned}\quad (3.20)$$

Nyní známe vše potřebné pro vyjádření odhadů \ddot{a} , $\ddot{\nu}$, \ddot{p} parametrů a , ν , p . Stačí nám dosadit odhady (3.20) do výše odvozených rovnic (3.17), (3.18), (3.19).

$$\frac{\psi'(\ddot{\nu})}{[\psi''(\ddot{\nu})]^{\frac{2}{3}}} = \frac{S_Y^2}{\widehat{\mu}_3^{\frac{2}{3}}}, \quad (3.21)$$

$$\ddot{p} = \left[\frac{\psi''(\ddot{\nu})}{\widehat{\mu}_3} \right]^{\frac{1}{3}}, \quad (3.22)$$

$$\ddot{a} = \exp \left\{ \bar{Y}_{in} - \left[\frac{\widehat{\mu}_3}{\psi''(\ddot{\nu})} \right]^{\frac{1}{3}} \psi(\ddot{\nu}) \right\}. \quad (3.23)$$

Stacy a Mihram (1965) uvádějí odlišné vyjádření odhadů parametrů, kde parametr p určují z rovnice (3.15) a následně pro vyjádření odhadu parametru ν se použije rovnice (3.16). Zde ovšem nastává nejednoznačnost řešení neboť parametr p se vyjádří jako

$$p = \pm \sqrt{\frac{\psi'(\nu)}{\mu_2(Y_i)}},$$

kde nelze nijak určit, zda máme volit odhad parametru p kladný či záporný. Níže jsou

uvedeny konkrétní odhady parametrů, tak jak to uvádějí Stacy a Mihram (1965).

$$\pm \widehat{\gamma}_y = \frac{\psi''(\check{\nu})}{[\psi'(\check{\nu})]^{\frac{3}{2}}}, \quad (3.24)$$

$$\check{p} = \pm \sqrt{\frac{\psi'(\check{\nu})}{S_y^2}}, \quad (3.25)$$

$$\check{a} = \exp \left\{ \bar{Y}_n \mp \sqrt{\frac{S_y^2}{\psi'(\check{\nu})}} \psi(\check{\nu}) \right\}, \quad (3.26)$$

kde $\widehat{\gamma}_y$ je odhad šikmosti náhodné veličiny Y_i .

$$\widehat{\gamma}_y = \frac{n}{(n-1)(n-2)S_y^3} \sum_1^n (Y_i - \bar{Y}_n)^3.$$

Námi zvolené odhady se nijak výrazně neliší od odhadů, které jsou zmíněny v Stacy a Mihram (1965). Tabulky 3.1 a 3.2 to ukazují na průměrech, směrodatných odchylkách a mediánech odhadů spočtených z 1000 opakování pro data generovaná ze zobecněného gama rozdělení o rozsahu výběru $n = 500$.

	\check{a}	$\check{\nu}$	\check{p}	\check{a}	$\check{\nu}$	\check{p}
Skutečná hodnota parametru	3	4	2	4	2	-2
Výběrový průměr	2.807	5.377	1.977	4.813	2.313	-2.012
Medián	2.805	4.309	1.926	4.199	2.109	-1.944
Směrodatná odchylka	1.240	6.733	0.552	2.702	1.038	0.524

Tabulka 3.1: Odhad dle Stacy a Mihram (1965)

	\check{a}	$\check{\nu}$	\check{p}	\check{a}	$\check{\nu}$	\check{p}
Skutečná hodnota parametru	3	4	2	4	2	-2
Výběrový průměr	2.780	5.438	1.965	4.866	2.338	-1.998
Medián	2.775	4.357	1.915	4.236	2.132	-1.932
Směrodatná odchylka	1.237	6.814	0.548	2.774	1.049	0.517

