Univerzita Karlova v Praze

Pedagogická fakulta

BAKALÁŘSKÁ PRÁCE

Bc. Barbora Šimková

Odhady parametrů rozdělení náhodných veličin

Katedra matematiky a didaktiky matematiky

Vedoucí bakalářské práce:	RNDr. František Mošna, Dr.
Studijní program:	Specializace v pedagogice
Studijní obor:	Matematika se zaměřením na vzdělávání a technická a informační výchova se zaměřením na vzdělávání

Praha 2013

Na tomto místě bych ráda poděkovala RNDr. Františku Mošnovi, Dr., vedoucímu mé bakalářské práce. Především za odborné vedení, věnovaný čas a cenné rady. Děkuji také svým rodičům, kteří mě vždy podporovali ve studiu.

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

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

V Praze dne 3.5.2013 Bc. Barbora Šimková

Název práce: Odhady parametrů rozdělení náhodných veličin

Autor: Bc. Barbora Šimková

Katedra: Katedra matematiky a didaktiky matematiky

Vedoucí bakalářské práce: RNDr. František Mošna, Dr.

Abstrakt: Předmětem této bakalářské práce je porovnání základních metod, kterými je možné spočítat bodové odhady spojitých a diskrétních pravděpodobnostních rozdělení. Práce se zabývá rozborem dvou metod, jedná se o momentovou metodu a metodu maximální věrohodnosti. Tyto metody se používají pro odhad bodových parametrů pravděpodobnostních rozdělení. Momentovou metodou rozumíme porovnání teoretických a výběrových momentů náhodné veličiny. Metodu maximální věrohodnosti bereme jako další alternativu při výpočtech bodových odhadů, která využívá klasický postup hledání maxima funkce s využitím vlastností náhodného výběru. Způsoby výpočtů vychází ze statistických metod a mohly by být vhodné pro rozšíření výuky základního kurzu pravděpodobnosti a statistiky na PedF UK. Práce je přehledem odhadů parametrů základních distribucí a srovnání kvality dvou základních metod pro jejich určení.

Klíčová slova: odhady parametrů, rozdělení náhodné veličiny, metoda maximální věrohodnosti, momentová metoda

Title: Parameter estimation of random variables distribution

Author: Bc. Barbora Šimková

Department: Department of Mathematics and Mathematical Education

Supervisor: RNDr. František Mošna, Dr.

Abstract: The subject of this thesis is to compare basic methods by which it is possible to calculate point estimates of discrete and continuous probability distributions. The work deals with the analysis of the two methods - the method of moments and maximum likelihood method. These methods are used for point estimates of probability distributions parameters. The method of moments studies the comparison between the theoretical and sample moments of a random variable. The method of maximum likelihood is another alternative for the calculation of point estimates, which uses the classical approach of finding the maximum of a function, using the properties of random selection. These methods of calculation are based on statistical methods and could be useful for extending the basic course on probability and statistics at Charles University´s Faculty of Education. The work is an overview of the estimated parameters of the basic distribution and compares the quality of two basic methods for their estimation.

Keywords: parameter estimation, distribution of random variables, maximum likelihood method, method of moments

Obsah

Úvod

V současném světě se setkáváme se statistikou, jejími důsledky a aplikacemi prakticky každý den. Ať už při studiu, v zaměstnání, při čtení denního tisku, při výběru hypotéky nebo při vybírání dovolené. Statistika nám pomáhá pochopit vztahy nás obklopující napříč vědními obory. Nejčastěji je spojována s matematikou, fyzikou, informatikou a ekonomií. Neméně ji využívají i humanitní obory jako je psychologie a sociologie. Statistika je naprosto nepostradatelnou součástí naší společnosti a bude tomu tak i v budoucnu.

Úkolem této bakalářské práce je pro účely výuky základního kurzu pravděpodobnosti a statistiky na PedF UK podrobně zpracovat a na datech replikovat statistické odhady a jejich vlastnosti pro určitá pravděpodobnostní rozdělení s využitím statistických metod. Metody použité v této práci se aplikují pro odhadování parametrů náhodných veličin a vektorů v široké škále reálných statistických, ekonomických, geografických, pojišťovnických a dalších problémů.

První kapitola popisuje základní pojmy z pravděpodobnosti rozdělení jako náhodné veličiny, základní druhy pravděpodobnostních rozdělení a podrobněji definuje ta rozdělení, které budou používána v dalších kapitolách, pomocí jejich hustoty nebo pravděpodobnostní funkce i s jejich základními momenty. Klíčovými pojmy matematické statistiky jsou náhodný výběr, parametrický model, bodový odhad parametru a jeho vlastnosti, které jsou ve statistické části této kapitoly. Definice pojmů jsme čerpali hlavně z [1] a [7].

Ve druhé kapitole této práce se zaměříme na teoretické odvození výpočtů bodových odhadů metodami, které jsou používané pro odhadování bodových parametrů v parametrických modelech. K odhadu parametrů budeme používat metodu maximální věrohodnosti (maximum-likelihood estimation) a metodu momentů (method of moments). Pro obě tyto metody sepíšeme předpoklady pro jejich použití a podrobný postup výpočtu. Vysvětlíme zde výběrové a teoretické momenty náhodných veličin, věrohodnostní funkci a způsob nalezení maximálně věrohodného odhadu.

Po vysvětlení teoretické a popisné stránky dané oblasti se budeme ve třetí kapitole věnovat výpočtu odhadů bodových parametrů náhodného výběru spojitých pravděpodobnostních rozdělení. Dále je zde uveden postup ověření základních vlastností jednotlivých odhadů a jejich následné srovnání. Ve výpočtech požadovaných odhadů jsme uvažovali pouze náhodné výběry o rozsahu n, jejichž náhodné veličiny mají spojité rozdělení (normální, gamma a logaritmicko-normální rozdělení).

Čtvrtá kapitola bude věnovaná diskrétním pravděpodobnostním rozdělením, přesněji geometrickému rozdělení, Poissonovu rozdělení a binomickému rozdělení. Pro tyto rozdělení odhadneme jejich parametry metodou momentů a metodou maximální věrohodnosti a ověříme vlastnosti těchto odhadů.

V páté kapitole uvedeme přehled všech odhadů parametrů včetně jejich vlastností, které jsme v bakalářské práci spočítali.

1. Základní pojmy

V této kapitole uvedeme některé pojmy z pravděpodobnosti a statistiky, které budeme používat v dalších částech práce. Vysvětlíme si zde, co je náhodná veličina, jak se popisuje její rozdělení, uvedeme některé druhy pravděpodobnostních rozdělení, a to jak spojitých, tak i diskrétních náhodných veličin a k nim jejich dva hlavní momenty. Ze statistiky si definujeme pojmy náhodný výběr, parametrický model, výběrový prostor, bodový odhad, vlastnosti bodového odhadu a další. Předpokládáme, že čtenář je již seznámen se základními pojmy z matematické analýzy, pravděpodobnosti a matematické statistiky. Vycházíme přitom ze zdrojů: [1], [2], [4], [6] a [7].

Náhodná veličina

Nechť je dán pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathcal{P})$. Měřitelné zobrazení $X \ : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$, kde $\mathcal{X}$ je nějaká množina a $\mathcal{B}$ nějaká σ-algebra na $\mathcal{X}$, nazveme náhodnou veličinou. Množinu $\mathcal{X}$ nazýváme výběrový prostor [4].

Rozdělení náhodné veličiny

Rozdělením náhodné veličiny $X \ : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ rozumíme indukovanou pravděpodobnostní míru P_X na $(\mathcal{X}, \mathcal{B})$ definovanou vztahem

$$P_X(B) = P(\{\omega \in \Omega \ : X(\omega) \in B\}), \ B \in \mathcal{B}.$$

Pravděpodobnostní prostor $(\Omega, \mathcal{A}, \mathcal{P})$ se pro danou náhodnou veličinu transformuje na pravděpodobnostní prostor $(\mathcal{X}, \mathcal{B}, P_X)$ [4].

Hustota

Nechť $X \ : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ je náhodná veličina, nechť μ je σ-konečná míra na $\mathcal{X}$ a nechť P_X je absolutně spojitá vzhledem k μ. Pak existuje reálná měřitelná nezáporná funkce $f_X(x)$ taková, že pro každou měřitelnou funkci $h : (\mathcal{X}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}_0)$ platí

$$\int_{\mathcal{X}} h(x) \ dP_X(x) = \int_{\mathcal{X}} h(x) f_X(x) \ d\mu(x).$$

Funkce $f_X(x)$ je určena jednoznačně μ-skoro všude [4]. Tato věta se někdy nazývá Radon-Nikodymova věta a funkce f_X z předchozí věty se nazývá hustotou náhodné veličiny X vzhledem k míře μ a dále ji v textu budeme značit $f(x)$.

Distribuční funkce

Funkci $F_X : \mathbb{R} \to \mathbb{R}$ definovanou vztahem $F_X(x) = P[X \leq x]$ nazýváme distribuční funkcí náhodné veličiny X. Distribuční funkce F_X jednoznačně charakterizuje rozdělení X viz [4].

Druhy náhodných veličin

Náhodné veličiny dělíme na různé druhy. V matematické statistice mají nejdůležitější uplatnění následující dva typy rozdělení. Pokud P_X je absolutně spojitá vzhledem k Lebesgueově míře λ, tak X je spojitá náhodná veličina, je-li P_X je absolutně spojitá vzhledem k čítací míře μ_S, kde S je nejvýše spočetná množina v $\mathbb{R}$, tak náhodná veličina X je diskrétní náhodná veličina více v literatuře [4].

U spojité náhodné veličiny platí $F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(t)\ d(t)$. Pro diskrétní náhodné veličiny s hodnotami v S máme $F_X(x) = \displaystyle\sum_{t \in S, t \leq x} P[X = t]$ v [4].

Momenty náhodné veličiny

Střední hodnotou EX (reálné) náhodné veličiny X rozumíme reálné číslo dané výrazem

$$EX = \int_{\Omega} X(\omega)\ dP_X(\omega),$$

pokud integrál na pravé straně existuje [4]. V práci budeme střední hodnotu značit μ' nebo EX.

Rozptyl Var$X$ náhodné veličiny X je její druhý centrální moment. Je definován následujícím vzorcem $\mathrm{Var}X = \mathrm{E}(X - \mathrm{E}X)^2$ [4] a budeme ho značit podle potřeby σ^2 nebo Var$X$.

