4.1.4 Најефикасни оценувачи

Општиот критериум за споредување на оценувачи за ист непознат параметар θ се темели врз споредбата на нивните средно квадратни грешки (MSE), односно нивните дисперзии доколку тие се непристрасни оценувачи (Дефиниција 4.6-4.6а). Подобриот оценувач се смета за **поефикасен оценувач**. Затоа, природно се наметнува задачата да се најде (ако постои) **најефикасен оценувач** меѓу сите непристрасни оценувачи, односно оценувачот со најмала дисперзија или барем да се одреди долната граница за вредностите на дисперзијата на сите можни непристрасни оценувачи за параметарот θ. Постоењето на таков оценувач се разгледува при одрдедени **услови за регуларност**.

Нека $\mathcal{P} = \{p(x,\theta) : \theta \in \Theta\}$ е допустлива фамилија од распределби за обележјето X, каде $p(x,\theta)$ е допустлива густина на распределба (ако X е непрекинато обележје) или допустлива распределба на веројатност (ако X е дискретно обележје). Нека $(X_1, X_2, ..., X_n)$ е примерок кој одговара на обележјето X. Функцијата

$$L(x,\theta) = p(x_1,\theta)p(x_2,\theta)...p(x_n,\theta),\ x = (x_1, x_2, ..., x_n),\ \theta \in \Theta, \tag{4.15}$$

се нарекува **функција на подобност**. Со неа (при фиксирана вредност на θ) е одредена распределбата на примерокот $(X_1, X_2, ..., X_n)$.

Дефиниција 4.10. Нека $f : \Theta \to \mathbb{R}$. Оценувачот $U = U(X_1, X_2, ..., X_n)$ е **регуларен оценувач** за $f(\theta)$ ако важат следните **услови за регуларност**:

(i) множеството $A = \{x = (x_1, x_2, ..., x_n) : L(x,\theta) > 0\}$ не зависи од θ,

(ii) за сите $x \in A$, $L(x,\theta)$ е диференцијабилна по θ,

(iii) може да се диференцира (по θ) под интегралот $\int_{\mathbb{R}^n} L(x,\theta)dx$, односно под сумата $\sum_{x\in\mathbb{R}^n} L(x,\theta)$,

(iv) може да се диференцира (по θ) под интегралот $\int_{\mathbb{R}^n} U(x)L(x,\theta)dx$, односно под сумата $\sum_{x\in\mathbb{R}^n} U(x)L(x,\theta)$,

(v) функцијата $f : \Theta \to \mathbb{R}$ е диференцијабилна.

Теорема 4.3 (Теорема на Рао-Крамер). *Нека* $\mathbf{X} = (X_1, X_2, ..., X_n)$ *е примерок и* $U = U(X_1, X_2, ..., X_n)$ *е непристрасен регуларен оценувач за* $f(\theta)$ *и нека* U *има конечен втор момент. Тогаш, за секој* $\theta \in \Theta$ *важи*

$$D(U) \geq \frac{(f'(\theta))^2}{E(\frac{\partial}{\partial\theta} \ln L(\mathbf{X},\theta))^2} = D_0. \tag{4.16}$$

Равенство важи ако и само ако $\frac{\partial}{\partial\theta} \ln L(\mathbf{X},\theta) = h(\theta)(U - f(\theta))$, *за некоја функција* $h : \Theta \to \mathbb{R}$.

Доказ. Бидејќи $L(x,\theta)$ е функција на подобност и U е непристрасен оценувач за $f(\theta)$, следи дека за секој $\theta \in \Theta$ важат следните равенства

$$\int_{\mathbb{R}^n} L(x,\theta)dx = 1 \text{ и } \int_{\mathbb{R}^n} U(x)L(x,\theta)dx = f(\theta). \tag{4.17}$$