Tabulka 3.2: Námi odvozený odhad

4. Simulační studie

Na závěr práce se zaměříme na posouzení kvality odhadů zmíněných v kapitole 3. Tedy půjde o odhad metodou maximální věrohodnosti (MLE), přímý odhad momentovou metodou (MM Přímá) a odhad momentovou metodou užitím transformace (MM Log). Odhady budeme porovnávat na výběru vygenerovaném ze zobecněného gama rozdělení. Rozsahy výběrů budou postupně 100, 200, 500 a 1000 hodnot a pro každý rozsah provedeme 1000 iterací těchto výběrů. Pro každé opakování odhadneme parametry a následně z těchto hodnot spočteme výběrové průměry a směrodatné odchyly. Na základě těchto hodnot budeme moct diskutovat konzistenci odhadu, pokud se nám bude s rostoucím rozsahem výběrů snižovat rozdíl mezi skutečnou hodnotou parametru a našimi průměry odhadů a zároveň se bude směrodatná odchyly blížit k nule. Na základě směrodatné odchyly pak budeme porovnávat přesnost odhadů. Ve výstupu budeme také vždy uvádět, pro kolik výběrů se nám podařilo odhady spočítat, protože v některých případech se příslušné soustavy rovnic nepodařilo numericky vyřešit.

Pro tvorbu simulací jsme využili program Wolfram Mathematica, kde jsme naprogramovali všechny potřebné funkce, včetně hustoty zobecněného gama rozdělení, jelikož v místní dokumentaci mají toto rozdělení definované jen pro parametr $p > 0$.

4.1 Simulace pro $p > 0$

Jako první jsme pro tvorbu náhodného výběru zvolili zobecněné gama rozdělení s parametry $a = 4$, $\nu = 3$ a $p = 2$. V Tabulce 4.1 a Obrázku 4.1 lze vidět výsledky naší simulace.

Metoda odhadu	Rozsah výběru	Průměry			Směrodatné odchyly			Úspěšné iterace
		\hat{a}	$\hat{\nu}$	\hat{p}	\hat{a}	$\hat{\nu}$	\hat{p}	
MLE	100	3.814	4.253	2.186	2.243	3.233	0.971	1000
	200	3.824	3.995	2.050	1.830	2.708	0.683	1000
	500	3.908	3.391	2.014	1.227	1.565	0.420	1000
	1000	3.972	3.164	2.014	0.895	0.968	0.299	1000
MM Přímá	100	4.227	3.665	2.261	2.042	2.803	0.871	970
	200	4.023	3.597	2.101	1.785	2.265	0.656	967
	500	3.986	3.292	2.039	1.210	1.500	0.419	952
	1000	4.014	3.123	2.029	0.912	0.936	0.306	945
MM Log	100	3.406	8.307	1.883	2.440	14.733	1.159	987
	200	3.527	4.979	1.903	1.998	4.800	0.791	1000
	500	3.746	3.606	1.973	1.968	1.792	0.495	1000
	1000	3.869	3.267	1.989	1.038	1.049	0.354	1000

Tabulka 4.1: Výsledky simulací pro výběry z $GF(4, 3, 2)$

Z výsledků lze usoudit, že všechny průměry odhadů s rostoucím rozsahem konvergují k skutečným hodnotám parametrů a zároveň se směrodatné odchyly blíží k nule a tudíž tyto výsledky naznačují konzistenci odhadů.

Nejlepší výsledky ve smyslu velikosti směrodatné odchyly a rozdílu mezi výběrovými průměry a skutečnou hodnotou parametru vykazuje metoda maximální věrohodnosti.

Přímá momentová metoda v tomto případě dává velmi podobné výsledky jako metoda

maximální věrohodnosti. Na rozdíl od ní je ovšem numericky náročnější na řešení a jak je vidět například v Tabulce 4.1, nejsou odhady pro všechny iterace numericky spočtené.

Momentová metoda využívající transformaci je pro nízké rozsahy nevyužitelná, neboť pro parametr ν vrací odhady s velkým rozptylem, což je patrné i z Obrázku 4.1b. Z obrázku a Tabulky 4.2 můžeme též nahlédnout, že medián odhadů parametru ν je i pro menší rozsahy výběrů blíže skutečné hodnotě než průměry, je to způsobeno tím, že několik odhadů vyšlo velmi odlišně od skutečné hodnoty parametru a tudíž výběrový průměr, který je na odlehlá pozorování citlivý, se jimi nechá vychýlit.