Pravděpodobnostní rozdělení

V práci se zaměřujeme pouze na náhodné veličiny se spojitým nebo diskrétním rozdělením (spojité nebo diskrétní náhodné veličiny) a nyní si některé definujeme pomocí jejich hustot, či pravděpodobnostních funkcí. Předpokládáme, že P_X je absolutně spojitá vzhledem k Lebesgueově míře λ u spojitých náhodných veličin a absolutně spojitá vzhledem k čítací míře μ_S u diskrétních náhodných veličin.

Spojitá rozdělení

Nyní si zadefinujeme některá spojitá rozdělení se kterými budeme později pracovat. Definice spojitých rozdělení jsme čerpali z [1].

Exponenciální rozdělení

Nechť $\lambda > 0$ je parametr. Exponenciální rozdělení, které značíme Ex(λ) má hustotu

$$f(x) = \frac{1}{\lambda} \exp\left[-\frac{x}{\lambda}\right],\ x > 0.$$

Má střední hodnotu $\mu' = \lambda$ a rozptyl $\sigma^2 = \lambda^2$. Exponenciální funkci v textu značíme buď $\exp[x]$ nebo e^x.

Spojité rovnoměrné rozdělení

Nechť $(\mu - h, h + \mu)$ je konečný nedegerovaný interval. Rovnoměrné rozdělení $\mathrm{R}(\mu - h, h + \mu)$ má hustotu

$$f(x) = \frac{1}{2h}, \text{ pro } \mu - h < x < \mu + h,$$

a platí pro něj $\mu' = \mu$, $\sigma^2 = \frac{h^2}{3}$.

Gama rozdělení

Nechť $a > 0$, $p > 0$. Gama rozdělení $\mathrm{Ga}(a, p)$ má hustotu

$$f(x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}, \ x > 0$$

a platí, že střední hodnota je $\mu' = \frac{p}{a}$ a rozptyl $\sigma^2 = \frac{p}{a^2}$.

Normální rozdělení

Nechť $\mu \in \mathbb{R}$, $\sigma^2 > 0$ jsou dané konstanty (parametry). Normální rozdělení je určeno hustotou

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

a označuje se symbolem $\mathrm{N}(\mu, \sigma^2)$ a platí, že střední hodnota je μ, a rozptyl σ^2.

Logaritmicko-normální rozdělení

Nechť $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ jsou dané parametry. Hustota logaritmicko-normálního rozdělení $\mathrm{LN}(\mu, \sigma^2)$je definována

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \text{ pro } x > 0.$$

Dále platí, že střední hodnota je $\mu' = e^{\mu + \frac{\sigma^2}{2}}$ a rozptyl $\sigma^2 = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma 2}$.

Diskrétní rozdělení

Geometrické rozdělení

Geometrické rozdělení $\mathrm{Ge}(p)$ má náhodná veličina X, která je dána jako počet pokusů, které musíme provést, aby nastal náhodný jev A, kde $P(A) = p$, $0 < p < 1$. Náhodná veličina X mající geometrické rozdělení má pravděpodobnostní funkci p, která je definována vztahem

$$P(X = k) = p(1-p)^{k-1}, \ k \in \mathbb{N}.$$

Pro základní číselné charakteristiky platí, že střední hodnota je $\mu' = \frac{1}{p}$ a rozptyl $\sigma^2 = \frac{1}{p}\left(\frac{1}{p} - 1\right)$ viz literatura [1].

Poissonovo rozdělení

Předpokládejme, že veličina X nabývá pouze hodnot $0, 1, \ldots$, a to s pravděpodobnostmi

$$P\left(X = x\right) = \frac{\lambda^x}{x!} \exp^{-\lambda},\ x = 0, 1, \ldots,$$

kde $\lambda > 0$ je dané číslo. Pak říkáme, že X má Poissonovo rozdělení $\mathrm{Po}(\lambda)$ s parametrem λ. Platí, že střední hodnota je $\mu' = \lambda$ a rozptyl $\sigma^2 = \lambda$ viz [7].

Binomické rozdělení

Binomické rozdělení $\mathrm{Bi}(m, p)$ má pravděpodobnostní funkci

$$P\left(X = x\right) = \binom{m}{x} p^k \left(1 - p\right)^{m-x},\ x = 0, 1, \ldots, m$$

a střední hodnota je $\mu' = mp$ a rozptyl $\sigma^2 = mp\left(1 - p\right)$ definice je z [7].

Náhodný výběr

Posloupnost $X_1, X_2, \ldots, X_n$ nezávislých stejně rozdělených náhodných veličin s distribuční funkcí F_0 se nazývá náhodný výběr z rozdělení F_0 [4]. Výběr pořizujeme proto, abychom se dozvěděli více o souboru dat, se kterými budeme pracovat. Budeme se soustředit na situaci, kdy budeme znát rozdělení souboru až na jeden nebo více parametrů (např. víme, že máme náhodně rozdělené náhodné veličiny, ale neznáme jejich střední hodnotu případně i rozptyl).

Model a parametr

Distribuční funkci F_0 neznáme. Chceme použít pozorování $X_1, \ldots, X_n$ k tomu, abychom se o ní něco potřebného dozvěděli. O distribuční funkci F_0 však předpokládáme, že patří do nějaké množiny rozdělení $\mathcal{F}$, které říkáme model pro data $X_1, \ldots, X_n$.

To, co se chceme o rozdělení F_0 dozvědět, nazýváme parametr. Vždy se jedná o nějakou konstantu, kterou bychom uměli zjistit, kdybychom znali F_0. Obecně tedy parametr budeme zapisovat $\boldsymbol{\theta}_0 \equiv t\left(F_0\right)$, kde t je nějaký funkcionál. Obvykle je $\boldsymbol{\theta}_0 \in \mathbb{R}^k$ pro $k \geq 1$ více o modelech v literatuře [4].

Parametrický model

Množina $\mathcal{F}$ je množina všech rozdělení s hustotami tvaru $f\left(x; \boldsymbol{\theta}\right)$ nebo pravděpodobnostními funkcemi $p\left(x; \boldsymbol{\theta}\right)$ pro $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^k$, kde $f\left(\cdot; \cdot\right)$ resp. $p\left(\cdot; \cdot\right)$ je známá funkce a $\boldsymbol{\theta}$ je neznámá konstanta (např. všechna exponenciální, normální, geometrická rozdělení). Tyto modely nazýváme parametrické. Parametry, které chceme odhadovat, jsou funkce složek $\boldsymbol{\theta}$. Definice je z [4], kde je parametrický model uváděn pouze pro rozdělení s hustotami.

Bodový odhad a jeho vlastnosti

Máme data $X_1, X_2 \dots, X_n$, model $\mathcal{F}$ a parametr $\boldsymbol{\theta} = t(F)$ pro $F \in \mathcal{F}$, který chceme v daném modelu odhadnout. Označme $\boldsymbol{X} = (X_1, \dots, X_n)^T$ (všechna data sestavená do jednoho vektoru). Nechť $F_0 \in \mathcal{F}$ je skutečné rozdělení náhodné veličiny X_i a $\boldsymbol{\theta}_0 \equiv t(F_0)$ je skutečná hodnota hledaného parametru. Odhadem paremetru $\boldsymbol{\theta}_0 \equiv t(F_0) \in \mathbb{R}^k$ rozumíme libovolnou měřitelnou funkci dat $\widehat{\boldsymbol{\theta}}_n \equiv \boldsymbol{T}_n(\boldsymbol{X}) \equiv \boldsymbol{T}_n(X_1, \dots, X_n)$ z [4].

Jinými slovy nám tato definice říká, že pro každou novou realizaci výběru obdržíme jiný bodový odhad. Bodový odhad nebývá početně náročný a obvykle ho používáme v případě, je-li rozsah výběrového souboru vzhledem k rozsahu základního souboru dostatečně velký.

Vlastnosti bodového odhadu $\widehat{\boldsymbol{\theta}}$ vypovídají o vhodnosti použité statistiky k odhadu hodnoty parametru $\boldsymbol{\theta}$. Dobrý bodový odhad musí splňovat určité vlastnosti: nestrannost (nevychýlenost, nezkleslenost), konzistenci, vydatnost (eficience) a dostatečnost.

Nestrannost

Odhad $\widehat{\boldsymbol{\theta}}_n \equiv \boldsymbol{T}_n(\boldsymbol{X})$ nazveme nestranným odhadem (unbiased estimator) parametru $\boldsymbol{\theta}_0$, právě když $\mathrm{E}\widehat{\boldsymbol{\theta}}_n = \boldsymbol{\theta}_0$ pro každé n viz [4]. Jinak řečeno, bodový odhad parametru $\boldsymbol{\theta}$ je nestranný, jestliže jeho střední hodnota se rovná tomuto parametru. Nestrannost říká, že odhad je "v průměru" (rozumíme přes všechny možné výběry) přesný.

Slabší formou nestrannosti je asymptotická nestrannost. Řekneme, že odhad je asymptoticky nestranný, pokud $\lim_{n\to\infty} \mathrm{E}\widehat{\boldsymbol{\theta}}_n = \boldsymbol{\theta}_0$.

Konzistence

Odhad $\widehat{\boldsymbol{\theta}}_n \equiv \boldsymbol{T}_n(\boldsymbol{X})$ nazveme konzistentním odhadem (consistent estimator) parametru $\boldsymbol{\theta}_0$, právě když $\widehat{\boldsymbol{\theta}}_n \xrightarrow{P} \boldsymbol{\theta}_0$ pro každé $n \to \infty$ [4]. Značkou $\xrightarrow{P}$ chápeme konvergenci v pravděpodobnosti k náhodné veličině, která je definována následovně: $X_n \xrightarrow{P} X$ pro $n \to \infty$

$$\lim_{n\to\infty} [\| X_n - X \| > \varepsilon] = 0 \text{ pro } \forall \varepsilon > 0.$$

Definice konzistence nám říká, že odhad je konzistentní, pokud se s rostoucím rozsahem výběru zpřesňuje. Konzistence je asymptotická vlastnost, která nic neříká o kvalitě odhadu při konečném n.