Ако равенствата (4.17) ги диференцираме по θ и користејќи ги условите за регуларност ќе добиеме

$$0 = \frac{\partial}{\partial\theta}\int_{\mathbb{R}^n} L(x,\theta)dx = \int_{\mathbb{R}^n} \frac{\partial L}{\partial\theta}\, dx = \int_{\mathbb{R}^n} (\frac{1}{L}\,\frac{\partial L}{\partial\theta})Ldx = \int_{\mathbb{R}^n} \frac{\partial \ln L}{\partial\theta}\, Ldx, \tag{4.18}$$

$$f'(\theta) = \frac{\partial}{\partial\theta}\int_{\mathbb{R}^n} U(x)L(x,\theta)dx = \int_{\mathbb{R}^n} U\,\frac{\partial L}{\partial\theta}\, dx = \int_{\mathbb{R}^n} U\,\frac{\partial \ln L}{\partial\theta}\, Ldx. \tag{4.19}$$

Од (4.18) и (4.19) имаме

$$\begin{aligned} f'(\theta) &= \int_{\mathbb{R}^n} U\,\frac{\partial \ln L}{\partial\theta}\, Ldx = \\ &= \int_{\mathbb{R}^n} U\,\frac{\partial \ln L}{\partial\theta}\, Ldx - f(\theta)\int_{\mathbb{R}^n} \frac{\partial \ln L}{\partial\theta}\, Ldx = \\ &= \int_{\mathbb{R}^n} (U - f(\theta))\,\frac{\partial \ln L}{\partial\theta}\, Ldx. \end{aligned}$$

Од непристрасноста на U и (4.18) имаме дека $E(U - f(\theta)) = 0$ и $E(\frac{\partial \ln L}{\partial\theta}) = 0$, па затоа

$$\begin{aligned} (f'(\theta))^2 &= \Big(\int_{\mathbb{R}^n} (U - f(\theta))\,\frac{\partial \ln L}{\partial\theta}\, Ldx\Big)^2 \leq \\ &\leq \int_{\mathbb{R}^n} (U - f(\theta))^2 Ldx \cdot \int_{\mathbb{R}^n} (\frac{\partial \ln L}{\partial\theta})^2 Ldx = DU \cdot E(\frac{\partial \ln L}{\partial\theta})^2. \end{aligned}$$

Значи,

$$DU \geq \frac{(f'(\theta))^2}{E(\frac{\partial \ln L}{\partial\theta})^2}$$

т.е. важи (4.16), што требаше да се покаже. ■

Неравенството (4.16) е познато како **неравенство на Рао-Крамер**. При тоа важи

$$E(\frac{\partial}{\partial\theta}\ln L(\mathbf{X},\theta))^2 = n\, I(\theta), \tag{4.20}$$

каде

$$I(\theta) = E(\frac{\partial}{\partial\theta}\ln p(x,\theta))^2, \tag{4.21}$$

е **информација на Фишер**.

Дефиниција 4.11. Регуларниот непристрасен оценувач U за $f(\theta)$ е **најефикасен** ако $D(U) = D_0$, односно ако за неговата дисперзија важи равенство во неравенството на Рао-Крамер.

Дефиниција 4.12. Нека U е регуларен непристрасен оценувач за $f(\theta)$ со конечна дисперзија, тогаш количникот

$$e(U) = \frac{D_0}{D(U)}$$

се нарекува **ефкасност** на оценувачот U.

Јасно е дека важи $0 \leq e(U) \leq 1$. За најефикасен оценувач U важи $e(U) = 1$. Ако $\lim_{n\to\infty} e(U_n) = 1$, тогаш за оценувачот U_n велиме дека е **асимптотски најефикасен оценувач**.

Пример 4.11. Нека $\mathbf{X} = (X_1, X_2, ..., X_n)$ е примерок кој одговара на обележјето $X \sim \mathcal{N}(\theta, \sigma^2)$, каде $\theta > 0$ е непознат параметар, додека σ^2 е познат параметар. Непознатиот параметар го оценуваме со средината на примерокот $\overline{X}_n$. Важат условите за регуларност (покажи!). Да испитаме дали за овој оценувач важи равенство во неравенството на Рао-Крамер. Имаме,

$$L(\mathbf{X}, \theta) = \Big(\frac{1}{2\pi\sigma^2}\Big)^{n/2} \exp\Big\{ -\frac{1}{2\sigma^2}\Big(\sum_{i=1}^n X_i^2 - 2\theta \sum_{i=1}^n X_i + n\theta^2\Big)\Big\},$$