	\hat{a}	$\hat{\nu}$	\hat{p}
100	3.413	3.605	1.818
200	3.579	3.406	1.879
500	3.805	3.143	1.941
1000	3.932	3.065	1.975

Tabulka 4.2: Mediány odhadů pro MM Log z $\Gamma(4, 3, 2)$

Tabulka 4.3 ukazuje výsledky pro druhou simulaci s $p > 0$. Tentokrát jsme jako parametry zvolili $a = 6$, $\nu = 5$ a $p = 5$. Závěry více méně odpovídají těm u simulace s volbou menších parametrů. Problematický je odhad parametru ν , který pro všechny metody vychází i pro rozsah 1000 asi o 0,5 vyšší než je jeho skutečná hodnota a směrodatné odchytky se pohybují kolem 2. Je zřejmé, že se zvyšujícími se hodnotami parametrů se přesnost všech metod zhoršuje.

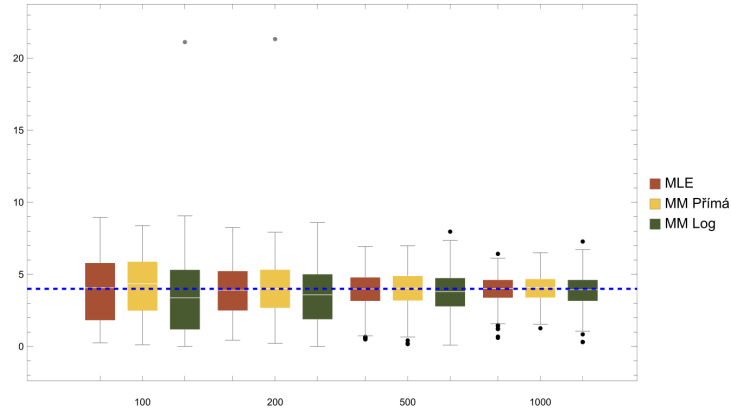
U momentové metody s transformací nám velmi odlehlé výsledky vycházejí pro nižší rozsahy výběrů i u parametru a , pro který nám v Tabulce 4.1 vycházely přijatelné hodnoty. Opět se zde ukazuje, že pro tuto metodu jsou mnohem vypovídající mediány nežli výběrové průměry viz Tabulka 4.4.

Metoda odhadu	Rozsah výběru	Průměry			Směrodatné odchytky			Úspěšné iterace
		\hat{a}	$\hat{\nu}$	\hat{p}	\hat{a}	$\hat{\nu}$	\hat{p}	
MLE	100	5.702	6.835	5.454	1.639	5.029	2.442	1000
	200	5.752	6.514	5.144	1.378	4.139	1.800	1000
	500	5.798	5.897	5.015	1.047	3.104	1.178	1000
	1000	5.888	5.452	5.005	0.758	1.927	0.858	997
MM Přímá	100	5.500	8.099	5.287	1.928	7.500	2.432	999
	200	5.565	7.417	5.077	1.648	6.305	1.868	1000
	500	5.775	5.999	5.004	1.104	3.299	1.208	1000
	1000	5.888	5.472	5.012	0.788	2.035	0.877	1000
MM Log	100	22.282	39.320	4.880	95.047	162.457	3.514	1000
	200	6.130	12.409	4.888	21.119	22.925	2.243	1000
	500	5.618	6.593	4.923	1.369	4.706	1.408	1000
	1000	5.812	5.630	4.967	0.929	2.291	0.993	1000