Postačující podmínky pro ověření konzistence dhadu $\widehat{\boldsymbol{\theta}}$ definujme následovně. Nechť $\mathrm{E}\widehat{\boldsymbol{\theta}}^2 < \infty$ pro každé přirozené n. Platí-li $\widehat{\boldsymbol{\theta}}$ je nestranný odhad nebo alespoň asymptoticky nestanný odhad parametru $\boldsymbol{\theta}$ a $\lim_{n\to\infty} \mathrm{Var}\ \widehat{\boldsymbol{\theta}} = 0$. Pak $\widehat{\boldsymbol{\theta}}$ je konzistentním odhadem parametru $\boldsymbol{\theta}$. Důkaz tohoto tvrzení je uveden v [1].

Vydatnost

Vydatný odhad (efficient estimator) je takový odhad $\widehat{\boldsymbol{\theta}}$, který má nejmenší rozptyl ze všech nestranných odhadů $\widehat{\boldsymbol{\theta}}_i$ parametru $\boldsymbol{\theta}$, tedy $\operatorname{Var} \widehat{\boldsymbol{\theta}} \leq \operatorname{Var} \widehat{\boldsymbol{\theta}}_i$, kde $i = 1, \ldots, l$, kde l je počet nestranných odhadů parametru $\boldsymbol{\theta}$ viz literatura [2].

Dostatečnost

Odhad je dostatečný, pokud obsahuje veškerou informaci o sledovaném parametru, kterou výběrový soubor poskytuje [2].

2. Metody bodového odhadu v parametrických modelech

2.1 Bodové odhady parametrů

Na základě hodnot náhodného výběru z rozdělení určitého typu odhadujeme parametry tohoto rozdělení, tak aby co nejlépe odpovídaly hodnotám náhodného výběru. Formulujeme tudíž úlohu, ve které hledáme odhady neznámých parametrů rozdělení.

Předpokládáme, že je dán náhodný výběr $X_1, \ldots, X_n$ z rozdělení s hustotou $f(x, \boldsymbol{\theta})$ nebo pravděpodobnostní funkcí $p(x, \boldsymbol{\theta})$, kde $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^T$ jsou parametry rozdělení. Na základě hodnot náhodného výběru odhadujeme parametry rozdělení. Odhad provádíme pomocí vhodně zvolené funkce náhodného výběru. O takové úloze mluvíme jako o bodovém odhadu parametrů.

2.2 Metody hledání bodových odhadů

Pro konstrukci bodových odhadů se nejčastěji používají dvě metody, a to metoda maximální věrohodnosti (maximum likelihood method) a momentová metoda (method of moments), jinak nazývaná metoda momentů.

U obou metod předpokládáme, že náhodné veličiny jsou spojité nebo diskrétní a budeme uvažovat parametrický model, kde máme náhodný výběr $\mathbf{X} = (X_1, \ldots, X_n)$ z rozdělení s hustotou $f(x; \boldsymbol{\theta})$ či pravděpodobnostní funkcí $p(x, \boldsymbol{\theta})$, kde tvar funkce $f(\cdot; \cdot)$ nebo $p(\cdot; \cdot)$ je známý a $\boldsymbol{\theta}$ je neznámý (vektorový) parametr, jenž leží v parametrickém prostoru $\Theta \subseteq \mathbb{R}^m$, $m \geq 1$. Pro spojité náhodné veličiny pracujeme tedy s modelem

$$\mathcal{F} = \{\text{rozdělení s hustotou } f(x; \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m\},$$

pro diskrétní náhodné veličiny s modelem

$$\mathcal{F} = \{\text{rozdělení s pravděpodobnostní funkcí } p(x; \boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m\}.$$

Při definici modelu jsme vycházeli z [4].

2.2.1 Momentová metoda

Metoda je založena na porovnání výběrových momentů s odpovídajícími teoretickými momenty předpokládaného rozdělení pravděpodobnosti. Využíváme faktu, že máme k dispozici konzistentní odhady momentů, což jsou výběrové momenty, a že momenty rozdělení X_i obvykle umíme vyjádřit jako funkce neznámých parametrů [4].

Metologie

Mějme náhodný výběr $\mathbf{X} = (X_1, \ldots, X_n)$ o rozsahu n z rozdělení ze třídy hustot $\{f(x;\boldsymbol{\theta}):\ \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m\}$, či třídy pravděpodobnostních funkcí $\{p(x;\boldsymbol{\theta}):\ \boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m\}$ a hledáme odhad parametru $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^T \in \Theta \subseteq \mathbb{R}^m$.

Pro odhad jednorozměrného parametru stačí předpokládat, že existuje prvních $m \geq 1$ teoretických počátečních momentů, které jsou pro spojitou náhodnou proměnnou definovány

$$\mu'_k = E\left(X^k\right),\ k = 1, \ldots, m.$$

Počáteční moment prvního řádu μ'_1 je roven střední hodnotě rozdělení EX. Pokud odhadujeme parametry porovnáním centrálních momentů, musíme přidat další předpoklad a to existenci prvních $m \geq 1$ teoretických centrálních momentů. Centrální moment k-tého řádu je definován

$$\mu_k = E\left(X - EX\right)^k,\ k = 1, \ldots, m.$$

Teoretický centrální moment druhého řádu je roven rozptylu rozdělení Var$X$. Je zřejmé, že momenty jsou funkcemi $\boldsymbol{\theta}$, píšeme proto $\mu'_k = \mu'_k(\boldsymbol{\theta})$ a $\mu_k = \mu_k(\boldsymbol{\theta})$.

Pro porovnání s teoretickými momenty potřebujeme výběrové momenty. Označme M'_k k-tý počáteční výběrový moment

$$M'_k = \frac{1}{n}\sum_{i=1}^{n}(X_i)^k,\ k = 1, \ldots, m.$$

První počáteční výběrový moment $M'_1 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i = \bar{X}_n$ nazýváme výběrový průměr někdy též aritmetický průměr. Výběrový centrální moment M_k k-tého řádu je definován

$$M_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^k,\ k = 1, \ldots, m.$$

Druhý výběrový centrální moment $M_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^k$ je označován jako empirický rozptyl. Momentovou metodou chceme odhadnout parametr $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^T$ tak, že budeme řešit soustavu rovnic

$$\begin{aligned} M'_k &= \mu'_k(\theta_1, \ldots, \theta_m),\ k = 1, \ldots, m, \\ M_k &= \mu_k(\theta_1, \ldots, \theta_m),\ k = 1, \ldots, m. \end{aligned} \tag{2.1}$$

Řešením soustavy $(\widehat{\theta}_1, \ldots, \widehat{\theta}_k)^T$ nazveme momentový odhad parametru $\boldsymbol{\theta}$, zřejmě $\widehat{\theta}_i = \widehat{\theta}_i(M'_1, \ldots, M'_m),\ i = 1, \ldots, m$. K odhadu je třeba získat tolik rovnic, kolik odhadujeme neznámých parametrů rozdělení. Nestačí-li m rovnic k jednoznačnému řešení soustavy (2.1), přidáme další rovnice pro $k = m+1, m+2, \ldots$

Výhodou momentové metody je její jednoduchost. Nevýhodou je často nízká eficience (velký rozptyl) výsledných odhadů. Tyto odhady se někdy používají např. jako počáteční aproximace pro jiné odhadovací metody.

2.2.2 Metoda maximální věrohodnosti

Metoda maximální věrohodnosti je obecná metoda pro konstrukci bodových odhadů parametrů. Používáme ji pro odhad parametrů, které určují rozdělení, např. střední hodnota a rozptyl normálního rozdělení, pravděpodobnost nastoupení jevu alternativního rozdělení apod. Princip metody maximální věrohodnosti spočívá v tom, že za odhad neznámého parametru vezmeme tu jeho hodnotu, ve které tzv. věrohodnostní funkce nabývá svého maxima viz literatura [6]. Vzorce pro metodu maximální věrohodnosti a postup výpočtu jsem čerpala z [4] a [7].

Metologie

Předpokládejme, že náhodný výběr byl pořízen z rozdělení, které je charakterizováno hustotou $f(x;\boldsymbol{\theta})$ nebo pravděpodobnostní funkcí $p(x;\boldsymbol{\theta})$. Sdružená hustota náhodného výběru $\mathbf{X} = (X_1, \ldots, X_n)$ je vzhledem k předpokládané nezávislosti dána předpisem

$$f(x_1, x_2, \ldots, x_n; \boldsymbol{\theta}) = f(\boldsymbol{x}; \boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta})$$

a sdružená pravděpodobnostní funkce je

$$p(x_1, x_2, \ldots, x_n; \boldsymbol{\theta}) = p(\boldsymbol{x}; \boldsymbol{\theta}) = \prod_{i=1}^{n} p(x_i; \boldsymbol{\theta}),$$

neboť složky náhodného výběru jsou nezávislé náhodné veličiny.

Pokud vyšetřujeme sdruženou hustotu $f(\boldsymbol{x};\boldsymbol{\theta})$ jako funkci $\boldsymbol{\theta}$ pro pevné x, nazýváme ji věrohodnostní funkcí a značíme ji symbolem $L(\boldsymbol{\theta})$, tedy

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(X_i; \boldsymbol{\theta}).$$

Analogicky v případě disktétního rozdělení náhodného výběru charakterizovaného sdruženou pravděpodobnostní funkcí $p(\boldsymbol{x};\boldsymbol{\theta})$, definujeme věrohodnostní funkci jako

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} p(X_i; \boldsymbol{\theta}).$$

Zpravidla je technicky výhodnější a při praktických výpočtech se místo věrohodnostní funkce pracuje s přirozeným logaritmem věrohodnostní funkce. Logaritmická věrohodnostní funkce je

$$l(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln f(X_i; \boldsymbol{\theta}),$$

$$l(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln p(X_i; \boldsymbol{\theta}),$$

přičemž tam, kde je $f(x) = 0$ nebo $p(x) = 0$, definujeme $l(\boldsymbol{\theta}) = -\infty$.

Věrohodnostní funkce nebo logaritmická věrohodnostní funkce je funkcí neznámého parametru $\boldsymbol{\theta}$, proto je nyní naším úkolem najít $\widehat{\boldsymbol{\theta}}$ tak, aby se maximalizovala hodnota věrohodnostní funkce. Vyjádřeno vzorcem dostáváme

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\theta\in\Theta} L(\boldsymbol{\theta}) = \arg\max_{\theta\in\Theta} l(\boldsymbol{\theta}).$$

Tento odhad nazýváme maximálně věrohodný odhad parametru $\boldsymbol{\theta}$ (maximum likelihood estimator). Tedy maximálně věrohodný odhad $\widehat{\boldsymbol{\theta}}$ parametru $\boldsymbol{\theta}$ je takový bod z Θ, který maximalizuje (přes všechny $\boldsymbol{\theta} \in \Theta_0$) sdruženou hustotu či sdruženou pravděpodobnostní funkci spočítanou v pozorovaných hodnotách $X_1, \ldots, X_n$.