од каде се добива дека

$$\ln L(\mathbf{X}, \theta) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\Big(\sum_{i=1}^n X_i^2 - 2\theta\sum_{i=1}^n X_i + n\theta^2\Big),$$

$$\frac{\partial}{\partial\theta}\ \ln L(\mathbf{X}, \theta) = -\frac{1}{2\sigma^2}\Big(-2\sum_{i=1}^n X_i + 2n\theta\Big) = \frac{n}{\sigma^2}(\overline{X}_n - \theta),$$

$$E\Big(\frac{\partial}{\partial\theta}\ \ln L(\mathbf{X}, \theta)\Big)^2 = \frac{n^2}{\sigma^4}E(\overline{X}_n - \theta)^2 = \frac{n^2}{\sigma^4}E(\overline{X}_n^2 - 2\theta\overline{X}_n + \theta^2).$$

Бидејќи $E(\overline{X}_n) = \theta$ и $D(\overline{X}_n) = \frac{\sigma^2}{n}$, следи дека $E(\overline{X}_n^2) = \frac{\sigma^2}{n} + \theta^2$. Затоа,

$$E\Big(\frac{\partial}{\partial\theta}\ \ln L(x, \theta)\Big)^2 = \frac{n^2}{\sigma^4}\cdot\Big(\frac{\sigma^2}{n} + \theta^2 - 2\theta^2 + \theta^2\Big) = \frac{n}{\sigma^2}.$$

Функцијата $f(\theta) = \theta$ (во неравенството на Рао-Крамер), па затоа $f'(\theta) = 1$, и

$$D_0 = \frac{(f'(\theta))^2}{E(\frac{\partial}{\partial\theta}\ln L(\mathbf{X}, \theta))^2} = \frac{1}{\frac{n}{\sigma^2}} = \frac{\sigma^2}{n} = D(\overline{X}_n),$$

од каде заклучуваме дека $\overline{X}_n$ е најефикасен оценувач за θ.

Од Теорема 4.1 следи следното својство за единственост на најефикасен оценувач.

Својство 4.3. *Најефикасниот оценувач за $f(\theta)$, ако постои, е единствен со веројатност 1.*

Доказ. Нека U и V се два најефикасни оценувачи за $f(\theta)$, тогаш тие се непристрасни оценувачи за $f(\theta)$ и $DU = DV = D_0$, каде D_0 е дефинирано со (4.16). Значи, U и V се оценувачи со минимална дисперзија, па од Теорема 4.1 следи дека $P\{U = V\} = 1$, што требаше да се докаже. ■

Од Теорема 4.3 следи следното својство за линеарна зависност меѓу две функции $f(\theta)$ и $g(\theta)$, ако постојат најефикасни оценувачи за двете од нив.

Својство 4.4. *Ако помеѓу функциите $f(\theta)$ и $g(\theta)$ не постои линеарна зависност, тогаш може да постои најефикасен оценувач за само една од нив.*

Доказ. Нека U е најефикасен оценувач за $f(\theta)$ и V е најефикасен оценувач за $g(\theta)$. Тогаш, од Теорема 4.3 следи дека

$$\frac{\partial}{\partial\theta}\ln L(\mathbf{X},\theta) = h_1(\theta)(U - f(\theta)),\ \frac{\partial}{\partial\theta}\ln L(\mathbf{X},\theta) = h_2(\theta)(V - g(\theta)),$$

од каде

$$h_1(\theta)(U - f(\theta)) = h_2(\theta)(V - g(\theta)),$$

односно

$$h_1(\theta)U - h_2(\theta)V - h_1(\theta)f(\theta) + h_2(\theta)g(\theta) = 0,$$

па U и V се линеарно зависни. ■

4.2 Доволни статистики

Нека $\mathcal{P} = \{F(x,\theta) : \theta \in \Theta\}$ е фамилијата од допустливи распределби за обележјето X, и нека $(X_1, X_2, ..., X_n)$ е примерок кој одговара на обележјето X. Нека $L(x,\theta)$ е распределбата на примерокот $(X_1, X_2, ..., X_n)$, односно функцијата на подобност.