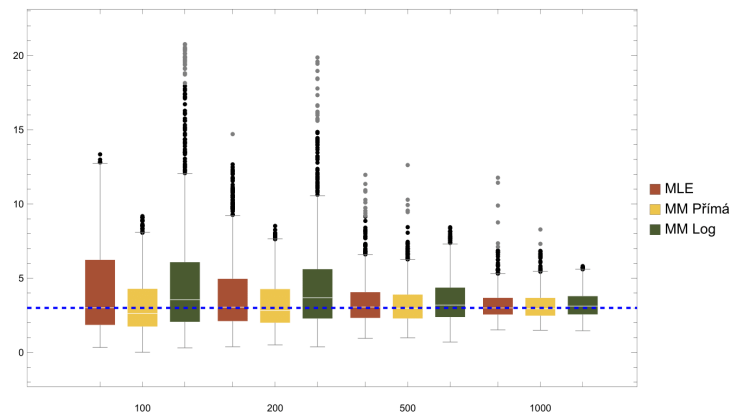
Tabulka 4.3: Výsledky simulací pro výběry z $\Gamma(6, 5, 5)$

	\hat{a}	$\hat{\nu}$	\hat{p}
100	5.846	5.861	4.680
200	5.743	5.664	4.720
500	5.901	5.189	4.870
1000	5.944	5.100	4.931

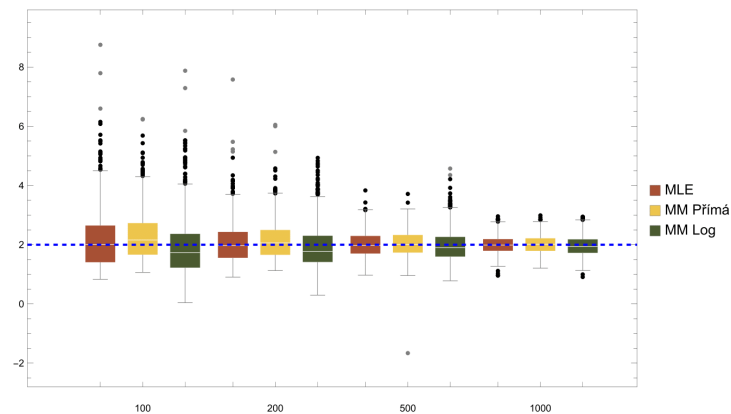
Tabulka 4.4: Mediány odhadů pro MM Log z $\Gamma(6, 5, 5)$



(a) Parametr a



(b) Parametr ν



(c) Parametr p

Obrázek 4.1: Rozložení odhadů pro skutečné hodnoty parametrů $a = 4$, $\nu = 3$ a $p = 2$ pro rozsahy výběrů 100, 200, 500 a 1000

4.2 Simulace pro $p < 0$

Odhady porovnáme také pro případ, kdy parametr p nabývá záporných hodnot. Budeme tedy simulovat výběr ze zobecněného gama rozdělení s parametry $a = 4$, $\nu = 2$ a $p = -3$, tak aby byla splněna podmínka, pro existenci třetích momentů $3 < -p\nu$, kterou vyžadujeme pro přímou momentovou metodu.

Metoda odhadu	Rozsah výběru	Průměry			Směrodatné odchylky			Úspěšné iterace
		\hat{a}	$\hat{\nu}$	\hat{p}	\hat{a}	$\hat{\nu}$	\hat{p}	
MLE	100	5.477	2.872	-3.209	3.670	2.295	1.392	1000
	200	4.729	2.458	-3.093	2.393	1.590	0.885	1000
	500	4.225	2.165	-3.034	1.019	0.823	0.547	1000
	1000	4.108	2.085	-3.008	0.561	0.512	0.384	1000
MM Přímá	100	19.068	6.231	-1.890	50.617	4.222	0.810	999
	200	11.304	4.967	-2.163	16.024	3.402	0.966	1000
	500	6.968	3.731	-2.457	5.136	2.226	1.043	1000
	1000	5.714	3.138	-2.605	3.110	1.615	0.991	1000
MM Log	100	14.626	4.492	-2.973	35.848	5.595	2.702	1000
	200	6.881	2.934	-2.977	13.675	2.385	1.417	1000
	500	4.484	2.307	-2.989	1.472	0.953	0.766	1000
	1000	4.244	2.172	-2.996	0.786	0.641	0.526	1000

Tabulka 4.5: Výsledky simulací pro výběry z $\text{GF}(4, 2, -3)$

Jak lze nahlédnout z Tabulky 4.5 pro výběr z rozdělení se zápornou hodnotou parametru p , výsledky opět naznačují konzistenci odhadů. V tomto případě se jako jasné nejlepší odhad nabízí odhad metodou maximální věrohodnosti.