Funkci $l(\boldsymbol{\theta})$ s proměnnou $\boldsymbol{\theta}$ se snažíme maximalizovat, a proto ji budeme derivovat podle $\boldsymbol{\theta}$. Tyto výrazy upravíme a položímě rovny nule a řešíme vzniklé rovnice. Obecně tedy řešíme soustavu rovnic, která je pro spojité náhodné veličiny

$$\sum_{i=1}^{n} \frac{\partial}{\partial\boldsymbol{\theta}} \ln f\left(X_i; \widehat{\boldsymbol{\theta}}\right) = \mathbf{0}$$

analogicky pro diskrétní náhodné veličiny

$$\sum_{i=1}^{n} \frac{\partial}{\partial\boldsymbol{\theta}} \ln p\left(X_i; \widehat{\boldsymbol{\theta}}\right) = \mathbf{0}.$$

Soustava se derivuje podle jednotlivých složek $\boldsymbol{\theta}$, protože $\boldsymbol{\theta} = (\theta_1, , \theta_m)^T$. Řešení této soustavy se obvykle hledá numericky. Řešení však nemusí existovat a ne každé řešení je maximálně věrohodný odhad.

3. Bodové odhady pro různá spojitá rozdělení

V této části bakalářské práce se zaměříme na některá spojitá pravděpodobnostní rozdělení a odhadneme jejich parametry. Nejprve momentovou metodou a následně metodou maximální věhodnosti, pokud tak lze učinit. Metody si nejdříve ukážeme na normálním rozdělení, u kterého lze použít k odhadu jeho parametrů obě metody. Rozepíšeme podrobně postup výpočtu odhadů, který už u dalších rozdělení budeme zkracovat.

3.1 Normální rozdělení

3.1.1 Momentová metoda

Je dán výběr $X_1, \ldots, X_n$ pocházející z normálního rozdělení. Momentovou metodou odhadneme parametry μ a σ^2. Hustota normálního rozdělení je tvaru

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \text{ kde } \mu \in \mathbb{R},\ \sigma^2 > 0.$$

U tohoto rozdělení víme, že parametr μ je střední hodnota, tj. počáteční (teoretický) moment prvního řádu a parametr σ^2 značí rozptyl, tedy teoretický centrální moment druhého řádu. Vyjádřeno matematicky počáteční moment prvního řádu normálního rozdělení je

$$\mu_1'\left(\mu, \sigma^2\right) = \mu$$

a centrální moment druhého řádu normálního rozdělení je

$$\mu_2\left(\mu, \sigma^2\right) = \sigma^2$$

Z druhé kapitoly víme, že první počáteční výběrový moment

$$M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$$

a druhý centrální výběrový moment

$$M_2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2.$$

Řešíme tedy soustavu rovnic

$$\begin{aligned} M_1' &= \mu_1'\left(\mu, \sigma^2\right) \\ M_2 &= \mu_2\left(\mu, \sigma^2\right). \end{aligned}$$

Po dosazení dostáváme

$$\begin{aligned} \bar{X}_n &= \mu, \\ \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 &= \sigma^2. \end{aligned}$$

Odhad parametru $\boldsymbol{\theta} = (\mu, \sigma^2)$ momentovou metodou je

$$\begin{aligned} \widehat{\mu} &= \bar{X}_n, \\ \widehat{\sigma^2} &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2. \end{aligned}$$

3.1.2 Metoda maximální věrohodnosti

Uvažujme opět náhodný výběr o rozsahu n, tedy máme nezávislé, stejně rozdělené náhodné veličiny pocházející z normálního rozdělení N (μ, σ^2), v němž budeme odhadovat oba neznámé parametry μ a σ^2. Hustota normálního rozdělení je

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \text{ kde } \mu \in \mathbb{R},\ \sigma^2 > 0.$$

Sdružená hustota je ve tvaru

$$f(x_1, x_2, \ldots, x_n; \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma}\exp -\frac{(x_i-\mu)^2}{2\sigma^2}.$$

Věrohodnostní funkce je definována

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} f(X_i, \mu, \sigma^2),$$

po dosazení

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \tfrac{1}{\sqrt{2\pi}\sigma}\exp\left[-\tfrac{(X_i-\mu)^2}{2\sigma^2}\right].$$

Logaritmická věrohodnostní funkce je definována jako

$$l(\mu, \sigma^2) = \ln L(\mu, \sigma^2),$$

po dosazení dostáváme

$$l(\mu, \sigma^2) = \ln\left(\prod_{i=1}^{n} \tfrac{1}{\sqrt{2\pi}\sigma}\exp\left[-\tfrac{(X_i-\mu)^2}{2\sigma^2}\right]\right).$$

Z vlastností logaritmů víme, že logaritmus součinu je součet logaritmů jednotlivých činitelů, upravíme si logaritmickou věrohodnostní funkci na co nejjednodušší tvar tedy

$$\begin{aligned} l\left(\mu,\sigma^2\right) &= \sum_{i=1}^{n}\left(\ln\left(2\pi\sigma^2\right)^{-\frac{1}{2}} - \frac{\left(X_i-\mu\right)^2}{2\sigma^2}\right) = \\ &= -\frac{n}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(X_i-\mu\right)^2 = \\ &= -\frac{n}{2}\ln\sigma^2 - n\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(X_i-\mu\right)^2. \end{aligned}$$

Tuto věrohodnostní funkci s proměnnými μ a σ^2 chceme maximalizovat, proto ji budeme derivovat nejprve podle μ a poté podle proměnné σ^2. Je potřeba si uvědomit, že druhá parciální derivace je podle σ^2. Derivováním dostáváme

$$\begin{aligned} \frac{\partial l\left(\mu,\sigma^2\right)}{\partial\mu} &= -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(-2)\left(X_i-\mu\right) = \frac{1}{\sigma^2}\sum_{i=1}^{n}\left(X_i-\mu\right) \\ \frac{\partial l\left(\mu,\sigma^2\right)}{\partial\sigma^2} &= -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(X_i-\mu\right)^2. \end{aligned}$$

Obě upravené derivace si položíme rovno nule a vyjádříme si z nich odhady parametrů μ a σ^2. Z první rovnice úpravami získáme

$$\begin{aligned} \frac{1}{\widehat{\sigma}^2}\sum_{i=1}^{n}\left(X_i-\widehat{\mu}\right) &= 0 \\ \sum_{i=1}^{n}X_i - n\widehat{\mu} &= 0 \\ \widehat{\mu} &= \frac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}_n. \end{aligned} \tag{3.1}$$

Z druhé rovnice úpravami a dosazením odhadu (3.1) získáme

$$\begin{aligned} -\frac{n}{2}\frac{1}{\widehat{\sigma}^2} + \frac{1}{2\widehat{\sigma}^4}\sum_{i=1}^{n}\left(X_i-\widehat{\mu}\right)^2 &= 0 \\ \sum_{i=1}^{n}\left(X_i-\widehat{\mu}\right)^2 &= \frac{n}{2\widehat{\sigma}^2}\frac{2\widehat{\sigma}^4}{1} \\ \sum_{i=1}^{n}\left(X_i-\widehat{\mu}\right)^2 &= n\widehat{\sigma}^2 \\ \widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2. \end{aligned}$$

Metodou maximální věrohodnoti dostáváme odhady parametrů μ a σ^2.

$$\begin{aligned}\widehat{\mu} &= \bar{X}_n,\\ \widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2.\end{aligned}$$

Oběma metodami jsme dostali stejné odhady. Nyní se zaměříme na vlastnosti těchto odhadů.

3.1.3 Vlastnosti odhadů

Z první kapitoly víme, že bodové odhady mají své vlastnosti, v této práci budeme věnovat pozornost zejména nestrannosti a konzistenci. Tyto vlastnosti si nejdříve ověříme u odhadu střední hodnoty, který je roven $\widehat{\mu} = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ a nazývá se výběrový průměr.

Nestrannost

Při zjišťování této vlastnosti, spočítáme střední hodnotu odhadu $\widehat{\mu}$, a tím ověříme, zda je odhad nestranný či nikoliv. Využijeme přitom vlastnost střední hodnoty $\mathrm{E}(a + bX) = a + b\mathrm{E}X$, kde $\forall\ a,\ b \in \mathbb{R}$.

$$\mathrm{E}\bar{X}_n = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(X_i\right) = \frac{1}{n}n\mu = \mu,$$

platí, že střední hodnota odhadu se rovná střední hodnotě rozdělení a tedy výběrový průměr je nestranným odhadem střední hodnoty.

Konzistence

Tato vlastnost se ověřuje pomocí zákona velkých čísel a to nejlépe pomocí Chinčinova slabého zákona velkých čísel, který definujme v následující větě. Nechť máme $X_1, \ldots, X_n$ posloupnost nezávislých stejně rozdělených náhodných veličin s $\mathrm{E}X_i = c < \infty$, kde c je konstanta, potom $\bar{X}_n \xrightarrow{P} c$.

V našem případě máme náhodný výběr z normálního rozdělení, což zaručuje stejně rozdělené náhodné veličiny, jehož prvky mají střední hodnotu rovno nějaké konstantě μ, tím máme splněny předpoklady Chinčinova slabého zákona velkých čísel. Tedy výběrový průměr je konzistentním odhadem střední hodnoty.

Dostatečnost a vydatnost

Výběrový průměr je dostatečným i vydatným odhadem střední hodnoty, viz literatura [2].

