

11. Teorie odhadu – úvod

V předchozí kapitole jsme popsali jeden ze základních principů matematické statistiky – výběr dat z populace. V obecném případě můžeme předpokládat, že daný výběr proběhne dvěma způsoby, prvky výběru jsou z populace vybírány s vrácením prvku do populace nebo jsou vybírány bez vrácení do populace. Právě oba tyto způsoby budeme nyní postupně studovat.

Obecně můžeme předpokládat, že zkoumaná data je možno popsat určitou náhodnou veličinou. Pro jisté druhy aplikací je podstatný druh této neznámé náhodné veličiny, ve většině případů je postačující dokonale popsat hledanou náhodnou veličinu pomocí neznámých čísel – parametrů této náhodné veličiny. Převážně jsou tyto parametry rovny některému ze základních momentů této neznámé náhodné veličiny.

Základní úlohou teorie bodového odhadu je vypracování metod na jejichž základě je možno z vybraných dat získat pokud možno nejlepší odhad hledaných parametrů náhodné veličiny, která popisuje rozdělení pravděpodobnosti populace.

11.1 Náhodný výběr

Definice 11.1

Náhodným výběrem ze základního souboru (populace) , která je popsána pomocí hustoty (kde Θ je neznámý parametr populace) , budeme nazývat posloupnost nezávislých náhodných veličin $X_1, X_2, \dots, X_n$, z nichž každá je popsána stejnou hustotou $f(x, \Theta)$.

Přirozené číslo n potom nazýváme **rozsahem výběru**.

Poznámka 11.2

V případě, že je populace popsána jako spojitě rozdělení je hodnota $f(x, \Theta)$ chápána jako hustota tohoto rozdělení, v případě diskrétního rozdělení popisuje dále $f(x, \Theta)$ hodnotu pravděpodobnosti rozdělení v bodě x .

Definice 11.3

Množinou všech možných výsledků (realizací) náhodného výběru $\mathbf{X}$ zapisujeme jako $\mathbf{x} = (x_1, \dots, x_n)$ a nazýváme ji **výběrovým prostorem**.

Definice 11.4

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr z populace s rozdělením $f(x, \Theta)$, realizovaný hodnotami $x = (x_1, \dots, x_n)$. Potom každou funkci

$$S_n = S(X_1, \dots, X_n) \quad (11.1)$$

definovanou na výběrovém prostoru, pomocí níž odhadujeme neznámý parametr Θ , budeme nazývat **výběrovým odhadem parametru nebo výběrovou statistiku**.

V dalším uvedeme výčet nejužívanějších druhů výběrových statistik, které budeme používat v dalším textu k odhadu jednotlivých parametrů.

Definice 11.5

Výběrový úhrn :

$$m = \sum_{i=1}^n x_i \quad (11.2)$$

Výběrový obecný moment k – tého stupně

$$m'_k = \frac{1}{n} \cdot \sum_{i=1}^n x_i^k \quad (11.3)$$

Výběrový průměr

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (11.4)$$

Výběrový centrální moment k – tého stupně

$$m_k = \frac{\sum_{i=1}^n (x_i - \bar{x})^k}{n}, \quad k = 1, 2, \dots \quad (11.5)$$

Výběrový rozptyl při známé střední hodnotě populace μ

$$s_0^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (11.6)$$

Výběrový rozptyl

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad (11.7)$$

V následujících oddílech popíšeme některé nejpoužívanější druhy náhodných výběrů.

11.1.1 Náhodný výběr bez vracení.

Předpokládejme dále, že populace má celkem N prvků a ve výběru bude právě $n \ll N$ prvků a dále, že populace základního souboru má neznámou střední hodnotu μ a rozptyl σ^2 .

