Spis treści

1 Wstęp

1.1 Statystyka matematyczna

STATYSTYKA
nauka poświęcona metodom badania (analizowania) zjawisk masowych; polega na systematyzowaniu obserwowanych cech ilościowych i jakościowych oraz przedstawianiu wyników w postaci zestawień tabelarycznych, wykresów, itp.; posługuje się rachunkiem prawdopodobieństwa.

STATYSTYKA MATEMATYCZNA
dział matematyki stosowanej oparty na rachunku prawdopodobieństwa; zajmuje się badaniem zbiorów na podstawie znajomości własności ich części.

Encyklopedia Popularna PWN, Warszawa 1982

1.2 Literatura

Literatura

- **Silvey S. D.** 1978: Wnioskowanie statystyczne,
  PWN Warszawa
- **Zieliński R.** 1990: Siedem wykładów wprowadzających do statystyki matematycznej, http://www.impan.pl/∼rziel/Books.html
- **Jaworski S., Zieliński W.** 2012: Zbiór zadań z rachunku prawdopodobieństwa i statystyki matematycznej, http://wojtek.zielinski.statystyka.info/dydaktyka.htm

Literatura

- **Krzyśko M.** 1994 Statystyka matematyczna, UAM Poznań
- **Krzyśko M.** 1997 Statystyka matematyczna, Część II, UAM Poznań
- **Rao C. R.** 1982: Modele liniowe statystyki matematycznej, PWN, Warszawa
- **Rao C. R.** 1994: Statystyka i prawda, PWN, Warszawa
- **Zieliński R., Zieliński W.** 1990 Tablice statystyczne, PWN Warszawa
- **Zieliński W.** 2001, Tablice Statystyczne, Wyd. V poprawione i uzupełnione, Fundacja Rozwój SGGW

1.3 Model statystyczny

Model statystyczny: przykład 1

Model probabilistyczny
Rzucamy n–krotnie **symetryczną** monetą.

$$(\{0,1,\ldots,n\}, Bin\,(n,0.5))$$

Model statystyczny
Rzucamy n–krotnie **jakąś** monetą.

$$(\{0,1,\ldots,n\}, \{Bin\,(n,\theta)\,,\ \theta\in[0,1]\})$$

Model statystyczny: przykład 2

Model probabilistyczny
W pewnym gatunku owadów jest 40% osobników męskich. Owady chwytamy do uzyskania ustalonej liczby k osobników męskich.

$$(\{k,k+1,k+2,\ldots\}, NB\,(k,0.4))$$

Model statystyczny
W pewnym gatunku owadów jest **pewien** odsetek osobników męskich. Owady chwytamy do uzyskania ustalonej liczby k osobników męskich.

$$(\{k,k+1,k+2,\ldots\}, \{NB\,(k,\theta)\,,\ \theta\in[0,1]\})$$

Model statystyczny: przykład 3

Model probabilistyczny
Z grupy N osób wśród których jest **znana** liczba M kobiet losujemy n osób.

$$(\{0,1,\ldots,M\}, H\left(N,n,M\right))$$

Model statystyczny
Z grupy N osób wśród których jest **nieznana** liczba M kobiet losujemy n osób.

$$(\{0,1,\ldots,M\}, \{H\left(N,n,M\right),\ M\in\{0,1,\ldots,N\}\})$$

Model statystyczny

Określenie

$$(\mathcal{X}, \mathcal{B}, \{P_\theta,\ \theta\in\Theta\})$$

Problemy

- Ile wynosi θ? (estymacja)
- W zbiorze Θ wyróżniony jest podzbiór Θ_0. Czy $\theta \in \Theta_0$? (weryfikacja hipotez statystycznych)

Model statystyczny

Idea wnioskowania statystycznego

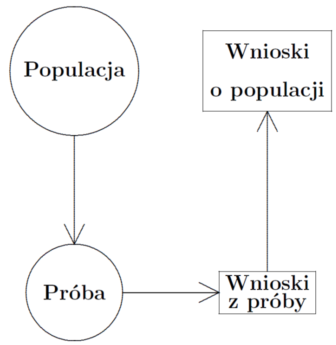
1.4 Preliminaria

Próba

Określenie
niezależne zmienne losowe $X_1, X_2, \ldots, X_n$ o jednakowym rozkładzie P_θ z dystrybuantą F_θ i gęstością p_θ

Dystrybuanta empiryczna

$$F_n(t) = \frac{\#\{1\le j\le n : X_j \le t\}}{n}$$



Dystrybuanta empiryczna

Ważne fakty

- Dla każdego $t \in \mathbb{R}$ zachodzi $E_\theta F_n(t) = F_\theta(t)$
- Dla każdego $t \in \mathbb{R}$ zachodzi $P_\theta\{\lim_{n\to\infty} F_n(t) = F_\theta(t)\} = 1$
- Dla każdego $t \in \mathbb{R}$ rozkład zmiennej losowej

$$\sqrt{n}\frac{F_n(t) - F_\theta(t)}{\sqrt{F_\theta(t)(1 - F_\theta(t))}}$$

  dąży do rozkładu $N(0,1)$ przy $n \to \infty$

Dystrybuanta empiryczna

Podstawowe twierdzenie statystyki matematycznej

Jeżeli próba $X_1, X_2, \ldots, X_n$ pochodzi z rozkładu F_θ, to zmienna losowa

$$D_n = \sup_{-\infty<t<\infty} |F_n(t) - F_\theta(t)|$$

dąży do zera z prawdopodobieństwem 1.