Následně určeme vlastnosti empirického odhadu rozptylu, který nám vyšel

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2.$$

Nestrannost

Výpočet provádíme za pomoci vlastností střední hodnoty, které jsme využili při ověřování nestrannosti výběrového průměru a dále faktu, že se jedná o nezávislé náhodné veličiny.

$$\begin{aligned}
\mathrm{E}\widehat{\sigma}^2 &= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2\right) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - 2X_i\bar{X}_n + \bar{X}_n^2\right)\right) = \\
&= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\frac{1}{n}\sum_{i=1}^{n}X_i\bar{X}_n + \frac{1}{n}\sum_{i=1}^{n}\bar{X}_n^2\right) = \\
&= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X}_n^2 + \frac{n}{n}\bar{X}_n^2\right) = \\
&= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}_n^2\right) = \mathrm{E}\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \mathrm{E}\bar{X}_n^2 = \\
&= \mathrm{E}X_i^2 - \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^{n}X_i\right)^2 = \mathrm{E}X_i^2 - \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}X_iX_j\right) \overset{\text{nezávislost}}{=} \\
&= \mathrm{E}X_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathrm{E}X_i^2 + n\left(n-1\right)\mu^2\right) = \\
&= \mathrm{E}X_i^2 - \frac{1}{n}\mathrm{E}X_i^2 - \frac{n-1}{n}\mu^2 = \\
&= \frac{n-1}{n}\mathrm{E}X_i^2 - \frac{n-1}{n}\mu^2 = \frac{n-1}{n}\sigma^2.
\end{aligned}$$

Vidíme, že $\mathrm{E}\widehat{\sigma}^2 \neq \sigma^2$, tedy empirický výběrový rozptyl není nestranným odhadem rozptylu. Další způsob ověření této vlastnosti je za pomoci triku (přičtení a odečtení μ).

$$\begin{aligned}
\mathrm{E}\widehat{\sigma}^2 &= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2\right) = \\
&= \mathrm{E}\left(\frac{1}{n}\left(\sum_{i=1}^{n}\left(X_i - \mu\right)^2 - n\left(\bar{X}_n - \mu\right)^2\right)\right) = \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\left(X_i - \mu\right)^2 - \frac{n}{n}\mathrm{E}\left(\bar{X}_n - \mu\right)^2 = \\
&= \frac{1}{n}n\sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2,
\end{aligned}$$

i tímto způsobem výpočtu jsme zjistili, že empirický výběrový rozptyl není nestranným odhadem parametru σ^2.

Ověříme ještě asymptotickou nestrannost

$$\lim_{n\to\infty} \mathrm{E}\widehat{\sigma^2} = \lim_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2.$$

a docházíme k závěru, že odhad rozptylu σ^2 je asymptoticky nestranný.

Konzistence

Jednoduchých výpočtem ověříme, že empirický odhad rozptylu je konzistentním odhadem

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}_n^2 \xrightarrow{P} \sigma^2.$$

3.2 Gama rozdělení

3.2.1 Momentová metoda

Je dán výběr $X_1, \ldots, X_n$ pocházející z gama rozdělení, jak již víme z první kapitoly gama rozdělení $\mathrm{Ga}(a,p)$, které má parametry $a > 0$, $p > 0$ má hustotu danou vzorcem

$$f(x) = \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1}, \; x > 0, \tag{3.2}$$

střední hodnota je $\mu_1' = \frac{p}{a}$ a rozptyl $\sigma^2 = \frac{p}{a^2}$ z [1].

Výpočet momentových odhadů provedeme v následujících krocích, kde nebude porovnávat druhý centrální moment s empirickým rozptylem, ale druhý centrální moment s výběrovým rozptylem $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2$, nazýván vypočtený rozptyl. Činíme z důvodu lepších výsledků. Tedy pro první dva počáteční momenty tohoto rozdělení platí

$$\begin{aligned} \mu_1'(a,p) &= \frac{p}{a}, \\ \mu_2(a,p) &= \frac{p}{a^2}. \end{aligned}$$

První počáteční výběrový moment a druhý centrální výběrový moment jsou tvaru

$$\begin{aligned} M_1' &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n, \\ M_2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^2 = S_n^2. \end{aligned} \tag{3.3}$$

Řešíme soustavu rovnic

$$\begin{aligned} \bar{X}_n &= \frac{p}{a}, \\ S_n^2 &= \frac{p}{a^2}. \end{aligned}$$

Nejdříve si spočítáme odhad parametru a

$$\frac{\bar{X}_n}{S_n^2} = \frac{\frac{\widehat{p}}{\widehat{a}}}{\frac{\widehat{p}}{\widehat{a}^2}} = \widehat{a},$$

následně odhad parametru p

$$\widehat{p} = \bar{X}_n \widehat{a} = \overline{X} \frac{\bar{X}_n}{S_n^2} = \frac{\bar{X}_n^2}{S_n^2}.$$

Metodou momentů dostáváme pro parametr $\theta = (a, p)$ odhady

$$\begin{aligned} \widehat{a} &= \frac{\bar{X}_n}{S_n^2} \\ \widehat{p} &= \frac{\bar{X}_n^2}{S_n^2}. \end{aligned}$$

3.2.2 Metoda maximální věrohodnosti

Odhadneme parametry gama rozdělení metodou maximální věrohodnosti. Hustota gama rozdělení je již uvedena v (3.2). Rovnou píšeme sdruženou hustotu

$$f(x_1, x_2, \dots x_n; a, p) = \prod_{i=1}^{n} \frac{a^p}{\Gamma(p)} e^{-ax_i} x_i^{p-1}.$$

Maximálně věrohodné odhady parametrů a a p dostaneme pomocí věrohodnostní funkce

$$L(a, p) = \prod_{i=1}^{n} \frac{a^p}{\Gamma(p)} e^{-aX_i} X_i^{p-1}.$$

Pro její logaritmus dostaneme vyjádření

$$\begin{aligned} l(a, p) &= \ln \left(\prod_{i=1}^{n} \frac{a^p}{\Gamma(p)} e^{-aX_i} X_i^{p-1} \right) = \\ &= \sum_{i=1}^{n} (p \ln a - \ln \Gamma(p) + (p-1) \ln X_i - aX_i). \end{aligned}$$

Derivací dostáváme

$$\begin{aligned} \frac{\partial l(a, p)}{\partial a} &= \sum_{i=1}^{n} \left(\frac{p}{a} - X_i \right) = n \frac{p}{a} - \sum_{i=1}^{n} X_i \\ \frac{\partial l(a, p)}{\partial p} &= \sum_{i=1}^{n} \left(\ln a - \frac{\Gamma'(p)}{\Gamma(p)} + \ln X_i \right) \\ &= n(\ln a - \psi(p)) + \sum_{i=1}^{n} \ln X_i, \end{aligned}$$

kde $\psi(p)$ nazýváme digama funkcí. První derivace položíme rovnu nule a vyjádříme si z ní odhad parametru a

$$\begin{aligned}
n\frac{\widehat{p}}{\widehat{a}} - \sum_{i=1}^{n} X_i &= 0 \\
\frac{\widehat{p}}{\widehat{a}} &= \frac{1}{n}\sum_{i=1}^{n} X_i \\
\widehat{a} &= \frac{\widehat{p}}{\bar{X}_n}.
\end{aligned} \tag{3.4}$$

Derivaci podle parametru p taktéž položíme rovnu nule a upravíme si ji

$$\begin{aligned}
n\left(\ln\widehat{a} - \psi\left(\widehat{p}\right)\right) + \sum_{i=1}^{n} \ln X_i &= 0 \\
-\ln\widehat{a} + \psi\left(\widehat{p}\right) &= \frac{1}{n}\sum_{i=1}^{n} \ln X_i
\end{aligned} \tag{3.5}$$

Nyní si rovnici (3.4) zlogaritmujeme. Dostáváme tak

$$-\ln\widehat{a} = \ln\bar{X}_n - \ln\widehat{p} \tag{3.6}$$

Rovnost (3.6) dosadíme do (3.5), dostáme

$$\ln\bar{X}_n - \ln\widehat{p} + \psi\left(p\right) = \frac{1}{n}\sum_{i=1}^{n} \ln X_i.$$

Na jednu stranu si dáme prvky, které obsahují odhadnutý parametr $\widehat{p}$ a rovnost si upravíme

$$\begin{aligned}
\ln\widehat{p} - \psi\left(\widehat{p}\right) &= \ln\bar{X}_n - \frac{1}{n}\sum_{i=1}^{n} \ln X_i = \\
&= \ln\bar{X}_n - \frac{1}{n}\ln\left(\prod_{i=1}^{n} X_i\right) = \\
&= \ln\frac{\bar{X}_n}{\sqrt[n]{\prod_{i=1}^{n} X_i}}.
\end{aligned}$$

Pokud bychom chtěli vědět přesnou hodnotu odhadu $\widehat{p}$, tak bychom využili nějaký program, který by odhad dopočítal numericky.

Vlastnosti odhadů gama rozdělení

Každou metodou nám vyšly jiné odhady. U metody maximální věrohodnosti vyšly odhady, které se musí softwarem dopočítat, proto ověříme jaké vlastnosti mají odhady odhadnuté pomocí metody momentů. Nejprve sepíšeme jaké vlastnosti má odhad

$$\widehat{a} = \frac{\bar{X}_n}{S_n^2}.$$

Nestrannost

Odhad $\widehat{a}$ je nestranným odhadem parametru a.

Konzistence

Z vlastností výběrového průměru, výběrového rozptylu a se splněním podmínek zákona velkých čísel dostáváme $\widehat{a} \xrightarrow{P} a$, jedná se o konzistentní odhad.

Nestrannost

Odhad $\widehat{p}$ není nestranný. Ověříme tedy asymptotickou nestrannost

$$\lim_{n\to\infty} \mathrm{E}\widehat{p} = \lim_{n\to\infty} \frac{1+np}{n} = p.$$

Tedy tento odhad je asymptoticky nestranný.

Konzistence

Platí $\widehat{p} \xrightarrow{P} p$ a tedy tento odhad je konzistentním odhadem parametru p.

3.3 Logaritmicko-normální rozdělení

Momentová metoda

Logaritmicko-normální rozdělení je příkladem, kdy se nepoužívá k výpočtu parametrů momentová metoda. Toto rozdělení nemá omezené momenty a tedy není zaručen jednoznačně vzájemný vztah mezi posloupností momentů a hustotou rozdělení. Formálním výpočtem rovnic pro počáteční teoretické a výběrové momenty dostaneme vzorce pro parametry, jejich řešení ale nemusí odpovídat skutečnosti.