Целта ни е да најдеме статистика $U = U(X_1, X_2, ..., X_n)$ која содржи иста информација за непознатиот параметар θ како и примерокот $(X_1, X_2, ..., X_n)$, односно при оценување на непознатиот параметар со помош на таа статистика да не изгубиме дел од информацијата која со себе ја носи примерокот. Во прилог на ова размислување е поимот за доволност.

Прв кој ја вовел доволноста е Фишер (1920) при оценување на дисперзијата σ^2 кај обележје $X \sim \mathcal{N}(m, \sigma^2)$. Тој ја оценувал σ^2 врз база примерокот $(X_1, X_2, ..., X_n)$ со помош на статистиките

$$U_1 = \sum_{i=1}^{n} |X_i - \overline{X}_n| \text{ и } U_2 = \sum_{i=1}^{n} (X_i - \overline{X}_n)^2,$$

каде $\overline{X}_n$ е средината на примерокот $(X_1, X_2, ..., X_n)$. Покажал дека условната распределба на U_1 при услов $U_2 = t$ не зависи од параметарот σ^2, додека условната распределба на U_2 при услов $U_1 = t$ зависи од σ^2. Со тоа, тој заклучил дека сета информација за σ^2 од примерокот е содржана во статистиката U_2, што значи дека оценувањето на σ^2 со U_1 може да се подобри со користење на информацијата во U_2, додека оценувањето со U_2 не може да биде подобрено користејќи се со U_1. Значи, оценувањето на σ^2 треба да биде базирано на U_2.

Дефиниција 4.13. Статистика $U = U(X_1, X_2, ..., X_n)$ е **доволна статистика** за параметарот θ, ако за секој $\theta \in \Theta$ условната распределба на примерокот $(X_1, X_2, ..., X_n)$ при услов $U = t$ т.е. $L(x, \theta | U = t)$ не зависи од θ.

Пример 4.12. Нека $(X_1, X_2, ..., X_n)$ е примерок кој одговара на обележјето $X \sim \mathcal{B}(m, \theta)$, каде $0 < \theta < 1$ е непознат параметар. Тогаш, распределбата на примерокот $(X_1, X_2, ..., X_n)$ е

$$L(x, \theta) = P((X_1, ..., X_n) = (x_1, ..., x_n)) = \prod_{i=1}^{n} \binom{m}{x_i} \theta^{x_i} (1-\theta)^{m-x_i},$$

за $x_i \in \{0, 1, ..., m\}$, $i = 1, 2, ..., n$. Да провериме дали статистиката $U = X_1 + X_2 + ... + X_n$ е доволна статистика за θ. Од независноста на $X_1, X_2, ..., X_n$ следи дека $U \sim \mathcal{B}(nm, \theta)$.

Нека $(x_1, ..., x_n)$ е произволен елемент од доменот и нека $t = x_1 + ... + x_n$. Тогаш, условната распределба на примерокот при услов $U = t$ е следната.

$$P((X_1, ..., X_n) = (x_1, ..., x_n) | U = t) = \frac{P((X_1, ..., X_n) = (x_1, ..., x_n), U = t)}{P(U = t)} =$$

$$= \frac{P((X_1, ..., X_n) = (x_1, ..., x_n), X_1 + ... + X_n = x_1 + ... + x_n)}{P(U = t)} =$$

$$= \frac{P((X_1, ..., X_n) = (x_1, ..., x_n))}{P(U = t)} = \frac{\prod_{i=1}^{n} \binom{m}{x_i} \theta^{x_i} (1-\theta)^{m-x_i}}{\binom{nm}{t} \theta^t (1-\theta)^{nm-t}} =$$

$$= \frac{\left(\prod_{i=1}^{n} \binom{m}{x_i}\right) \theta^{x_1 + ... + x_n} (1-\theta)^{nm-(x_1+...+x_n)}}{\binom{nm}{t} \theta^t (1-\theta)^{nm-t}} = \frac{\prod_{i=1}^{n} \binom{m}{x_i}}{\binom{nm}{t}},$$

што значи не зависи од θ. Во случај кога $t \neq x_1 + ... + x_n$ имаме дека условната распределба е 0, па тривијално следи дека не зависи од θ. Значи, според дефиницијата за доволна статистика, заклучуваме дека $U = X_1 + ... + X_n$ е доволна статитика за θ.

