Základy teorie odhadu parametrů – bodový odhad

Jiří Neubauer

Katedra ekonometrie, FVL, UO Brno
kancelář 69a, tel. 973 442029
email:Jiri.Neubauer@unob.cz

Odhady parametrů

Úkolem výběrového šetření je podat informaci o neznámé hodnotě charakteristiky základního souboru či o parametrech rozdělení základního souboru na základě náhodného výběru.

Charakteristiky základního souboru nazýváme **parametry** (příp. teoretické charakteristiky) a značíme je řeckými písmeny (μ, σ^2, θ, …).

Charakteristiky náhodného výběru nazýváme **výběrové charakteristiky** nebo **statistiky** a značíme je latinskými písmeny ($\overline{X}$, S^2, …).

Odhady parametrů

Budeme konstruovat 2 typy odhadů:

- bodový odhad
- intervalový odhad

Bodový odhad

Definice

Bodovým odhadem parametru θ rozumíme statistiku

$$T = T(X_1, X_2, \dots, X_n) = T(\mathbf{X}),$$

jejíž hodnoty kolísají kolem θ. Bodový odhad parametru θ tedy spočívá v jeho nahrazení jedním číslem (bodem). Potom píšeme $\hat{\theta} = T(\mathbf{X})$ a čteme: odhadem parametru θ je statistika $T(\mathbf{X})$.

Nestranný odhad

Definice

Statistika T je **nestranným (nevychýleným, nezkresleným)** odhadem parametru θ, platí-li

$$E\left[T(\mathbf{X})\right] = \theta.$$

Tento požadavek vyjadřuje skutečnost, že použitý bodový odhad skutečnou hodnotu charakteristiky ani nenadhodnocuje ani nepodhodnocuje.

Rozdíl

$$B(\theta, T) = E\left[T(\mathbf{X})\right] - \theta$$

se nazývá **vychýlení (zkreslení)** odhadu T.

Nestranný odhad

Definice

Statistika T je **nestranným (nevychýleným, nezkresleným)** odhadem parametru θ, platí-li

$$E\left[T(\mathbf{X})\right] = \theta.$$

Tento požadavek vyjadřuje skutečnost, že použitý bodový odhad skutečnou hodnotu charakteristiky ani nenadhodnocuje ani nepodhodnocuje.

Rozdíl

$$B(\theta, T) = E\left[T(\mathbf{X})\right] - \theta$$

se nazývá **vychýlení (zkreslení)** odhadu T.

Příklad

Mějme náhodný výběr $X_1, X_2, \ldots, X_n$ z rozdělení se střední hodnotou μ a rozptylem σ^2.

- Výběrový průměr $\overline{X}$ je nestranným odhadem parametru μ, neboť

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu.$$

- Výběrový rozptyl S^2 je nestranným odhadem parametru σ^2, neboť

$$E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \cdots = \sigma^2.$$

Příklad

Mějme náhodný výběr $X_1, X_2, \ldots, X_n$ z rozdělení se střední hodnotou μ a rozptylem σ^2.

- Výběrový průměr $\overline{X}$ je nestranným odhadem parametru μ, neboť

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \mu.$$

- Výběrový rozptyl S^2 je nestranným odhadem parametru σ^2, neboť

$$E(S^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \cdots = \sigma^2.$$

Příklad

Mějme náhodný výběr $X_1, X_2, \ldots, X_n$ z rozdělení se střední hodnotou μ a rozptylem σ^2.

- Momentový rozptyl S_n^2 není nestranným odhadem parametru σ^2, neboť

$$E(S_n^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \cdots = \frac{n-1}{n}\sigma^2.$$

Vychýlení odhadu S_n^2 parametru σ^2 je rovno

$$B(\sigma^2, S_n^2) = E(S_n^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \frac{1}{n}\sigma^2.$$

Se zvyšujícím se n se vychýlení zmenšuje.

Příklad

Mějme náhodný výběr $X_1, X_2, \ldots, X_n$ z rozdělení se střední hodnotou μ a rozptylem σ^2.

- Momentový rozptyl S_n^2 není nestranným odhadem parametru σ^2, neboť

$$E(S_n^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) = \cdots = \frac{n-1}{n}\sigma^2.$$

Vychýlení odhadu S_n^2 parametru σ^2 je rovno

$$B(\sigma^2, S_n^2) = E(S_n^2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \frac{1}{n}\sigma^2.$$

Se zvyšujícím se n se vychýlení zmenšuje.