Logaritmicko-normální rozdělení $\mathrm{LN}(\mu, \sigma^2)$ se dvěma parametry $\mu \in \mathbb{R}$ a $\sigma^2 > 0$, které má náhodná veličina X jejíž transformace $\ln X$ má normální rozdělení $\mathrm{N}(\mu, \sigma^2)$. Pro její hustotu dostaneme vyjádření ve tvaru

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \text{ pro } x > 0.$$

Počáteční moment prvního řádu a centrální moment druhého řádu tohoto rozdělení je

$$\begin{aligned} \mu_1'\left(\mu, \sigma^2\right) &= \exp\left[\mu + \frac{\sigma^2}{2}\right], \\ \mu_2\left(\mu, \sigma^2\right) &= \left(\exp\left[\sigma^2\right] - 1\right) \exp\left[2\mu + \sigma^2\right]. \end{aligned}$$

Výběrový průměr a empirický rozptyl jsou

$$\begin{aligned} M_1' &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n, \\ M_2 &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2. \end{aligned}$$

Dostáváme na porovnáním rovnice, ve kterých pro zjednodušení a zkrácení vzorců nebudeme vyjadřovat hodnoty M_1' a M_2

$$\begin{aligned} \bar{X}_n &= \exp\left[\mu + \frac{\sigma^2}{2}\right] = M_1', \\ \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2 &= \left(\exp\left[\sigma^2\right] - 1\right)\exp\left[2\mu + \sigma^2\right] = M_2. \qquad (3.7) \end{aligned}$$

První rovnici zlogaritmuji a vyjádřím si z ní μ

$$\begin{aligned} \ln M_1' &= \mu + \frac{\sigma^2}{2} \\ \mu &= \ln M_1' - \frac{\sigma^2}{2}. \qquad (3.8) \end{aligned}$$

Druhou rovnici ze soustavy (3.7) si také zlogaritmuji a dosadim za $2\mu + \sigma^2$ výraz $2\ln M_1'$ dostáváme tak

$$\begin{aligned} \ln M_2 &= \ln\left(\exp\left[\sigma^2\right] - 1\right) + 2\mu + \sigma^2 \\ \ln M_2 &= \ln\left(\exp\left[\sigma^2\right] - 1\right) + 2\ln M_1' \qquad (3.9) \end{aligned}$$

postupně si z (3.9) vyjádřím σ^2

$$\begin{aligned} \exp\left[\ln M_2\right]\exp\left[-2\ln M_1'\right] &= \exp\left[\sigma^2\right] - 1 \\ \frac{M_2}{(M_1')^2} + 1 &= \exp\left[\sigma^2\right] \\ \sigma^2 &= \ln\left(\frac{M_2}{(M_1')^2} + 1\right) \qquad (3.10) \end{aligned}$$

Pro vyjádření odhadu parametru μ dosadím (3.10) do první zlogaritmované rovnice (3.8) čili

$$\begin{aligned} \ln M_1' &= \mu\frac{1}{2}\ln\left(\frac{M_2}{(M_1')^2} + 1\right) \\ \mu &= \ln M_1' - \frac{1}{2}\ln\left(\frac{M_2}{(M_1')^2} + 1\right) \end{aligned}$$

Pro náhodný výběr o rozsahu n, jehož náhodné veličiny mají logaritmicko-normální rozdělení dostáváme metodou momentů odhady pro jeho parametry

$$\widehat{\mu} = \ln\left(M_1'\right) - \frac{1}{2}\ln\left(1 + \frac{M_2}{(M_1')^2}\right)$$

$$\widehat{\sigma^2} = \ln\left(1 + \frac{M_2}{M_1'}\right).$$

Metoda maximální věrohodnosti

Nyní máme za úkol určit odhad parametrů náhodné veličiny X, která má logaritmicko-normální rozdělení s neznámými parametry $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ pomocí metody maximální věrohodnosti. Rovnou napíšeme sdruženou hustotu

$$f(x_1, x_2, \ldots, x_n; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma x_i} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right].$$

Věrohodnostní funkce

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \tfrac{1}{\sqrt{2\pi}\sigma X_i} \exp\left[-\tfrac{(\ln X_i - \mu)^2}{2\sigma^2}\right].$$

Dostáváme se k logaritmické věrohodnostní funkci, kterou máme definovanou $l(\mu, \sigma^2) = \ln L(\mu, \sigma^2)$

$$l(\mu, \sigma^2) = \ln\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma X_i} \exp\left[-\frac{(\ln X_i - \mu)^2}{2\sigma^2}\right]\right),$$

postupně upravujeme

$$\begin{aligned}
l(\mu, \sigma^2) &= \sum_{i=1}^{n}\left(\ln\frac{1}{\sqrt{2\pi}\sigma X_i} + \ln\left(\exp\left[\frac{(\ln X_i - \mu)^2}{2\sigma^2}\right]\right)\right) = \\
&= \sum_{i=1}^{n}\left(-\ln\sqrt{2\pi} - \ln\sigma - \ln X_i - \frac{(\ln X_i - \mu)^2}{2\sigma^2}\right) = \\
&= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \sum_{i=1}^{n}\ln X_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln X_i - \mu)^2.
\end{aligned}$$

Soustavu věrohodnostních rovnic derivujeme podle parametrů μ a σ^2

$$\begin{aligned}
\frac{\partial l(\mu, \sigma^2)}{\partial \mu} &= -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(\ln X_i - \mu)(-1) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln X_i - \mu) \\
\frac{\partial l(\mu, \sigma^2)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(\ln X_i - \mu)^2.
\end{aligned}$$

Derivace položíme rovno nule a vyjádříme si z nich odhady parametrů μ a σ^2

$$\begin{aligned}
\frac{1}{\widehat{\sigma}^2}\sum_{i=1}^{n}(\ln X_i - \widehat{\mu}) &= 0 \\
\sum_{i=1}^{n}\ln X_i - n\widehat{\mu} &= 0 \\
\widehat{\mu} &= \frac{1}{n}\sum_{i=1}^{n}\ln X_i
\end{aligned}$$

a pro odhad σ^2 dostáváme

$$\begin{aligned}
-\frac{n}{2\widehat{\sigma}^2} + \frac{1}{2\widehat{\sigma}^4}\sum_{i=1}^{n}(\ln X_i - \widehat{\mu})^2 &= 0 \\
\sum_{i=1}^{n}(\ln X_i - \widehat{\mu})^2 &= \frac{n}{2\widehat{\sigma}^2}\frac{2\widehat{\sigma}^4}{1} \\
\sum_{i=1}^{n}(\ln X_i - \widehat{\mu})^2 &= n\widehat{\sigma}^2 \\
\widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}(\ln X_i - \widehat{\mu})^2.
\end{aligned}$$

Metodou maximální věrohodnosti jsme dostali tyto odhady

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln X_i$$

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(\ln X_i - \widehat{\mu})^2.$$

Vlastnosti odhadů vypočítané metodou maximální věrohodnosti

Ověříme jaké vlastnosti má odhad $\widehat{\mu}$ spočtený metodou maximální věrohodnosti.

Nestrannost

$$\mathrm{E}\,(\widehat{\mu}) = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\ln X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}\,(\ln X_i) = \frac{1}{n}n\mu = \mu,$$

tedy odhad $\widehat{\mu}$ je nestranným odhadem parametru μ.

Konzistence

Tuto vlastnost ověřujeme za pomoci Chinčinova slabého zákona velkých čísel. Jeho předpoklady máme splněny, jelikož máme náhodný výběr z rozdělení, které má střední hodnotu rovnou nějaké konstantě a tedy odhad $\widehat{\mu}$ je konzistentním odhadem parametru μ.

3.4 Exponenciální rozdělení

3.4.1 Momentová metoda

Exponenciální rozdělení $\mathrm{Ex}(\lambda)$, kde $\lambda > 0$ má hustotu definovanou

$$f(x) = \frac{1}{\lambda}\exp\left[-\frac{x}{\lambda}\right],\ x > 0,$$

střední hodnota $\mu' = \lambda$. Metodou momentů budeme odhadovat parametr λ, a tudíž odhad parametru λ získáme z porovnání $M_1' = \mu_1'$, kde $M_1' = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$, tedy odhadem parametru λ je hodnota

$$\widehat{\lambda} = \bar{X}_n.$$

3.4.2 Metoda maximální věrohodnosti

Věrohodnostní funkce je rovna

$$L(\lambda) = \prod_{i=1}^{n} \frac{1}{\lambda}\exp\left[-\frac{X_i}{\lambda}\right]$$

a její logaritmická věrohodnostní funkce je rovna

$$\begin{aligned} l(\lambda) &= \ln\left(\prod_{i=1}^{n} \frac{1}{\lambda}\exp\left[\frac{-X_i}{\lambda}\right]\right) \\ &= \sum_{i=1}^{n}\left(-\ln\lambda - \frac{1}{\lambda}X_i\right) = -n\ln\lambda - \frac{1}{\lambda}\sum_{i=1}^{n} X_i. \end{aligned}$$

Odtud dostáváme věrohodnostní rovnici

$$\frac{\partial l(\lambda)}{\partial\lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\sum_{i=1}^{n} X_i.$$

Věrohodnostní rovnici položíme rovnu nule a úpravami dostáváme maximálně věrohodný odhad parametru λ

$$\begin{aligned} -\frac{n}{\widehat{\lambda}} + \frac{1}{\widehat{\lambda}^2}\sum_{i=1}^{n} X_i &= 0 \\ \frac{n}{\widehat{\lambda}} &= \frac{1}{\widehat{\lambda}^2}\sum_{i=1}^{n} X_i \\ \widehat{\lambda} &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n. \end{aligned}$$

3.4.3 Vlastnosti odhadů

Oběma metodami nám vyšly stejné odhady. Odhad střední hodnoty je výběrový průměr, jehož vlastnosti jsme již ověřili v části (3.1.3) a víme, že tento odhad je nestranný a konzistentní.