Výběr n prvků lze tedy provést z populace celkem $\binom{N}{n}$ způsoby. Předpokládáme, že každá n -tice má stejnou pravděpodobnost výběru. Tato pravděpodobnost je tedy dána hodnotou $\frac{1}{\binom{N}{n}}$. Proto, abychom mohli jednoduše vyšetřit základní vlastnosti tohoto druhu výběru,

zavedeme pomocné náhodné veličiny Q_i , které jsou pro daný výběr v definovány takto:

$$Q_i = Q_i(v) = \begin{cases} 1, & i \in v \\ 0, & i \notin v \end{cases} \quad (11.8)$$

Věta 11.6

Pro náhodné veličiny $Q_1, \dots, Q_n$ platí

$$P(Q_i = 1) = \frac{n}{N} \quad (11.9)$$

$$E(Q_i) = \frac{n}{N} \quad (11.10)$$

$$VAR(Q_i) = \frac{n}{N} \cdot \frac{N-n}{N} \quad (11.11)$$

$$E(Q_i Q_j) = \frac{n(n-1)}{N(N-1)}, \quad i \neq j \quad (11.12)$$

$$\text{cov}(Q_i, Q_j) = -\frac{n \cdot (N-n)}{N^2 \cdot (N-n)}, \quad i \neq j \quad (11.13)$$

Důkaz ponechávám jako cvičení.

Všimněme si, že vyšetřované náhodné veličiny jsou typu alternativní rozdělení, ale jejich součet není binomické rozdělení, neboť je stále konstantní a rovný n . Ověřte proč je tomu tak!

Jak již vyplývá z předchozích kapitol jsou nejvýznamnější bodové odhady střední hodnoty a rozptylu. Jak dále uvidíme, neoptimálnějšími statistikami pro tvorbu těchto odhadů se jeví výběrový průměr $\bar{X}$ jako bodový odhad střední hodnoty a výběrová směrodatná odchylka S^2 jako bodový odhad rozptylu. Proto se v této a následných částech budeme soustřeďovat na určení vlastností právě těchto veličin.

Věta 11.7

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr bez vracení z konečné populace o N prvcích o neznámé střední hodnotě μ a neznámém rozptylu σ^2 . Potom pro výběrový průměr $\bar{X}$ platí následující výroky

$$E(\bar{X}) = \mu \quad (11.14)$$

$$\text{VAR}(\bar{X}) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n} \quad (11.15)$$

Důkaz:

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = E\left(\frac{1}{n} \cdot \sum_{i=1}^n x_i \cdot Q_i\right) = \frac{1}{n} \cdot \sum_{i=1}^n x_i \cdot \frac{n}{N} = \frac{1}{N} \cdot \sum_{i=1}^n x_i = \mu. \\ \text{VAR}(\bar{X}) &= E(\bar{X} - \mu)^2 = E\left(\frac{1}{n} \cdot \sum_{i=1}^n x_i \cdot Q_i - \mu \cdot \frac{1}{n} \cdot \sum_{i=1}^n Q_i\right)^2 = \frac{1}{n^2} \cdot E\left(\sum_{i=1}^n (x_i - \mu) \cdot Q_i\right)^2 = \\ &= \frac{1}{n^2} \cdot \left(\sum_{i=1}^n (x_i - \mu)^2 \cdot E Q_i^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^n (x_i - \mu) \cdot (x_j - \mu) \cdot E Q_i Q_j \right) = \\ &= \frac{1}{n^2} \cdot \left(\frac{n}{N} \cdot \sum_{i=1}^n (x_i - \mu)^2 + \frac{n \cdot (n-1)}{N \cdot (N-1)} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^n (x_i - \mu) \cdot (x_j - \mu) \right) = \\ &= \frac{1}{n^2} \cdot \left(\frac{n}{N} - \frac{n \cdot (n-1)}{N \cdot (N-1)} \right) \cdot N \cdot \sigma^2 = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}. \end{aligned}$$

Q.E.D.

Zlomek $\frac{N-n}{N-1}$ se nazývá **konečnostní násobitel nebo faktor**. Ze tvrzení věty je zřejmé, že vliv tohoto faktoru se v případě mnohem větších hodnot N než n výrazně zmenšuje. Tento faktor budeme dále užívat v korekcích (opravách) na spojitost, tedy v případech, kdy budeme nahrazovat diskrétní rozdělení rozdělením spojitým (většinou rozdělením typu $N(\mu, \sigma^2)$).

Podobně v následující větě budeme vyšetřovat vlastnost výběrového rozptylu S^2 .