Naopak přímá momentová metoda se pro tuto volbu parametrů chová nepřesně a i pro rozsah 1000 pozorování má výrazné směrodatné odchylky a průměry odhadů jsou velmi vzdálené od skutečných hodnot parametrů. To, že jsou tyto odhady značně nepřesné ukazují i Obrázky 4.2, kde je vidět, že medián odhadů leží výrazně mimo skutečné hodnoty parametrů.

Momentová metoda využívající transformaci náhodné veličiny má stále podobné vlastnosti, kdy se odhady chovají, jak bychom požadovali až pro výběry o rozsahu 500. Jak je ovšem vidět i na Obrázcích 4.2 problémem této metody je opět zejména to, že několik odhadů je velmi odlišných od skutečné hodnoty parametru. To je patrné i při pohledu na mediány těchto odhadů, které jsou skutečným hodnotám parametrů mnohem bližší i pro nižší rozsahy výběru viz Tabulky 4.6 a 4.8.

	\hat{a}	$\hat{\nu}$	\hat{p}
100	4.623	2.544	-2.601
200	4.269	2.250	-2.823
500	4.150	2.138	-2.890
1000	4.084	2.075	-2.932

Tabulka 4.6: Mediány odhadů pro MM Log z $\text{GF}(4, 2, -3)$

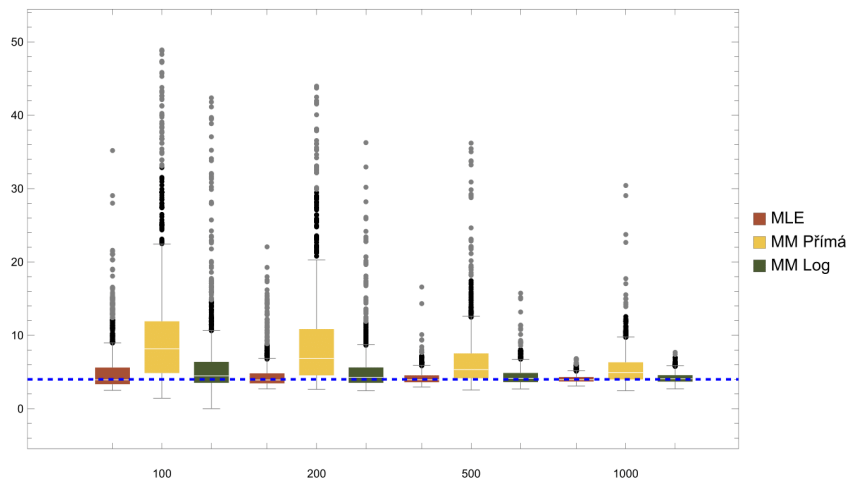
Simulaci jsme opět zopakovali pro jinou volbu parametrů se záporným p a to pro $a = 4$, $\nu = 4$ a $p = -6$ viz Tabulka 4.7. Výsledky jsou opět horší, než u předchozí volby parametrů, kdy jsme pro ně volili v absolutní hodnotě nižší hodnoty.

Metoda odhadu	Rozsah výběru	Průměry			Směrodatné odchylky			Úspěšné iterace
		\hat{a}	$\hat{\nu}$	\hat{p}	\hat{a}	$\hat{\nu}$	\hat{p}	
MLE	100	4.737	6.079	-6.616	1.982	5.537	3.166	1000
	200	4.368	5.062	-6.307	1.293	3.574	2.069	1000
	500	4.161	4.506	-6.100	0.659	2.117	1.347	1000
	1000	4.101	4.321	-6.014	0.451	1.473	0.984	1000
MM Přímá	100	8.761	10.359	-5.193	8.519	8.365	5.186	993
	200	6.778	8.138	-5.775	6.181	7.309	4.736	998
	500	4.737	5.501	-6.199	2.807	3.917	4.541	997
	1000	4.553	4.736	-6.317	2.243	2.657	3.570	997
MM Log	100	20.237	21.507	-5.545	64.758	62.023	4.209	1000
	200	10.302	8.796	-5.854	39.561	16.898	2.532	1000
	500	4.472	4.968	-5.931	2.138	2.902	1.595	1000
	1000	4.190	4.501	-5.934	0.664	1.733	1.143	1000