Asymptoticky nestranný odhad

Některé odhady jsou sice zkreslené, ale s rostoucím rozsahem výběru se jejich zkreslení zmenšuje.

Definice

Je-li $T(\mathbf{X})$ odhad založený na n pozorováních a jestliže platí

$$\lim_{n \to \infty} E\left[T(\mathbf{X})\right] = \theta,$$

pak říkáme, že $T(\mathbf{X})$ je **asymptoticky nestranným odhadem** parametru θ.

Pro asymptoticky nestranný odhad tedy platí

$$\lim_{n \to \infty} E\left[T(\mathbf{X}) - \theta\right] = 0.$$

Příklad

Momentový rozptyl je asymptoticky nestranným odhadem parametru σ^2, neboť

$$\lim_{n\to\infty} E(S_n^2) = \lim_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2.$$

Konzistentní odhad

V některých případech jsme nuceni pracovat s vychýlenými odhady. Pak požadujeme, aby se odhad s rostoucím rozsahem výběru blížil odhadovanému parametru, tedy aby byl **konzistentní**.

Definice

Statistika $T(\mathbf{X})$ je **konzistentním** odhadem parametru θ, platí-li pro každé $\epsilon > 0$

$$\lim_{n\to\infty} P(|T(\mathbf{X}) - \theta| < \epsilon) = 1.$$

Je-li

$$\lim_{n\to\infty} B(\theta, T) = 0 \quad \text{a} \quad \lim_{n\to\infty} D[T(\mathbf{X})] = 0,$$

pak statistika $T(\mathbf{X})$ je konzistentní odhad parametru θ.

Konzistentní odhad

V některých případech jsme nuceni pracovat s vychýlenými odhady. Pak požadujeme, aby se odhad s rostoucím rozsahem výběru blížil odhadovanému parametru, tedy aby byl **konzistentní**.

Definice

Statistika $T(\mathbf{X})$ je **konzistentním** odhadem parametru θ, platí-li pro každé $\epsilon > 0$

$$\lim_{n\to\infty} P(|T(\mathbf{X}) - \theta| < \epsilon) = 1.$$

Je-li

$$\lim_{n\to\infty} B(\theta, T) = 0 \quad \text{a} \quad \lim_{n\to\infty} D[T(\mathbf{X})] = 0,$$

pak statistika $T(\mathbf{X})$ je konzistentní odhad parametru θ.

Příklad

Ukažte, že výběrový průměr je konzistentním odhadem střední hodnoty μ.

Vzhledem k tomu, že $E(\overline{X}) = \mu$ a $D(\overline{X}) = \sigma^2/n$ dostáváme

$$B(\mu, \overline{X}) = E(\overline{X}) - \mu = 0 \quad \text{a} \quad \lim_{n\to\infty} D(\overline{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0.$$

Příklad

Ukažte, že výběrový průměr je konzistentním odhadem střední hodnoty μ.

Vzhledem k tomu, že $E(\overline{X}) = \mu$ a $D(\overline{X}) = \sigma^2/n$ dostáváme

$$B(\mu, \overline{X}) = E(\overline{X}) - \mu = 0 \quad \text{a} \quad \lim_{n\to\infty} D(\overline{X}) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0.$$

Vydatnost odhadů

V některých případech lze najít více statistik, které jsou nestranné a konzistentní. V takovém případě použijeme k odhadování parametru tu z nich, která má nejmenší rozptyl. Statistika, která má ze všech nestranných odhadů nejmenší rozptyl je **vydatným** (**nejlepší nestranným**) odhadem parametru θ.

Nechť T a U jsou 2 nestranné odhady parametru θ, pak **vydatnost odhadu** T **vzhledem k odhadu** U je definována vztahem

$$e(T, U) = \frac{D(U)}{D(T)}.$$

Vydatnost odhadů

V některých případech lze najít více statistik, které jsou nestranné a konzistentní. V takovém případě použijeme k odhadování parametru tu z nich, která má nejmenší rozptyl. Statistika, která má ze všech nestranných odhadů nejmenší rozptyl je **vydatným** (**nejlepší nestranným**) odhadem parametru θ.