Věta 11.8

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr bez vracení z konečné populace o N prvcích o neznámé střední hodnotě μ a neznámém rozptylu σ^2 . Potom pro výběrový rozptyl S^2 platí

$$ES^2 = \frac{N}{N-1} \cdot \sigma^2 \quad (11.16)$$

Důkaz:

$$\begin{aligned} ES^2 &= E\left(\frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \bar{X})^2\right) = E\left(\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \mu)^2 \cdot Q_i - \frac{n}{n-1} \cdot (\bar{X} - \mu)^2\right) = \\ &= \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \mu)^2 \cdot EQ_i - \frac{n}{n-1} \cdot E(\bar{X} - \mu)^2 = \frac{1}{n-1} \cdot N \cdot \sigma^2 \cdot \frac{n}{N} - \frac{n}{n-1} \cdot \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n} = \\ &= \frac{n \cdot \sigma^2}{n-1} - \frac{(N-n) \cdot \sigma^2}{(n-1) \cdot (N-1)} = \sigma^2 \cdot \frac{n \cdot N - N}{(n-1) \cdot (N-1)} = \sigma^2 \cdot \frac{N}{N-1} \end{aligned}$$

Q.E.D.

Z tvrzení věty vyplývá, že pro výběrový průměr v tomto případě neplatí obdoba vztahu (11.14).

11.1.2 Náhodný výběr s vracením

V této části budeme studovat opět náhodný výběr z populace (konečné nebo nekonečné), ale tentokrát budeme prvek po jeho zvolení opět vracet do základního souboru. Budeme však požadovat, aby všechny výběry byly prováděny za stejných podmínek a byly nezávislé. Budeme dále předpokládat, že populace i náhodné veličiny jsou popsány stejnými náhodnými veličinami se stejnou střední hodnotou μ a rozptylem σ^2 . Za takovýchto předpokladů budeme nyní vyšetřovat výběrový průměr $\bar{X}$ a výběrový rozptyl S^2 .

Věta 11.9

Pro výběrový průměr spočítaný z náhodného výběru z populace s vracením platí následující tvrzení

$$E(\bar{X}) = \mu \quad (11.17)$$

$$VAR(\bar{X}) = \frac{\sigma^2}{n} \quad (11.18)$$

Důkaz:

$$\begin{aligned} E(\bar{X}) &= E\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot \sum_{i=1}^n E(X_i) = \frac{n \cdot \mu}{n} = \mu \\ VAR(\bar{X}) &= VAR\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot \sum_{i=1}^n VAR(X_i) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

Q.E.D.

V druhé části důkazu jsme využili především nezávislost náhodných veličin X_i . Je zřejmé, že výběrový průměr i v tomto případě splňuje vlastnost nestrannosti (viz definice 11.12), v případě výběru s vracením hodnota rozptylu odhadu menší než u výběru bez vracení.

Podobně zformulujeme tvrzení pro případ výběrového rozptylu.

Věta 11.10

Nechť pro populaci existuje čtvrtý centrální moment μ_4 , potom pro výběrový rozptyl spočítaný z náhodného výběru z populace s vrácením platí následující tvrzení

$$E(S^2) = \sigma^2 \quad (11.19)$$

$$VAR(S^2) = \frac{1}{n} \cdot \left(\mu_4 - \frac{n-3}{n-1} \cdot \sigma^4 \right) \quad (11.20)$$

Důkaz:

Využijeme opět přičtení a odečtení hodnoty μ ve vztahu pro S^2 . Potom je

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

Druhou část tvrzení nebudeme dokazovat odkazujeme se na větu 5.8.3 v [].
Q.E.D.

Z tvrzení věty je zřejmé, že v tomto případě je výběrový průměr S^2 nestranný.

Velmi důležitá pro další aplikace viz např. intervalový odhad v následující 11. části, je podmínka při níž je možno zanedbat rozdíl mezi výběrem s vrácením a výběrem bez vrácení.

Tato je splněna tehdy když $\frac{n}{N} \leq 0,05$, tedy tehdy když je podíl vybraných prvků maximálně 5% z celku!

11.2 Vlastnosti odhadů

V této části se budeme zabývat základními vlastnostmi bodového odhadu jednoho či více neznámých parametrů populace. Pro další využití forem bodového odhadu je nutné nejdříve uvést několik definic nutných pro vyšetřování vlastností zkonstruovaných bodových odhadů.

Definice 11.11

Nechť Θ je vyšetřovaný parametr, potom označíme Ω množinu všech možných hodnot tohoto neznámého parametru a nazveme ji **parametrickým prostorem**.