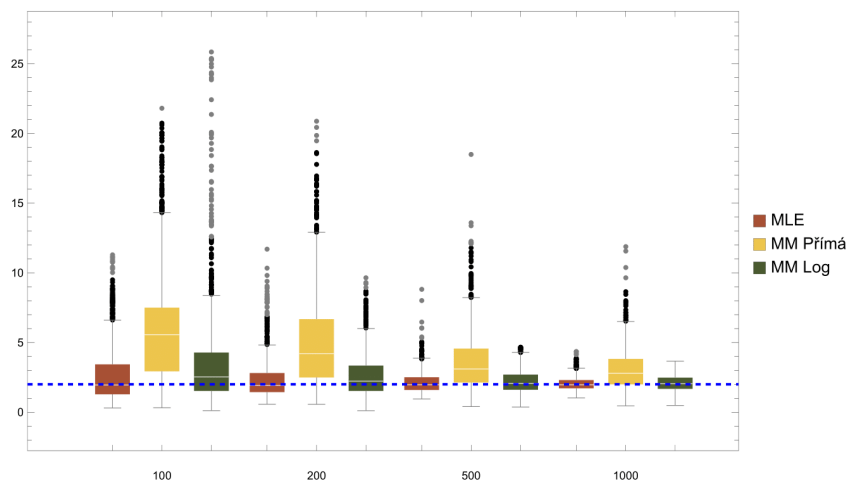
Tabulka 4.7: Výsledky simulací pro výběry z $\text{G}\Gamma(4, 4, -6)$

	\hat{a}	$\hat{\nu}$	\hat{p}
100	4.163	4.922	-5.304
200	4.104	4.418	-5.747
500	4.064	4.218	-5.848
1000	4.049	4.135	-5.888

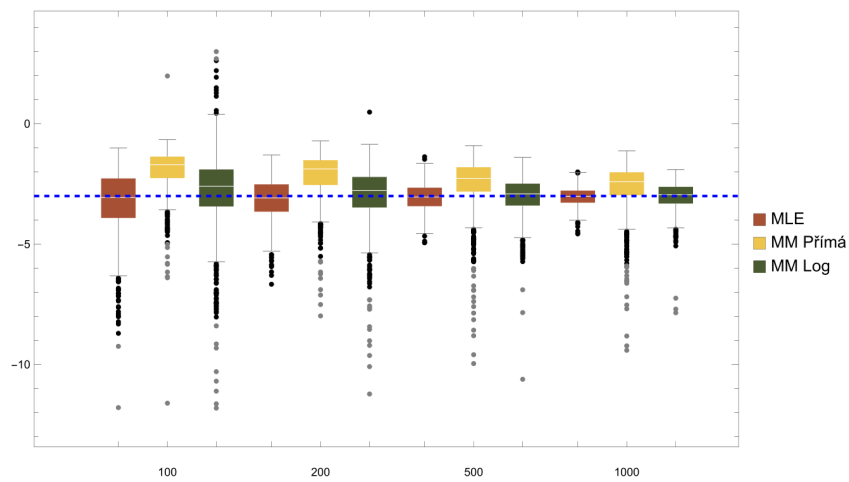
Tabulka 4.8: Mediány odhadů pro MM Log z $\text{G}\Gamma(4, 4, -6)$



(a) Parametr a



(b) Parametr ν



(c) Parametr p

Obrázek 4.2: Rozložení odhadů pro skutečné hodnoty parametrů $a = 4$, $\nu = 2$ a $p = -3$ pro rozsahy výběrů 100, 200, 500 a 1000