3.5 Rovnoměrné rozdělení

3.5.1 Momentová metoda

U této motedy si můžeme zvolit jaké teoretické momenty budeme porovnávat s výběrovými momenty. Tedy u tohoto rozdělení nebude porovnávat druhý centrální moment s empirickým rozptylem, ale využijeme druhý výběrový moment M_2', který nám usnadní výpočet. Momentová metoda

Rovnoměrné rozdělení je definované na intervalu $(\mu - h, h + \mu)$ a má hustotu danou výrazem

$$f(x) = \frac{1}{2h}, \ \mu - h < x < \mu + h.$$

První a druhý teoretický počáteční moment je

$$\begin{aligned} \mu_1'(\mu, h) &= \mu, \\ \mu_2'(\mu, h) &= \mu^2 + \frac{h^2}{3}. \end{aligned}$$

První a druhý výběrový moment je

$$\begin{aligned} M_1' &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n \\ M_2' &= \frac{1}{n}\sum_{i=1}^{n} X_i^2. \end{aligned}$$

Porovnáním prvního momentu s výběrovým průměrem a druhého momentu s druhým výběrovým momentem získáme odhady parametrů μ a h. Budeme tedy vycházet z následující soustavy rovnic

$$\begin{aligned} M_1' &= \mu_1'(\mu, h) \\ M_2' &= \mu_2'(\mu, h). \end{aligned}$$

Dosadíme a dostáváme odhad parametru μ

$$\widehat{\mu} = \bar{X}_n,$$

a s využitím odhadu $\widehat{\mu}$ vypočítáme odhad pro parametr h

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} X_i^2 &= \widehat{\mu}^2 + \frac{\widehat{h}^2}{3} \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 &= \left(\bar{X}_n\right)^2 + \frac{\widehat{h}^2}{3} \\ \widehat{h}^2 &= 3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\bar{X}_n\right)^2\right) \\ \widehat{h} &= \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\bar{X}_n\right)^2\right)}. \end{aligned}$$

3.5.2 Metoda maximální věrohodnosti

Věrohodnostní funkce pro náhodný výběr z rovnoměrného rozdělení je rovna

$$L(\mu, h) = \prod_{i=1}^{n} \frac{1}{2h} = \frac{1}{(2h^n)}$$

její logaritmus je roven

$$\begin{aligned} l(\mu, h) &= \ln \frac{1}{(2h)^n} \\ &= -n \ln (2h). \end{aligned}$$

Tato funkce má maximální hodnotu pro minimální hodnotu parametru h. Maximálně věrohodný odhad parametru h je roven

$$\widehat{h} = \frac{1}{2} \left(\max \{X_i;\ 1 \le i \le n\} - \min \{X_i;\ 1 \le i \le n\} \right).$$

Pro maximálně věrohodný odhad parametru μ dostáváme

$$\widehat{\mu} = \frac{1}{2} \left(\max \{X_i;\ 1 \le i \le n\} + \min \{X_i;\ 1 \le i \le n\} \right),$$

což jsou hodnoty odlišné od odhadů, které jsme dostali metodou momentů.

3.5.3 Vlastnosti odhadů

U odhadu spočtené metodou maximální věrohodnosti jsme obdrželi $\widehat{\mu}$, který je nestranný odhad a odhadu parametru h ověříme nestrannost

$$\begin{aligned} \mathrm{E}\left(\widehat{h}\right) &= \mathrm{E}\left(\sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\bar{X}_n\right)^2\right)}\right) = \sqrt{3}\sqrt{\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\bar{X}_n\right)^2\right)} \\ &= \sqrt{3}\sqrt{\frac{n-1}{n}\sigma^2} = \sqrt{3}\sqrt{\lambda^2} = \sqrt{3\lambda^2}. \end{aligned}$$

Tedy odhad $\widehat{h}$ není nestranný, ale je asymptoticky nestranný a oba odhady jsou konzistentní viz [3].

4. Bodové odhady pro různá diskrétní rozdělení

Tato kapitola se věnuje odhadům parametrů diskrétních pravděpodobnostních rozdělení. Podobně jako ve třetí kapitole odhadneme parametry momentovou metodou a metodou maximální věhodnosti.

4.1 Geometrické rozdělení

4.1.1 Momentová metoda

Náhodná veličina X má geometrické rozdělení definované pravděpodobnostní funkci

$$P\left(X=x\right)=p\left(1-p\right)^{x-1},\ x\in\mathbb{N}.$$

Pro základní číselné charakteristiky platí, že střední hodnota je $\mu'=\frac{1}{p}$ a rozptyl $\sigma^2=\frac{1}{p}\left(\frac{1}{p}-1\right)$. Je-li $X_1,X_2,\ldots,X_n$ náhodný výběr z geometrického rozdělení, tak metodou momentů porovnáváme pouze výběrový průměr a střední hodnotu geometrického rozdělení, protože pro geometrické rozdělení odhadujeme jeden parametr p. Matematicky vyjádřeno

$$M_1'=\mu_1'$$

dosadíme za $\mu_1'=\frac{1}{p}$ a za $M_1'=\frac{1}{n}\sum_{i=1}^{n}X_i=\bar{X}_n$. Tedy odhad parametru p je

$$\widehat{p}=\frac{1}{\bar{X}_n}.$$

4.1.2 Metoda maximální věrohodnosti

Maximálně věrohodný odhad spočítáme z věrohodnostní funkce, která je

$$L\left(p\right)=\prod_{i=1}^{n}p\left(1-p\right)^{X_i-1},\ \text{kde } X_i\in\mathbb{N}.$$

Logaritimická věrohodnostní funkce se rovná

$$\begin{aligned}
l\left(p\right) &= \ln\left(\prod_{i=1}^{n}p\left(1-p\right)^{X_i-1}\right)=\sum_{i=1}^{n}\left(\ln p+\ln\left(1-p\right)^{X_i-1}\right)\\
&= \sum_{i=1}^{n}\left(\ln p+\left(X_i-1\right)\ln\left(1-p\right)\right)=n\ln p+\left(\sum_{i=1}^{n}X_i-n\right)\ln\left(1-p\right),
\end{aligned}$$

tudíž věrohodnostní rovnice

$$\frac{\partial l(p)}{\partial p} = \frac{n}{p} - \left(\sum_{i=1}^{n} X_i - n\right)\frac{1}{1-p}.$$

Věrohodnostní rovnici položíme rovnu nule, abychom mohli najít maximálně věrohodný odhad parametru

$$\begin{aligned}
\frac{n}{\widehat{p}} - \left(\sum_{i=1}^{n} X_i - n\right)\frac{1}{1-\widehat{p}} &= 0\\
\frac{\sum_{i=1}^{n} X_i - n}{1-\widehat{p}} &= \frac{n}{\widehat{p}}\\
\widehat{p}\left(\sum_{i=1}^{n} X_i - n\right) &= n - n\widehat{p}\\
\widehat{p} &= \frac{n}{\sum_{i=1}^{n} X_i} = \frac{n}{\frac{n}{n}\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}_n}.
\end{aligned} \tag{4.2}$$

4.1.3 Vlastnosti odhadů

Metodou momentů i metodou maximální věrohodnosti nám vyšel stejný odhad parametru p a to převrácená hodnota výběrového průměru. Tento odhad je nestranný a konzistentním odhadem střední hodnoty viz literatura [1].

4.2 Poissonovo rozdělení

4.2.1 Momentová metoda

Předpokládejme, že veličina X nabývá pouze hodnot $0, 1, \ldots$, a to s pravděpodobnostmi

$$P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda},\ x = 0, 1, \ldots,$$

kde $\lambda > 0$ je dané číslo. Pak říkáme, že X má Poissonovo rozdělení Po(λ) s parametrem λ. Platí, že střední hodnota je $\mu' = \lambda$. Tedy pro odhad parametru λ porovnáme střední hdonotu s výběrovým průměrem $\mu_1' = M_1'$, po dosazení odhadem parametru λ je odhad $\widehat{\lambda} = \bar{X}_n$.

4.2.2 Metoda maximální věrohodnosti

Věrohodnostní funkce je rovna

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i}}{X_i!}e^{-\lambda}$$

a její logaritmus je roven

$$\begin{aligned} l(\lambda) &= \ln\left(\prod_{i=1}^{n}\frac{\lambda^{X_i}}{X_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\bigl(\ln\lambda_i^X - \ln(X_i!) - \lambda\bigr) = \\ &= \ln\lambda\sum_{i=1}^{n}X_i - \sum_{i=1}^{n}\ln(X_i!) - n\lambda. \end{aligned}$$

Zderivováním logaritmické věrohodnostní funkce dostáváme věrohodnostní rovnici

$$\frac{\partial l}{\partial\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}X_i - n.$$

Věrohodnostní rovnici položíme rovnu nule a dostaneme maximálně věrohodný odhad parametru λ, který je roven

$$\begin{aligned} \frac{1}{\widehat{\lambda}}\sum_{i=1}^{n}X_i - n &= 0 \\ \widehat{\lambda} &= \frac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}_n. \end{aligned}$$

4.2.3 Vlastnosti odhadů

Oběma metodami nám vyšly stejné odhady. Odhad střední hodnoty je výběrový průměr, jehož vlastnosti jsme již ověřili a víme, že tento odhad je nestranný a konzistentní.

4.3 Binomické rozdělení

4.3.1 Momentová metoda

Binomické rozdělení $\mathrm{Bi}(m,p)$ má pravděpodobnostní funkci

$$P(X=x) = \binom{m}{x}p^k(1-p)^{m-x}, \ x = 0,1,\ldots,m$$

střední hodnota je $\mu' = mp$ a rozptyl $\sigma^2 = mp(1-p)$. Momentovou metodou dostaneme pro porovnání $M_1' = \mu_1'$ tedy

$$\bar{X}_n = mp$$

odtud dostaneme odhad parametru p

$$\widehat{p} = \frac{\bar{X}_n}{m} = \frac{1}{nm}\sum_{i=1}^{n}X_i.$$

4.3.2 Metoda maximální věrohodnosti

Věrohodnostní funkce je rovna

$$L(n,p) = \prod_{i=1}^{n} \binom{m}{X_i} p^{X_i} (1-p)^{m-X_i}.$$

Logaritmická věrohodnostní funkce je

$$\begin{aligned} l(m,p) &= \ln\left(\prod_{i=1}^{n} \binom{m}{X_i} p^{X_i} (1-p)^{m-X_i}\right) \\ &= \ln\left(\prod_{i=1}^{n} \binom{m}{X_i}\right) + \ln p \sum_{i=1}^{n} X_i + \ln(1-p) \sum_{i=1}^{n} (m - X_i) \\ &= \ln\left(\prod_{i=1}^{n} \binom{m}{X_i}\right) + \ln p \sum_{i=1}^{n} X_i + \ln(1-p) \left(mn - \sum_{i=1}^{n} X_i\right). \end{aligned}$$

Protože je

$$\lim_{p \to 0_+} l(m,p) = \lim_{p \to 1_-} l(m,p) = -\infty$$

má věrohodnostní funkce maximum ve stacionárním bodě. Pro něj dostaneme věrohodnostní rovnici

$$\frac{\partial l(m,p)}{\partial p} = \frac{\sum_{i=1}^{n} X_i}{p} - \frac{mn - \sum_{i=1}^{n} X_i}{1-p}.$$

Věrohodnostní rovnici položíme rovnu nule a vyjádříme si maximálně věrohodný odhad parametru p

$$\begin{aligned} \frac{\sum_{i=1}^{n} X_i}{\widehat{p}} - \frac{mn - \sum_{i=1}^{n} X_i}{1-\widehat{p}} &= 0 \\ \frac{\sum_{i=1}^{n} X_i}{\widehat{p}} &= \frac{mn - \sum_{i=1}^{n} X_i}{1-\widehat{p}} \\ \sum_{i=1}^{n} X_i - \widehat{p} \sum_{i=1}^{n} X_i &= \widehat{p} mn - \widehat{p} \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i &= \widehat{p} mn \\ \widehat{p} &= \frac{1}{mn} \sum_{i=1}^{n} X_i = \frac{\bar{X}_n}{m}. \end{aligned}$$

4.3.3 Vlastnosti odhadů

Oběma metodami nám vyšly stejné odhady. Odhad $\widehat{p}$ není nestranný odhad viz [1].

$$\mathrm{E}\widehat{p} = \mathrm{E}\frac{\bar{X}_n}{m} = \frac{1}{m}\mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\frac{1}{n} n \mathrm{E} X_i}{m} = \frac{mp}{m} = p,$$

tedy odhad parametru p je nestranný. Tento odhad je i konzistentní.