Definice 11.12

Odhad S_n , pro který platí $E(S_n) = \Theta$, nazýváme **nevychýleným odhadem parametru Θ** . Jestliže platí nerovnost $E(S_n) > \Theta$, nazýváme odhad kladně vychýleným odhadem parametru Θ . Jestliže platí opačná nerovnost $E(S_n) < \Theta$, nazýváme odhad záporně vychýleným odhadem parametru Θ .

Příklad 11.13

Nechť $S_n = \frac{\sum_{i=1}^n X_i}{n}$ je bodový odhad střední hodnoty μ populace. Potom podle věty je tento odhad S_n nevychýleným odhadem, neboť $E(S_n) = \mu$.

Příklad 11.14

Nechť $S_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ je statistika zkonstruovaná jako bodový odhad rozptylu σ^2 . Potom je podle způsobu výběru buď nevychýleným odhadem tohoto parametru (viz věta 11.10) nebo je záporně vychýleným odhadem tohoto parametru (viz věta 11.8).

Uvedená vlastnost bodových odhadů zaručuje, že pomocí takového odhadu je možno vyloučit systematické chyby, ale tato vlastnost sama o sobě ještě negarantuje to, že odhad bude hodnotný.

Definice 11.15

Odhad parametru Θ , pro který platí $\lim_{n \rightarrow \infty} E(S_n) = \Theta$ nazýváme **asymptoticky nevychýleným odhadem parametru Θ** . Podobně bychom mohli definovat kladně či záporně asymptoticky vychýlené odhady parametru Θ .

Příklad 11.16

Nechť $S_n = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ je statistika zkonstruovaná jako bodový odhad rozptylu σ^2 . V případě výběru bez vracení je tento odhad odhadem asymptoticky nevychýleným.

Definice 11.17

Bodový odhad S_n nazveme **konzistentním odhadem parametru Θ** , jestliže náhodná veličina S_n konverguje podle pravděpodobnosti k parametru Θ neboli

$$\forall \varepsilon > 0 \lim_{n \rightarrow \infty} P(|S_n - \Theta| \geq \varepsilon) = 0 \quad (11.21)$$

Jestliže je bodový odhad konzistentní, potom je malá pravděpodobnost velké chyby v odhadu parametru Θ , jestliže rozsah výběru roste. Následující věta udává postačující podmínku konzistentnosti daného bodového odhadu.

Věta 11.18

Nechť je S_n nevychýleným nebo asymptoticky nevychýleným odhadem parametru Θ a $E(S_n^2) < \infty$ pro každé přirozené n . Nechť dále platí $\lim_{n \rightarrow \infty} VAR(S_n) = 0$. Potom je S konzistentním odhadem parametru Θ .

Důkaz:

Zvolme $\varepsilon > 0$ libovolné. Z Čebyševovy nerovnosti vyplývá

$$P(|S_n - E(S_n)| \geq \varepsilon) \leq \frac{VAR(S_n)}{\varepsilon^2}.$$

Dále protože $\lim_{n \rightarrow \infty} E(S_n) = \Theta$, existuje N takové, že pro $n > N$ platí

- $\varepsilon < S_n - E(S_n) < \varepsilon$. Nyní již můžeme provést odhad

$$P(|S_n - \Theta| \geq 2\varepsilon) = 1 - P(|S_n - \Theta| < 2\varepsilon) = 1 - P(|S_n - E(S_n) + E(S_n) - \Theta| < 2\varepsilon) \leq$$

$\leq 1 - P(|S_n - E(S_n)| < \varepsilon) = P(|S_n - E(S_n)| \geq \varepsilon) \leq \frac{VAR(S_n)}{\varepsilon^2}$. Nyní již provedeme jen limitní přechod .
Q.E.D.

Pro vlastní hodnocení bodových odhadů se ještě často užívá hodnocení kvality odhadu. Kvalitu odhadu budeme vyšetřovat pomocí rozptylu takového odhadu.

Definice 11.19

Nechť S_n je nevychýlený odhad parametru Θ . Nechť dále T_n je libovolný nevychýlený odhad stejného parametru. Řekneme, že S_n je **nejlepším nevychýleným odhadem parametru Θ** , jestliže platí

$$VAR(S_n) \leq VAR(T_n) \quad (11.22)$$

11.3 Konstrukce nejlepších odhadů

Posuzování kvality bodových odhadů z hlediska vydatnosti usnadňuje znalost dolní hranice rozptylů nevychýlených bodových odhadů. Abychom ji vyšetřili zavedeme základní definici.