4.3 Odhady pro klasické gama rozdělení

Zobecněné gama rozdělení zavádíme proto, abychom měli flexibilnější model, který by dokázal lépe a přesněji aproximovat skutečné rozdělení v populaci. Tato navýšená flexibilita, kterou získáme přidáním parametru p , ovšem zřejmě způsobí, že při odhadu parametrů budeme dostávat méně přesné výsledky, než kdybychom uvažovali pouze klasické gama rozdělení. Jak je tento rozdíl výrazný budeme porovnávat na simulacích, které provedeme podobně jako výše, kdy výběr budeme generovat ze zobecněného gama rozdělení s parametry $a = 3$, $\nu = 2$ a $p = 1$, tedy z klasického gama rozdělení s parametry $a = 3$ a $\nu = 2$. V Tabulce 4.9 jsou vidět odhady pro parametry zobecněného gama rozdělení a v Tabulce 4.10 jsme parametry odhadli metodou maximální věrohodnosti a momentovou metodou pro klasické gama rozdělení.

Metoda odhadu	Rozsah výběru	Průměry			Směrodatné odchyly			Úspěšné iterace
		\hat{a}	$\hat{\nu}$	\hat{p}	\hat{a}	$\hat{\nu}$	\hat{p}	
MLE	100	3.454	3.130	1.074	2.688	3.262	0.461	996
	200	3.145	2.543	1.022	2.063	1.806	0.306	1000
	500	3.052	2.169	1.008	1.333	0.793	0.182	1000
	1000	3.050	2.069	1.007	0.985	0.496	0.129	1000
MM Přímá	100	5.381	1.476	1.371	2.742	0.779	0.471	980
	200	4.536	1.675	1.217	2.264	0.750	0.335	983
	500	3.819	1.861	1.736	1.736	0.694	0.229	995
	1000	3.480	1.945	1.059	1.423	0.608	0.180	995
MM Log	100	3.507	4.668	1.031	13.877	7.353	0.965	988
	200	2.803	2.941	0.972	2.215	2.104	0.357	992
	500	2.893	2.326	0.989	1.635	0.961	0.247	1000
	1000	2.946	2.154	0.995	1.250	0.610	0.175	1000

Tabulka 4.9: Výsledky simulací pro výběry z $\mathbf{G}\Gamma(3, 2, 1)$

Metoda odhadu	Rozsah výběru	Průměry		Směrodatné odchyly	
		\hat{a}	$\hat{\nu}$	\hat{a}	$\hat{\nu}$
MLE	100	2.975	2.050	0.447	0.279
	200	2.989	2.025	0.312	0.190
	500	2.992	2.011	0.197	0.119
	1000	2.997	2.004	0.143	0.082
MM	100	2.998	2.062	0.577	0.366
	200	2.992	2.036	0.399	0.248
	500	2.994	2.014	0.250	0.156
	1000	2.997	2.007	0.178	0.110

Tabulka 4.10: Výsledky simulací pro klasické gama rozdělení $\Gamma(3, 2)$

Z tabulek můžeme vyčíst, že odhady uvažující klasické gama rozdělení jsou někdy takřka o řád přesnější ve smyslu velikosti směrodatné odchyly než ty předpokládající zobecněné gama rozdělení.

4.4 Shrnutí simulací

Na základě simulací lze odvodit, že pro odhad parametrů zobecněného gama rozdělení je potřebné mít dostatečný rozsah výběru o velikosti alespoň 500 pozorování. Zároveň se ale ukazuje, že pro v absolutní hodnotě rostoucí hodnoty parametrů se zhoršují odhady zejména parametru ν .

Jako nejspolehlivější metoda, z těch které jsme uváděli, pro odhad parametrů je metoda maximální věrohodnosti. Ta ve všech případech prokazovala konzistenci a takřka vždy pro ní vycházela nejnižší směrodatná odchylka a také rozdíly od skutečných hodnot parametrů. Při jednotlivých simulacích ukazovala i numerickou stabilitu, kdy se rovnice podařilo vyřešit takřka ve všech případech a nenastávaly velmi odlehlé výsledky.