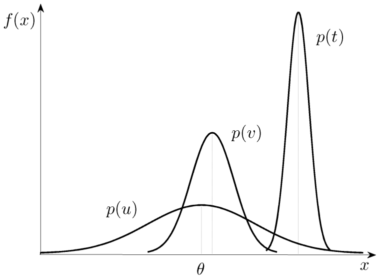
Nechť T a U jsou 2 nestranné odhady parametru θ, pak **vydatnost odhadu** T **vzhledem k odhadu** U je definována vztahem

$$e(T, U) = \frac{D(U)}{D(T)}.$$

Vydatnost odhadů

Předpokládejme nyní, že srovnáváme vychýlené i nestranné odhady parametru θ. V takovém případě nemusí být vhodné vybrat odhad s nejmenším rozptylem.

Odhad T má sice nejmenší rozptyl, ale má velké vychýlení. Ani odhad s nejmenším vychýlením nemusí být nejvhodnější. Odhad U má nulové vychýlení, ale má příliš velký rozptyl. Jako nejlepší se jeví odhad V.



Příklad

Pro parametr λ Poissonova rozdělení lze nalézt 2 nestranné odhady

$$E(\overline{X}) = \lambda \quad \text{a} \quad E(S^2) = \lambda.$$

Lze ale ukázat, že

$$D(\overline{X}) < D(S^2),$$

proto je $\overline{X}$ lepším (vydatnějším) nestranným odhadem než S^2.

Přesnost odhadu

Přesnost bodového odhadu lze měřit pomocí **střední kvadratické chyby** $MSE(T)$ statistiky T.

Definice

Střední kvadratická chyba statistiky T pro odhad parametru θ je definován jako

$$MSE(T) = E(T - \theta)^2 = D(T) + B^2(\theta, T)$$

$$(MSE \text{odhadu} = (\text{rozptyl odhadu} + \text{jeho vychýlení})^2),$$

kde $T - \theta$ je výběrová chyba.

Přesnost odhadu

Střední kvadratická chyba

- charakterizuje, jaká je „průměrná“ výběrová chyba odhadů přicházející v úvahu při všech různých výběrech daného rozsahu.
- je kombinací 2 požadovaných vlastností (malého vychýlení a malého rozptylu), proto je univerzálním kritériem.

Je-li statistika T nestranným odhadem, potom $MSE(T) = D(T)$ (střední kvadratická chyba je rovna rozptylu). Přesnost můžeme měřit pomocí směrodatné odchylky $SE = \sqrt{D(T)}$, která se nazývá **směrodatná (střední) chyba odhadu**.

Přesnost odhadu

Střední kvadratická chyba

- charakterizuje, jaká je „průměrná“ výběrová chyba odhadů přicházející v úvahu při všech různých výběrech daného rozsahu.
- je kombinací 2 požadovaných vlastností (malého vychýlení a malého rozptylu), proto je univerzálním kritériem.

Je-li statistika T nestranným odhadem, potom $MSE(T) = D(T)$ (střední kvadratická chyba je rovna rozptylu). Přesnost můžeme měřit pomocí směrodatné odchylky $SE = \sqrt{D(T)}$, která se nazývá **směrodatná (střední) chyba odhadu**.

Přesnost odhadu

Střední kvadratická chyba

- charakterizuje, jaká je „průměrná“ výběrová chyba odhadů přicházející v úvahu při všech různých výběrech daného rozsahu.
- je kombinací 2 požadovaných vlastností (malého vychýlení a malého rozptylu), proto je univerzálním kritériem.

Je-li statistika T nestranným odhadem, potom $MSE(T) = D(T)$ (střední kvadratická chyba je rovna rozptylu). Přesnost můžeme měřit pomocí směrodatné odchylky $SE = \sqrt{D(T)}$, která se nazývá **směrodatná (střední) chyba odhadu**.