Definice 11.20

Systém hustot $f(x, \Theta), \Theta \in \Omega$ budeme nazývat **regulárním** jestliže:

- Množina Ω je otevřená
- Výběrový prostor $V = \{x; f(x, \Theta) > 0\}$ nezávisí na Θ
- Pro každé $\Theta \in \Omega$ existuje derivace $\frac{\partial f(x, \Theta)}{\partial \Theta}$
- Pro každé $\Theta \in \Omega$ platí $E\left(\frac{\partial \ln f(x, \Theta)}{\partial \Theta}\right) = 0$
- Výraz $I(\Theta) = VAR\left(\frac{\partial \ln f(x, \Theta)}{\partial \Theta}\right) = E\left(\left(\frac{\partial \ln f(x, \Theta)}{\partial \Theta}\right)^2\right)$ je pro všechna Θ konečný a kladný.

Podobně nazveme nezkreslený odhad $\tau(\mathbf{X})$ parametrické funkce $\gamma(\Theta)$ regulárním, jestliže pro každé $\Theta \in \Omega$

$$a) \text{ Existuje derivace } \frac{d\gamma(\Theta)}{d\Theta}$$

$$b) \text{ Platí } \frac{d}{d\Theta} \int_{V^n} \tau(x) \cdot f(x, \Theta) dx = \int_{V^n} \tau(x) \cdot \frac{\partial f(x, \Theta)}{\partial \Theta} dx.$$

Pro další ilustraci uveďme několik příkladů.

Příklad 11.21

Ukažme, že systém hustot $f(x, \Theta) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\Theta)^2}{2}}$, pro $\Theta \in (-\infty, \infty)$ je regulární.

- Množina Ω je otevřená
- Výběrový prostor $V = \{x; f(x, \Theta) > 0\}$ nezávisí na Θ a je to celé $\mathbb{R}_1$

c) Pro každé Θ existuje derivace $\frac{\partial f(x, \Theta)}{\partial \Theta} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\Theta)^2}{2}} \cdot (x - \Theta)$

d) $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\Theta)^2}{2}} \cdot (x - \Theta) dx = 0$

e) $I(\Theta) = \int_{-\infty}^{\infty} (x - \Theta)^2 \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\Theta)^2}{2}} dx = 1.$

V dalším příkladě prostudujeme případ diskrétního rozdělení.

Příklad 11.22

Ukažme, že systém hustot $f(x, \Theta) = \frac{e^{-\Theta} \cdot \Theta^x}{x!}$, pro $\Theta \in (0, +\infty)$ je regulární.

a) Množina Ω je otevřená

b) Výběrový prostor $V = \{x; f(x, \Theta) > 0\}$ nezávisí na Θ a je to celé $\mathbb{R}_1$

c) Pro každé Θ existuje derivace $\frac{\partial f(x, \Theta)}{\partial \Theta} = \frac{-\Theta^x \cdot e^{-\Theta} + x \cdot \Theta^{x-1} \cdot e^{-\Theta}}{x!}$

d) $\ln f(x, \Theta) = -\Theta + x \cdot \ln \Theta - \ln x!$

$$\frac{\partial \ln f(x, \Theta)}{\partial \Theta} = \frac{x - \Theta}{\Theta}, \quad E\left(\frac{X - \Theta}{\Theta}\right) = \frac{1}{\Theta} \cdot (E(X) - \Theta) = 0$$

e) $I(\Theta) = E\left(\left(\frac{X - \Theta}{\Theta}\right)^2\right) = \frac{1}{\Theta^2} \cdot \text{VAR}(X) = \frac{\Theta}{\Theta^2} = \frac{1}{\Theta}.$

V předchozích příkladech jsme studovali spojité rozdělení typu $N(\mu, 1)$ a diskrétní rozdělení $\text{Po}(\lambda)$.

Definice 11.23

Výše uvedenou hodnotu $I(\Theta)$ budeme nazývat **Fischerovou mírou informace** parametru Θ .