5. Přehled bododových odhadů

Tato kapitola bude shrnutím všech odhadů parametrů, které jsme napočítali. V tabulkách bude vždy určité rozdělení a pro něj uvedeny odhady metodou momentů, kterou budeme značit MM a metodu maximální věrohodnosti označíme jako MMV. U těch odhadů u kterých jsme ověřili nestrannost uvedeme tuto vlastnost v tabulce. Všechny odhady, která jsem spočítali jsou konzistentní.

Spojitá rozdělení

U normálního rozdělenení nám vyšly oběma metodami stejné odhady pro oba parametry. Odhad $\widehat{\mu}$ se nazývá výběrový průměr a odhad $\widehat{\sigma}^2$ je empirický rozptyl.

Rozdělení	MM	MMV	Nestrannost
$\mathrm{N}(\mu, \sigma^2)$	$\widehat{\mu} = \bar{X}_n$	$\widehat{\mu} = \bar{X}_n$	ano
$\mathrm{N}(\mu, \sigma^2)$	$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$	$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}_n\right)^2$	ne

Tabulka A: Normální rozdělení.

Pro gama rozdělení nám vyšly rozdílné odhady a navíc pro výpočet odhadů parametrů a a p metodou maximální věrohodnosti potřebujeme software, který by nám numericky dopočítal hodnoty těchto odhadů. Tedy pro gama rozdělení je vhodnější použití momentové metody.

Rozdělení	MM	Nestrannost	MMV
$\mathrm{Ga}(a, p)$	$\widehat{a} = \frac{X_n}{S_n^2}$	ano	$\widehat{a} = \frac{\widehat{p}}{\bar{X}_n}$
$\mathrm{Ga}(a, p)$	$\widehat{p} = \frac{X_n^2}{S_n^2}$	ne	$\ln \widehat{p} - \psi\left(\widehat{p}\right) = \ln \frac{X_n}{\sqrt{n} \prod_{i=1}^{n} X_i}$

Tabulka B: Gama rozdělení.

U logaritmicko-normálního rozdělení nám vyšly různé odhady parametrů μ a σ^2 a pro toto rozdělení je lepší použití metody maximální věrohodnosti. Pro parametr μ dává metoda maximální věrohodnosti nestranný odhad, ale odhad parametru σ^2 není nestranný.

Rozdělení	MM	MMV
$\mathrm{LN}(\mu, \sigma^2)$	$\widehat{\mu} = \ln\left(M_1'\right) - \frac{1}{2} \ln\left(1 + \frac{M_2}{\left(M_1'\right)^2}\right)$	$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \ln X_i$
$\mathrm{LN}(\mu, \sigma^2)$	$\widehat{\sigma}^2 = \ln\left(1 + \frac{M_2}{M_1'}\right)$	$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(\ln X_i - \widehat{\mu}\right)^2$

Tabulka C: Logaritmicko-normální rozdělení.

Pro exponenciální rozdělení jsme odhadovali parametr λ a u obou metod nám vyšel výběrový průměr. Tedy obě metody jsou stejně vhodné pro výpočet tohoto odhadu.

Rozdělení	MM	Nestrannost	MMV	Nestrannost
$\mathrm{Ex}(\lambda)$	$\widehat{\lambda} = \bar{X}_n$	ano	$\widehat{\lambda} = \bar{X}_n$	ano

Tabulka D: Exponenciální rozdělení.

Momentovou metodou nám pro rovnoměrné rozdělení vyšly odhad $\widehat{\mu}$, který je nestranný a odhad $\widehat{h}$ není nestranný. U odhadů metodou maximální věrohodnosti musí pro X_i platit $1 \leq i \leq n$.

Rozdělení	MM	MMV
$\mathrm{R}(\mu, h)$	$\widehat{\mu} = \bar{X}_n$	$\widehat{h} = \frac{1}{2}\left(\max\{X_i\} - \min\{X_i\}\right)$
$\mathrm{R}(\mu, h)$	$\widehat{h} = \sqrt{3\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 + \left(\bar{X}_n\right)^2\right)}$	$\widehat{\mu} = \frac{1}{2}\left(\max\{X_i\} - \min\{X_i\}\right)$

Tabulka E: Rovnoměrné rozdělení.

Diskrétní rozdělení

U geometrického rozdělení odhad parametru p vyšla převrácená hodnota výběrovového průměru.

Rozdělení	MM	Nestrannost	MMV	Nestrannost
$\mathrm{Ge}(p)$	$\widehat{p} = \frac{1}{\bar{X}_n}$	ano	$\widehat{p} = \frac{1}{\bar{X}_n}$	ano

Tabulka F: Geometrické rozdělení.

Poissonovo rozdělení má parametr λ a odhad tohoto parametru je výběrový průměr, který vyšel oběma metodami.

Rozdělení	MM	Nestrannost	MMV	Nestrannost
$\mathrm{Po}(\lambda)$	$\widehat{\lambda} = \bar{X}_n$	ano	$\widehat{\lambda} = \bar{X}_n$	ano

Tabulka G: Poissonovo rozdělení.

Pro parametr binomického rozdělění dává metoda momentů i metoda maximální věrohodnosti stejný výsledek.

Rozdělení	MM	Nestrannost	MMV	Nestrannost
$\mathrm{Bi}(m, p)$	$\widehat{p} = \frac{X_n}{m}$	ano	$\widehat{p} = \frac{X_n}{m}$	ano

Tabulka H: Binomické rozdělení.

Závěr

V první kapitole jsme se seznámili s tím, co jsou to rozdělení, jak mohou vypadat některé jejich hustoty a zadefinovali jsme si základní pojmy týkající se bodového odhadu. V další kapitole jsme si představili, jak lze odhadovat parametry spojitých distribucí. Ukázali jsme, že lze pro bodový odhad použít momentovou metodu nebo metodu maximální věrohodnosti. V třetí kapitole jsme napočítali odhady parametrů spojitých rozdělení pomocí dvou dříve zmíněných metod a zjistili jsme, že i přes velkou početní náročnost metody maximální věrohodnosti existuje efektivní řešení odhadování parametrů v podobě metody momentů. Čtvrtá kapitola zahrnuje totéž jako třetí kapitola, pouze s tím rozdílem, že se věnujeme diskrétním náhodným veličinám. Pátá kapitola je přehledem spočítaných odhadů.

Pro základní rozdělení dávají obě metody skoro vždy shodné výsledky, vyjímkou je gama rozdělení, logaritmicko-normální rozdělení a rovnoměrné rozdělení. Nelze jednoznačně rozhodnout, která z metod poskytuje lepší výsledky. Rozhodování provádíme podle dané situace, nejčastěji rozhoduje jednoduchost získaných vzorců. Metoda momentů zohledňuje všechna data z výběru a volíme ji v případech, kdy je soustava věrohodnostních rovnic obtížně řešitelná.

Informace, dostupné v této bakalářské práci, lze využít při rozhodování, kterou metodu pro bodový odhad použít, neboť u zmíněných metod je postup výpočtu a základní princip. Bude záležet jen na rozhodnutí odhadujícího pracovníka, kterou metodu použije a která splňuje dané požadavky.

Z této práce vyplývájí poznatky o tom, že je nutné aby student nebo pracovník, který bude pracovat s daty a provádět odhad, měl základní představu o pravděpodobnostním rozdělení dat, o příkladu nebo o analýze, kterou bude řešit a ovládá problematiku daného oboru. Pokud tomu tak nebude a obtížnost projektu bude odhadovat řádně nepoučená osoba, může to vést ke špatným vysledkům a nesplnitelným odhadům.

Téma této bakalářské práce pro mne bylo velkou výzvou, neboť jsem o problematice odhadování bodových parametrů slyšela, ale doposud jsem je nikde nenalezla spočítané odhady, abych se seznámila s funkčností metod. Poznání metod je pro mě velkým přínosem, především vhodnost jejich použití. Metoda maximální věrohodnosti mne oslovila především svou jednoduchostí a poměrně přesnými výsledky.

Seznam použité literatury

[1] ANDĚL, Jiří: *Základy matematické statistiky*, Praha: Matfyzpress, 2011, 358 s., ISBN 978-80-7378-162-0

[2] BISKUP, Roman: *Bodový odhad parametrů rozdělení*, Praha, 2012, http://home.ef.jcu.cz/ birom/stat/

[3] FRANK Jakub: *Odhady parametrů vícerozměrného Studentova rozdělení*, Brno, 2009. 41 s. Diplomová práce. Masarykova Univerzita v Brně.

[4] KULICH, Michal: *Malý větníček*, Praha, 2013, http://www.karlin.mff.cuni.cz/ kulich/

[5] REISS Martin: *Odhady parametrů některých rozdělení pravděpodobnosti a jejich vlastnosti*, Brno, 2011. 32 s. Bakalářská práce. Masarykova univerzita.

[6] VAŇKÁTOVÁ Kristýna: *Metody odhadu regresních parametrů*, Olomouc, 2012. 95 s. Bakalářská práce. Univerzita Palackého v Olomouci.

[7] ZVÁRA, Karel; ŠTĚPÁN, Josef: *Pravděpodobnost a matematická statistika*, Praha : Matfyzpress, 2006. 230 s., ISBN 80-86732-71-7