Příklad

Průměr je nestranným odhadem střední hodnoty, proto směrodatná chyba odhadu je rovna směrodatné odchylce výběrového průměru, tj.

$$SE = \sqrt{D(\overline{X})} = \sigma(\overline{X}) = \frac{\sigma(X)}{\sqrt{n}}.$$

Protože $\sigma(X)$ neznáme, odhadneme směrodatnou chybu pomocí výběrové střední chyby

$$\widehat{SE} = \frac{\hat{\sigma}(X)}{\sqrt{n}} = \frac{S}{\sqrt{n}}.$$

Příklad

Spočtěte střední kvadratickou chybu statistiky S^2 a statistiky S_n^2. Uvažujme nejprve statistiku S^2, která je nestranný odhadem parametru σ^2. Platí, že

$$MSE(S^2) = D(S^2) = E(S^2 - \sigma^2)^2 = E(S^4) - 2\sigma^2 E(\sigma^2) + \sigma^4 = \\ = E(S^4) - \sigma^4 = \tfrac{2\sigma^4}{n-1}.$$

Pro střední kvadratickou chybu statistiky S_n^2 dostaneme

$$MSE(S_n^2) = E(S_n^2 - \sigma^2)^2 = E(S_n^4) - 2\tfrac{n-1}{n}\sigma^4 + \sigma^4 = \\ = E(S_n^4) - \tfrac{2-n}{n}\sigma^4 = \tfrac{2n-1}{n^2}\sigma^4,$$

to je méně než $MSE(S^2)$, neboť $\frac{2n-1}{n^2} < \frac{2}{n-1}$. Každý z těchto odhadů je lepší v jiném smyslu.

Příklad

Spočtěte střední kvadratickou chybu statistiky S^2 a statistiky S_n^2. Uvažujme nejprve statistiku S^2, která je nestranný odhadem parametru σ^2. Platí, že

$$MSE(S^2) = D(S^2) = E(S^2 - \sigma^2)^2 = E(S^4) - 2\sigma^2 E(\sigma^2) + \sigma^4 = \\ = E(S^4) - \sigma^4 = \tfrac{2\sigma^4}{n-1}.$$

Pro střední kvadratickou chybu statistiky S_n^2 dostaneme

$$MSE(S_n^2) = E(S_n^2 - \sigma^2)^2 = E(S_n^4) - 2\tfrac{n-1}{n}\sigma^4 + \sigma^4 = \\ = E(S_n^4) - \tfrac{2-n}{n}\sigma^4 = \tfrac{2n-1}{n^2}\sigma^4,$$

to je méně než $MSE(S^2)$, neboť $\frac{2n-1}{n^2} < \frac{2}{n-1}$. Každý z těchto odhadů je lepší v jiném smyslu.

Metody bodových odhadů

Doposud jsme popisovali vlastnosti odhadů, nezabývali jsem se však otázkou, jak odhady získat. Uvedeme 2 nejčastěji používané metody

- metodu momentů,
- metodu maximální věrohodnosti.

Metoda momentů

Uvažujme rozdělení, které závisí na $m \geq 1$ reálných parametrech $\theta_1, \theta_2, \ldots, \theta_m$ a mějme náhodný výběr $X_1, X_2, \ldots, X_n$ z tohoto rozdělení. Předpokládejme, že existují obecné momenty

$$\mu'_r = E(X_i^r) \quad \text{pro} \quad r = 1, 2, \ldots, m.$$

Tyto momenty obecně závisí na parametrech $\theta_1, \theta_2, \ldots, \theta_m$. Výběrové momenty jsou dány vztahem

$$M'_r = \frac{1}{n} \sum_{i=1}^{n} X_i^r, \quad r = 0, 1 \ldots.$$

Metoda momentů

Momentová metoda odhadu parametrů $\theta_1, \theta_2, \ldots, \theta_m$ spočívá v tom, že za jejich odhad vezmeme řešení rovnic

$$\mu_r' = M_r'.$$

Příklad

Odhad parametru λ Poissonova rozdělení.