Zásadní význam pro naše hledání optimálních odhadů má následující věta

Věta 11.24 Rao – Cramerova nerovnost

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr z regulárního rozdělení $f(x, \Theta)$. Nechť $\tau(\Theta)$ je daná parametrická funkce. Potom pro každý regulární nevychýlený bodový odhad $S(X)$ parametrické funkce $\tau(\Theta)$ platí

$$\text{VAR}(S(X)) \geq \frac{(\tau'(\Theta))^2}{n \cdot I(\Theta)} = C(\Theta), \quad \text{kde hodnota } \tau'(\Theta) \text{ je derivace funkce } \tau \text{ podle}$$

parametru Θ .

Důkaz tvrzení nebudeme provádět, lze ho najít například v [].

Na základě tohoto tvrzení je zřejmé, že nejlepším nevychýleným odhadem v případě prostého parametru tj. $\tau(\Theta) = \Theta$, je odhad jehož rozptyl bude roven číslu $\frac{1}{n \cdot I(\Theta)}$.

Uveďme dále dva příklady.

Příklad 11.25

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr z Poissonova rozdělení s neznámým parametrem λ . Podle předchozí věty je hodnota $C(\lambda) = \frac{1}{n \cdot I(\lambda)} = \frac{1}{n \cdot \frac{1}{\lambda}} = \frac{\lambda}{n}$. Tedy

bodový odhad, jehož rozptyl bude roven této hodnotě je v daném slova smyslu nejlepší. Využili jsme znalosti $I(\lambda)$ z příkladu 11.22.

Příklad 11.26

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr z normálního rozdělení $N(\mu, \sigma^2)$, kde hodnotu σ^2 známe. Opět podle věty je dolní hodnota $C(\mu)$ rovna $\frac{1}{n \cdot I(\mu)}$. Ale podle příkladu 11.21 (proved'te samostatně úpravu rozptylu) je Fischerova míra informace rovna $I(\mu) = \frac{1}{\sigma^2}$. Tedy $C(\mu) = \frac{1}{n \cdot \frac{1}{\sigma^2}} = \frac{\sigma^2}{n}$. Pokud se nám podaří nalézt odhad

parametru μ při známé hodnotě σ^2 , který bude mít rozptyl roven $\frac{\sigma^2}{n}$, bude to nejvydatnější odhad. Všimněme si, že zřejmě v tomto případě je tím správným bodovým odhadem přímo výběrový průměr, neboť podle věty 11.9 má tyto správné vlastnosti.

11.4 Základní metody bodového odhadu

V této části se na rozdíl od předchozího textu budeme systematicky zabývat odvozováním nejlepších bodových odhadů parametrů populace. Obecně existují dvě základní metody tvorby statistik s danými vlastnostmi. Je to metoda maximální věrohodnosti a metoda momentová. Obě tyto metody budeme dále studovat.

11.4.1 Metoda maximální věrohodnosti

Tato metoda umožňuje získávat velmi dobré odhady jednotlivých parametrů populace, proto je častěji užívána. Základním pojmem je zde tzv. věrohodnostní funkce.

Definice 11.27

Nechť $X = (X_1, \dots, X_n)$ je náhodný výběr a $x = (x_1, \dots, x_n)$ je jeho realizace. Nechť je dále popsána populace pomocí regulární hustoty $f(x, \Theta)$, kde Θ je neznámý parametr. Potom funkci

$$L(x, \Theta) = L(x_1, \dots, x_n, \Theta) = f(x_1, \Theta) \cdot f(x_2, \Theta) \cdot \dots \cdot f(x_n, \Theta) \quad (11.23)$$

budeme nazývat **věrohodnostní funkcí**.

Poznámka 11.28

Funkce je označovaná L jako likelihood z anglického věrohodnost.

Metoda maximální věrohodnosti spočívá v tom , že se za odhad neznámého parametru (neznámých parametrů) zvolí hodnota $\hat{\Theta}$, která při daných hodnotách maximalizuje funkci věrohodnosti.

Definice 11.29

Jestliže existuje bod $\hat{\Theta}$ z parametrického prostoru ,takový že pro všechny hodnoty parametru Θ z parametrického prostoru platí

$$L(X, \Theta) \leq L(X, \hat{\Theta}) \quad (11.24)$$

potom nazveme tento bod **maximálně věrohodným odhadem** neznámé hodnoty parametru Θ .