Проверката со помош на дефиницијата дали една статистика е доволна наидува на проблем при наоѓање на условната распределба посебно кога обележјето X е непрекинато. Постои поедноставен критериум за проверка дали една статистика е доволна - Теоремата за факторизација.

Теорема 4.4 (Теорема за факторизација). *Статистика $U = U(X_1, X_2, ..., X_n)$ е доволна статистика за параметарот θ ако и само ако функцијата на подобност $L(x, \theta)$ може да се запише во облик*

$$L(x, \theta) = g(U(x), \theta) \cdot h(x),$$

каде функцијата h не зависи од параметарот θ.

Доказ. Доказот ќе го споведене за случајна променлива од дискретен тип. Нека U е доволна статистика и нека $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ е таков што $U(x_1, x_2, ..., x_n) = t$. Тогаш,

$$\begin{aligned} L(x,\theta) &= P\{(X_1, X_2, ..., X_n) = x\} = P\{(X_1, X_2, ..., X_n) = x,\ U = t\} = \\ &= P\{U = t\} \cdot P\{(X_1, X_2, ..., X_n) = x \mid U = t\} = \\ &= g(t,\theta) \cdot L(x, \theta \mid U = t) = g(U(x), \theta) \cdot h(x), \end{aligned}$$

затоа што $L(x, \theta \mid U = t)$ не зависи од θ.

Нека сега $L(x, \theta) = g(U(x), \theta) \cdot h(x)$ и нека $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ е таков што $U(x_1, x_2, ..., x_n) = t$. Тогаш,

$$\begin{aligned} L(x, \theta \mid U = t) &= P\{(X_1, X_2, ..., X_n) = x \mid U = t\} = \\ &= \frac{P\{(X_1, X_2, ..., X_n) = x\ U = t\}}{P\{U = t\}} = \frac{L(x,\theta)}{\sum_{y:\ U(y)=t} L(y,\theta)} = \\ &= \frac{g(U(x),\theta) \cdot h(x)}{\sum_{y:\ U(y)=t} g(U(y),\theta) \cdot h(y)} = \frac{h(x)}{\sum_{y:\ U(y)=t} h(y)}, \end{aligned}$$

не зависи од θ, па селди дека U е доволна статистика за θ.

Ако $U(x) \neq t$, тогаш $P\{(X_1, X_2, ..., X_n) = x \mid U = t\} = 0$, и доказот е тривијален. ∎

Пример 4.13. Нека $(X_1, X_2, ..., X_n)$ е примерок кој одговара на обележјето $X \sim \mathcal{U}(0,\theta)$, каде $\theta > 0$ е непознат параметар. Да се обидеме со критериумот за факторизација да најдеме една доволна статистика за θ. Распределбата на примерокот $(X_1, X_2, ..., X_n)$ е

$$L(x,\theta) = \frac{1}{\theta^n},\ 0 < x_i < \theta,\ i = 1, ..., n,$$

па таа може да се запише како

$$L(x,\theta) = \frac{1}{\theta^n} I\{0 < x_1 < \theta, ..., 0 < x_n < \theta\} =$$

$$= \frac{1}{\theta^n} I\{\max_{1\le i\le n} x_i < \theta\} I\{\min_{1\le i\le n} x_i > 0\} =$$
$$= g(\max\{x_1, ..., x_n\}, \theta) h(x),$$

од каде, според теоремата за факторизација, следи дека подредената статистика $X_{(n)} = \max\{X_1, ..., X_n\}$ е доволна статистика за θ.

Според дефиницијата за доволност, доволните статистики не се единствени. Се покажува дека ако U е доволна статистика и g е биекција, тогаш и $g(U)$ е доволна статистика. Со ова се потенцира поголемото значење на просторот на примерокот индуциран од доволната статистика U (т.е. просторот составен од множествата $\{x : U(x) = t\}$), отколку самата доволна статистика.