V případě náhodného výběru z Poissonova rozdělení $Po(\lambda)$ dostaneme rovnici

$$\mu_1' = M_1' \quad \Rightarrow \quad E(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

takže odhadem $\hat{\lambda}$ parametru λ získaným metodou momentů je

$$\hat{\lambda} = \overline{X}.$$

Příklad

Odhad parametrů μ a σ^2 normálního rozdělení $N(\mu, \sigma^2)$.

$$\mu_1' = M_1' \quad \Rightarrow \quad E(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\mu_2' = M_2' \quad \Rightarrow \quad E(X_i^2) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \Leftrightarrow D(X_i) + E(X_i)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Získané odhady jsou

$$\hat{\mu} = \overline{X}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_n^2 = \frac{n-1}{n}S^2$$

Příklad

Odhad parametrů μ a σ^2 normálního rozdělení $N(\mu, \sigma^2)$.

$$\mu_1' = M_1' \quad \Rightarrow \quad E(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\mu_2' = M_2' \quad \Rightarrow \quad E(X_i^2) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \Leftrightarrow D(X_i) + E(X_i)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Získané odhady jsou

$$\hat{\mu} = \overline{X}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_n^2 = \frac{n-1}{n}S^2$$

Příklad

Odhad parametrů μ a σ^2 normálního rozdělení $N(\mu, \sigma^2)$.

$$\mu_1' = M_1' \quad \Rightarrow \quad E(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\mu_2' = M_2' \quad \Rightarrow \quad E(X_i^2) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \Leftrightarrow D(X_i) + E(X_i)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

$$\sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Získané odhady jsou

$$\hat{\mu} = \overline{X}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_n^2 = \frac{n-1}{n} S^2$$

Metoda maximální věrohodnosti

Nechť $X_1, X_2, \ldots, X_n$ je náhodný výběr z rozdělení s hustotou $f(x, \theta)$, resp. s pravěpodobnostní funkcí $p(x, \boldsymbol{\theta})$ obsahující neznámý parametr $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$. Náhodný vektor $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ má sdruženou hustotu rozdělení resp. sdruženou pravděpodobnostní funkci

$$g(\mathbf{x}, \boldsymbol{\theta}) = g(x_1, x_2, \ldots, x_n, \boldsymbol{\theta}) = f(x_1, \boldsymbol{\theta}) f(x_2, \boldsymbol{\theta}) \cdots f(x_n, \boldsymbol{\theta})$$

resp.

$$g(\mathbf{x}, \boldsymbol{\theta}) = g(x_1, x_2, \ldots, x_n, \boldsymbol{\theta}) = p(x_1, \boldsymbol{\theta}) p(x_2, \boldsymbol{\theta}) \cdots p(x_n, \boldsymbol{\theta}).$$

Metoda maximální věrohodnosti

Hustota $g(\mathbf{x}, \boldsymbol{\theta})$ reprezentuje funkci proměnné $\mathbf{x}$ při pevně dané hodnotě $\boldsymbol{\theta}$. Při každé pevné hodnotě $\mathbf{x}$ lze $g(\mathbf{x}, \boldsymbol{\theta})$ chápat jako funkci proměnné $\boldsymbol{\theta}$. Pro tuto funkci budeme používat značení $\mathcal{L}(\boldsymbol{\theta}, \mathbf{x})$ a nazývat jí **věrohodnostní funkce**.

Existuje-li takové $\hat{\boldsymbol{\theta}}$, že pro každé $\boldsymbol{\theta}$ platí

$$\mathcal{L}(\hat{\boldsymbol{\theta}}, \mathbf{x}) \geq \mathcal{L}(\boldsymbol{\theta}, \mathbf{x}),$$

pak $\hat{\boldsymbol{\theta}}$ nazýváme **maximálně věrohodným odhadem** parametru $\boldsymbol{\theta}$.

Místo věrohodnostní funkce je někdy výhodnější pracovat s jejím logaritmem. Potom buhdeme mluvit o **logaritmické věrohodnostní funkci** $L(\theta, \mathbf{x}) = \ln \mathcal{L}(\theta, \mathbf{x})$. Pro maximálně věrohodný odhad také platí

$$L(\hat{\theta}, \mathbf{x}) \geq L(\theta, \mathbf{x}),$$

neboť logaritmus je rostoucí funkcí.

Metoda maximální věrohodnosti

Hustota $g(\mathbf{x}, \boldsymbol{\theta})$ reprezentuje funkci proměnné $\mathbf{x}$ při pevně dané hodnotě $\boldsymbol{\theta}$. Při každé pevné hodnotě $\mathbf{x}$ lze $g(\mathbf{x}, \boldsymbol{\theta})$ chápat jako funkci proměnné $\boldsymbol{\theta}$. Pro tuto funkci budeme používat značení $\mathcal{L}(\boldsymbol{\theta}, \mathbf{x})$ a nazývat jí **věrohodnostní funkce**.