Pokud je funkce maximální věrohodnosti dostatečně hladká , pak je možné hledání maximálních hodnot zjednodušit na prosté hledání maxim pomocí parciálních derivací. V mnoha případech nemusíme pracovat přímo s věrohodnostní funkcí $L(x_1, \dots, x_n, \Theta)$, ale s jejím logaritmem (součin hustot bude nulový jen na míře nula , proto můžeme použít logaritmus) . Protože je funkce logaritmus prostá a rostoucí , budou všechny maximální hodnoty funkce $\ln(L(x_1, \dots, x_n, \Theta))$ stejné jako funkce $L(x_1, \dots, x_n, \Theta)$.

V další části se budeme věnovat přímo konkrétním případům .

Příklad 11.30

Nechť je populace popsána pomocí Poissonova rozdělení s parametrem neznámým λ . Odhadneme tento parametr pomocí metody maximální věrohodnosti.

$$\ln(L(x_1, \dots, x_n, \lambda)) = \ln\left(\prod_{i=1}^n \frac{\lambda^{x_i} \cdot e^{-\lambda}}{x_i!}\right) = \left(\sum_{i=1}^n x_i\right) \cdot \ln(\lambda) - n \cdot \lambda - \sum_{i=1}^n \ln(x_i!)$$

Tedy zjistíme stacionární body

$$\frac{\partial \ln(L(x_1, \dots, x_n, \lambda))}{\partial \lambda} = \frac{1}{\lambda} \cdot \sum_{i=1}^n x_i - n = 0 \Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}.$$

Jednoduchým výpočtem zjistíme druhá derivace v bodě $\hat{\lambda}$ je záporná , tedy jde o maximum.

Takovýto odhad je nevychýlený , konzistentní a nejvydatnější podle předchozí kapitoly.

Příklad 11.31

Nechť je populace popsána pomocí $N(\mu, \sigma^2)$, kde je parametr σ^2 znám. Odhadneme parametr μ pomocí metody maximální věrohodnosti:

$$\ln(L(x_1, \dots, x_n, \mu)) = \ln\left(\prod_{i=1}^n \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}\right) = -\frac{n}{2} \cdot \ln(2\pi) - \frac{n}{2} \cdot \ln \sigma^2 - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu)^2 \quad (11.25)$$

Maximálně věrohodný odhad parametru $\hat{\mu}$ musí být řešením rovnice

$$\frac{\partial \ln(L(x, \mu))}{\partial \mu} = \frac{1}{\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu) = 0 \Rightarrow \hat{\mu} = \frac{1}{n} \cdot \sum_{i=1}^n x_i = \bar{x}.$$

Druhá derivace logaritmické věrohodnostní funkce v bodě $\hat{\mu}$ je záporná , nastává v něm tedy maximum. Zjistili jsme tedy , že podobně jako v předchozím případě je nejlepším bodovým odhadem střední hodnoty výběrový průměr.

Příklad 11.32

Nechť je populace popsána pomocí alternativního rozdělení $\mathbf{A}(\mathbf{p})$, s neznámých parametrem p . Provedeme odhad parametru p pomocí metody maximální věrohodnosti.

Vytvoříme nejprve logaritmickou věrohodnostní funkci jako v předchozích případech.

$$\ln(L(x_1, \dots, x_n, p)) = \ln \left(p^{\sum_{i=1}^n x_i} \cdot (1-p)^{n-\sum_{i=1}^n x_i} \right) = \sum_{i=1}^n x_i \cdot \ln(p) + (n - \sum_{i=1}^n x_i) \cdot \ln(1-p)$$

$$\frac{\partial \ln(L(x, p))}{\partial p} = \sum_{i=1}^n x_i \cdot \frac{1}{p} - \left(n - \sum_{i=1}^n x_i \right) \cdot \frac{1}{1-p} = 0 \Rightarrow \hat{p} = \frac{1}{n} \cdot \sum_{i=1}^n x_i = \bar{x}$$

Výpočet druhé derivace udává zápornou hodnotu tedy v bodě $\hat{p}$ nastává maximum. I v tomto případě je nejlepším bodovým odhadem střední hodnoty výběrový průměr.