Својство 4.5. *Ако $U = U(X_1, X_2, ..., X_n)$ е доволна статистика за параметарот θ и g е биекција, тогаш и статистиката $V = g(U)$ е доволна статистика за θ.*

Доказ. Нека U е доволна статистика за θ. Од Теорема 4.4, следи дека

$$L(x,\theta) = g(U(x),\theta) \cdot h(x).$$

Нека $V = g(U)$, при што g е биекција. Нека s е инверзната функција на g т.е. $U = s(V)$. Тогаш,

$$L(x,\theta) = g(s(V(x)),\theta) \cdot h(x) = f_1(V(x),\theta) \cdot h(x),$$

па според Теорема 4.4 следи дека V е доволна статистика за θ. ■

Доволните статистики ни помагаат да најдеме непристрасен оценувач со помала дисперзија од дисперзијата на некој даден непристрасен оценувач, имено тој оценувач е функција од некоја доволна статистика. Исто така важи дека, ако постои оценувач со минимална дисперзија, тогаш тој е функција од некоја доволна статистика.

Ирена Стојковска

Теорема 4.5. *Нека T е доволна статистика за параметарот θ и U е непристрасен оценувач за θ. Тогаш, условното математичко очекување $Q(T) = E(U|T)$ е непристрасен оценувач за θ и за секој $\theta \in \Theta$ важи $D(Q(T)) \leq D(U)$. Равенство важи ако и само ако $P\{U = Q(T)\} = 1$.*

Доказ. Од тоа што T е доволна статистика за θ следи дека условната распределба на U при услов $T = t$ не зависи од θ, а со тоа и $E(U|T) = Q(T)$ не зависи од θ и може да се смета за оценувач за θ. Од непристрасноста на U, имаме

$$E(Q(T)) = E(E(U|T)) = E(U) = \theta,$$

од каде следи дека $Q(T)$ е непристрасен оценувач за θ. Понатаму,

$$D(Q(T)) = D(E(U|T) \leq D(U),$$

а равенство се достигнува акоми само ако $Q(T)$ и U се оценувачи со минимална дисперзија, што според Теорема 4.1 е еквивалентно со $P\{U = Q(T)\} = 1$, што требаше да се докаже. ∎

4.3 Методи за наоѓање на оценки

4.3.1 Метод на моменти

Нека $(X_1, X_2, ..., X_n)$ е случаен примерок кој одговара на обележјето X со функција на распределба F која зависи од непознатите праметри $\theta_1, ..., \theta_r$. Се прашуваме дали може принципот на замена да се примени при оценување на непознатите параметри $\theta_1, ..., \theta_r$ во еден параметарски модел.

Еден пристап (за кој зборевме претходно) е ако може да се изразат $\theta_1, ..., \theta_r$ како функционали од F, а потоа со замена на оценувач $\widehat{F}$ на местото од F се добиваат оценувачи $\widehat{\theta}_1, ..., \widehat{\theta}_r$ за $\theta_1, ..., \theta_r$ соодветно.

Обопштување на овој пристап е да при оценување на $\theta = (\theta_1, ..., \theta_r)$ успееме да најдеме r функционални параметри од F, $\eta_1(F), ..., \eta_r(F)$ кои зависат од θ, односно

$$\eta_k(F) = g_k(\theta), \ k = 1, ..., r, \tag{4.22}$$

каде $g_1, ..., g_r$ се познати функции. На овој начин се добива систем од r равенки со r непознати $\theta_1, ..., \theta_r$. Користејќи го примципот на замена, може да ги оцениме $\eta_1(F), ..., \eta_r(F)$ со $\eta_1(\widehat{F}), ..., \eta_r(\widehat{F})$ соодветно. Ако за секои $\eta_1(F), ..., \eta_r(F)$ постои единствено решение на системот (4.22), тогаш оценувачот $\widehat{\theta}$ на θ го дефинираме така да го задоволува системот

$$\eta_k(\widehat{F}) = g_k(\widehat{\theta}), \ k = 1, ..., r. \tag{4.23}$$