Existuje-li takové $\hat{\boldsymbol{\theta}}$, že pro každé $\boldsymbol{\theta}$ platí

$$\mathcal{L}(\hat{\boldsymbol{\theta}}, \mathbf{x}) \geq \mathcal{L}(\boldsymbol{\theta}, \mathbf{x}),$$

pak $\hat{\boldsymbol{\theta}}$ nazýváme **maximálně věrohodným odhadem** parametru $\boldsymbol{\theta}$.

Místo věrohodnostní funkce je někdy výhodnější pracovat s jejím logaritmem. Potom buhdeme mluvit o **logaritmické věrohodnostní funkci** $L(\boldsymbol{\theta}, \mathbf{x}) = \ln \mathcal{L}(\boldsymbol{\theta}, \mathbf{x})$. Pro maximálně věrohodný odhad také platí

$$L(\hat{\boldsymbol{\theta}}, \mathbf{x}) \geq L(\boldsymbol{\theta}, \mathbf{x}),$$

neboť logaritmus je rostoucí funkcí.

Metoda maximální věrohodnosti

Maximálně věrohodný odhad (není obecně nestranný) vektoru $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ je určen řešením soustavy věrohodnostních rovnic

$$\frac{\partial L(\boldsymbol{\theta}, \mathbf{x})}{\partial \theta_i} = 0, \quad i = 1, 2, \ldots, m.$$

Příklad

Metodou maximální věrohodnosti odhadněte parametr π z alternativního rozdělení $A(\pi)$.

$$\mathcal{L}(\pi, \mathbf{x}) = \pi^{x_1}(1-\pi)^{1-x_1}\pi^{x_2}(1-\pi)^{1-x_2}\dots\pi^{x_n}(1-\pi)^{1-x_n} =$$
$$= \pi^{\sum_{i=1}^n x_i}(1-\pi)^{n-\sum_{i=1}^n x_i}$$

Logaritmická věrohodnostní funkce má tvar

$$L(\pi, \mathbf{x}) = \sum_{i=1}^n x_i \ln \pi + \left(n - \sum_{i=1}^n x_i\right) \ln(1-\pi).$$

Příklad

Položíme-li derivaci této funkce rovnu nule

$$\frac{dL(\pi, \mathbf{x})}{d\pi} = \frac{\sum_{i=1}^{n} x_i}{\pi} - \frac{n - \sum_{i=1}^{n} x_i}{1 - \pi} = 0,$$

dostáváme odhad

$$\hat{\pi} = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}.$$

Příklad

Metodou maximální věrohodnosti odhadněte parametr λ z Poissonova rozdělení $Po(\lambda)$.

$$\mathcal{L}(\lambda, \mathbf{x}) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1! x_2! \cdots x_n!},$$

$$L(\lambda, \mathbf{x}) = \ln \mathcal{L}(\lambda, \mathbf{x}) = -n\lambda + \sum_{i=1}^{n} x_i \ln \lambda - \ln(x_1! x_2! \cdots x_n!)$$

$$\frac{dL(\lambda, \mathbf{x})}{d\lambda} = -n + \sum_{i=1}^{n} x_i \cdot \frac{1}{\lambda} = 0$$

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}.$$

Příklad

Metodou maximální věrohodnosti odhadněte parametry μ a σ^2 normálního rozdělení $N(\mu, \sigma^2)$.

$$\mathcal{L}(\mu, \sigma^2, \mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$L(\mu, \sigma^2, \mathbf{x}) = \ln \mathcal{L}(\mu, \sigma^2, \mathbf{x}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}$$

$$\frac{dL(\mu, \sigma^2)}{d\mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}$$

Příklad

Metodou maximální věrohodnosti odhadněte parametry μ a σ^2 normálního rozdělení $N(\mu, \sigma^2)$.

$$L(\mu, \sigma^2, \mathbf{x}) = \ln \mathcal{L}(\mu, \sigma^2, \mathbf{x}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2\sigma^2}$$

$$\frac{dL(\mu, \sigma^2)}{d\sigma^2} = \frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n}(x_i - \mu) = 0$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n}(x_i - \overline{x})^2$$