Příklad 11.33

Nechť je populace popsána pomocí $N(\mu, \sigma^2)$, kde je parametry μ a σ^2 nejsou známy. Odhadneme tyto parametry pomocí metody maximální věrohodnosti:

Využijeme funkce (11.25) z příkladu 11.31 a získáme soustavu dvou rovnic pro dvě neznámé:

$$\frac{\partial \ln(L(x, \mu, \sigma^2))}{\partial \mu} = \frac{1}{\sigma^2} \cdot \sum_{i=1}^n (x_i - \mu) = 0 \Rightarrow \hat{\mu} = \frac{1}{n} \cdot \sum_{i=1}^n x_i = \bar{x}$$

$$\frac{\partial \ln(L(x, \mu, \sigma^2))}{\partial \sigma^2} = -\frac{n}{2} \cdot \frac{1}{\sigma^2} + \frac{1}{2 \cdot \sigma^4} \cdot \sum_{i=1}^n (x_i - \mu)^2 = 0 \Rightarrow \widehat{\sigma^2} = \frac{1}{n} \cdot \sum_{i=1}^n (X_i - \bar{X})^2$$

V předchozím výrazu se provádí derivace podle σ^2 nikoli podle σ ! Zjistíme ještě matici druhých derivací

$$\begin{pmatrix} -\frac{n}{\widehat{\sigma^2}^2} & 0 \\ 0 & -\frac{n}{2 \cdot (\widehat{\sigma^2})^2} \end{pmatrix}$$

Tato matice je negativně definitní, tedy v daných bodech nastává maximum.

11.4.2 Momentová metoda

Druhou základní metodou konstrukce bodových odhadů je metoda momentová. Tato metoda poskytuje obecně méně hodnotné bodové odhady. Je u nich zaručena obecně jen konzistence. Někdy slouží tato metoda jako odhad první aproximace při řešení metodou maximální věrohodnosti.

Nechť tedy $X = (X_1, \dots, X_n)$ je náhodný výběr z populace, která závisí na obecně k parametrech $\Theta = (\Theta_1, \dots, \Theta_k)$. Předpokládejme, že pro každé Θ z parametrického prostoru existují momenty

$$\mu_m = E(X_i^m), \text{ pro } m = 1, 2, \dots, k \quad (11.26)$$

o nichž budeme předpokládat, že závisí na Θ . Výběrové momenty m'_k jsou definovány podle 11.3.

Odhady parametrů $\Theta_1, \dots, \Theta_k$ metodou momentů získáme tak, že za odhady těchto parametrů vezmeme řešení rovnic

$$\mu_i = m'_i \text{ pro } i = 1, 2, \dots, k \quad (11.27)$$

Někdy tento počet rovnic nestačí k určení všech parametrů, můžeme si vypomoci přidáním podobných vztahů jako v (11.27) pro hodnoty $i = k+1, k+2, \dots$

Celý tento postup je založen na existenci obecných momentů z (11.26).

Příklad 11.34

Nechť je populace popsána binomickým rozdělením $Bi(n, p)$ s neznámým parametrem p . Zapišeme jednu rovnici ze vztahu (11.27):

$$\mu_1 = n \cdot p = m'_1 = \frac{1}{n} \cdot \sum_{i=1}^n X_i = \bar{X} \Rightarrow \hat{p} = \frac{\bar{X}}{n},$$

v tomto případě stačilo vyřešit je jednu rovnici a získali jsme odhad parametru p . Vlastnosti tohoto odhadu ponechávám studentům jako domácí cvičení.

Příklad 11.35

Nechť je populace popsána pomocí $N(\mu, \sigma^2)$, kde je parametry μ a σ^2 nejsou známy. Odhadneme tyto neznámé parametry pomocí metody momentů.

$$\mu_1 = E(X) = \mu = \frac{1}{n} \cdot \sum_{i=1}^n X_i \Rightarrow \hat{\mu} = \bar{X}$$

$$\mu_2 = \sigma^2 = m'_2 - (m'_1)^2 = \frac{1}{n} \cdot \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \cdot \sum_{i=1}^n x_i \right)^2 = \frac{1}{n} \cdot \sum_{i=1}^n x_i^2 - \bar{x}^2 \Rightarrow$$

$$\hat{\sigma}^2 = \frac{1}{n} \cdot \sum_{i=1}^n (x_i - \bar{x})^2 = s^{*2}.$$

Odhad střední hodnoty je nevychýlený, konzistentní a nejlepší. Odhad rozptylu je asymptotický, nevychýlený a konzistentní.