Přímá momentová metoda se zdá být pro numerické výpočty nejnáročnější, kdy se nejčastěji nepodařilo určit odhady. Pro případ, kdy bychom uvažovali, že p může být záporné bych tuto metodu nedoporučil. Není zde totiž pro všechny kombinace parametrů zajištěna existence momentů a navíc i výsledky pro přípustné volby $p < 0$ byly velmi nepřesné. Metoda se zdála být pro $p > 0$ přijatelná, ovšem případ, kdy jsme zvolili $p = 1$ ukázal, že je její přesnost závislá na konkrétní volbě parametrů.

Momentová metoda využívající transformaci dává podobné výsledky jak pro $p > 0$ tak pro $p < 0$. Ve všech případech u ní pro rozsah pozorování 100 nebo i 200 nastávali velmi odlehlé výsledky, které způsobovaly velké rozdíly mezi průměry odhadů a skutečnými hodnotami parametrů. Většina odhadů však byla poměrně přesná, což jsme ukázali mediány hodnot. Dalo by se tedy říct, že tato metoda by mohla být asi v 95 % přesnější než přímá metoda, oproti které má ještě tu výhodu, že ji lze použít pro všechny volby parametrů.

V praxi by tedy bylo vhodné zamyslet se, zda máme k dispozici dostatek dat, abychom mohli uvažovat zobecněné gama rozdělení, jako náš model. Pokud bychom se touto cestou vydali, doporučil bych pro odhad parametrů využít metodu maximální věrohodnosti. V případě, že bychom neměli dostatečné množství dat, například méně než 500, by bylo vhodné uvážit, zda nepoužít k jejich modelování některé dvouparametrické rozdělení.

Závěr

V bakalářské práci jsme popsali zobecněné gama rozdělení tak, jak ho definuje Stacy a Mihram (1965). Popsali jsme jeho souvislost s běžným gama rozdělením a dalšími známými rozděleními. Následně jsme odvodili jeho vlastnosti a diskutovali jsme za jaké volby parametrů je rozdělení unimodální.

V další části práce jsme odvozovali odhady jednotlivých parametrů. Nejprve jsme uvedli metodu maximální věrohodnosti, u které jsme oproti Johnson a kol. (1994) odhady zobecnili pro případ, kdy je $p < 0$, a zároveň jsme opravili chybu, kterou v této knize uvádějí. Jako další jsme odvodili momentovou metodu, která byla většinou autorů opomíjena. Na závěr jsme uvedli jiný odhad momentovou metodou tak, jak je popsán v článku Stacy a Mihram (1965). U něj jsme odvodili jiné rovnice pro odhad parametrů tak, aby nenastával problém při určení znaménka u parametru p .

Odhady jsme následně porovnali v simulační studii. Naše výsledky naznačovali konzistenci u všech odhadů, pro jejich přesnost je však vždy třeba velké množství dat. Jako nejpřesnější metoda se poté ukázala metoda maximální věrohodnosti.

Pro zobecněné gama rozdělení existují také další parametrizace, které jsme již v práci neuvedli. Jednu z nich například uvádí a využívá pro analýzu doby do události Cox a kol. (2007).

Seznam použité literatury

- COX, C., CHU, H., SCHNEIDER, M. F. a MUÑOZ, A. (2007). Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*, **26**, 4353–4374. URL <https://api.semanticscholar.org/CorpusID:20211482>.
- DUPAČ, V. a HUŠKOVÁ, M. (2013). *Pravděpodobnost a matematická statistika*. Skripta Univerzity Karlovy v Praze. Karolinum, Praha, 2., upravené vydání edition. ISBN 978-80-246-2208-8.
- HAGER, H. W. (1969). *Statistical inferences for the generalized gamma distribution*. PhD thesis, University of Missouri–Rolla.
- JOHNSON, N. L., KOTZ, S. a BALAKRISHNAN, N. (1994). *Continuous univariate distributions*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley-Interscience publication John Wiley Sons, New York, second edition. ISBN 0-471-58495-9.
- STACY, E. W. a MIHRAM, G. A. (1965). Parameter estimation for a generalized gamma distribution. *Technometrics*, **7**(3), 349–358. ISSN 00401706. URL <http://www.jstor.org/stable/1266594>.