Dystrybuanta empiryczna

Podstawowe twierdzenie statystyki matematycznej - ilustracja

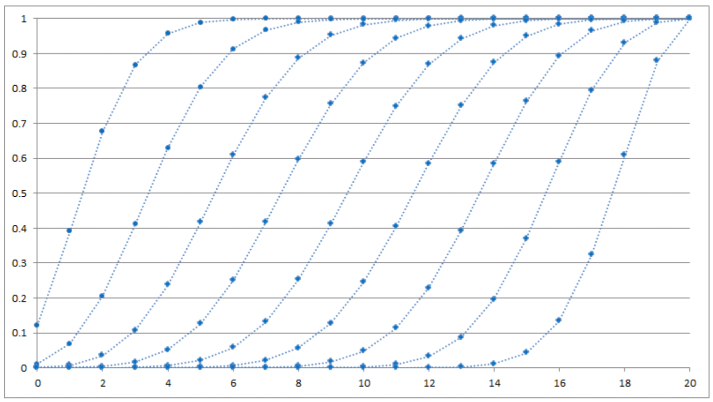
$$(\{0, 1, \ldots, 20\}, \{Bin\,(20, \theta)\,,\ \theta \in [0, 1]\})$$

„Dystrybuanty” dla

$$\theta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$$

Dystrybuanta empiryczna

Podstawowe twierdzenie statystyki matematycznej - ilustracja



Dystrybuanta empiryczna

Podstawowe twierdzenie - ilustracja $n = 10$
Próba: $9, 10, 13, 9, 10, 11, 5, 5, 12, 9$

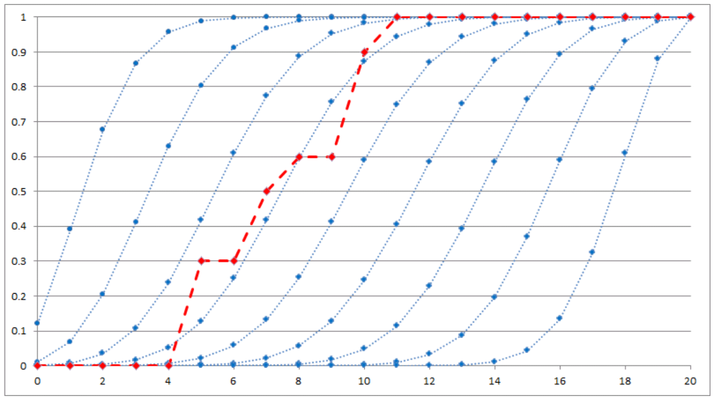
$$F_{10}(t) = \frac{\#\{1 \leq j \leq 10 : X_j \leq t\}}{10} = \begin{cases} 0, & t < 5, \\ 2/10, & t < 9, \\ 5/10, & t < 10, \\ 7/10, & t < 11, \\ 8/10, & t < 12, \\ 9/10, & t < 12, \\ 1, & t \geq 13, \end{cases}$$

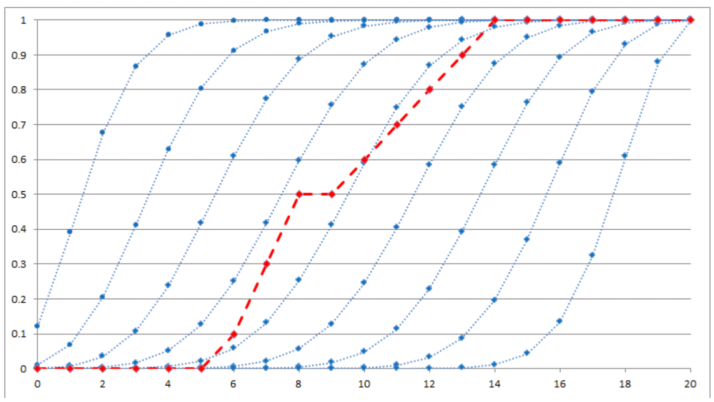
Dystrybuanta empiryczna

Podstawowe twierdzenie - ilustracja $n = 10$

Dystrybuanta empiryczna

Podstawowe twierdzenie - ilustracja $n = 10$





Dystrybuanta empiryczna

Podstawowe twierdzenie - ilustracja $n = 10$

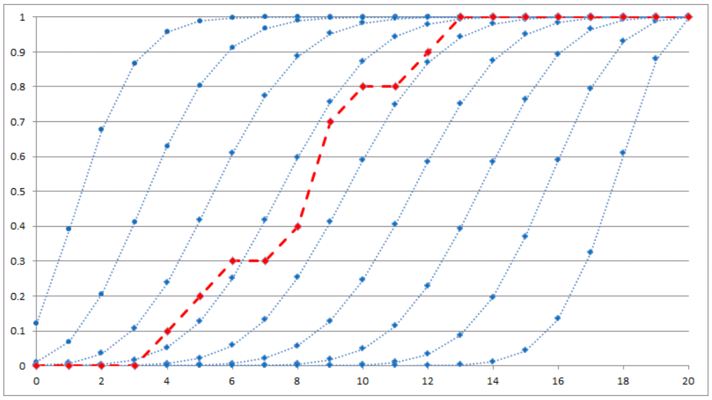
Dystrybuanta empiryczna

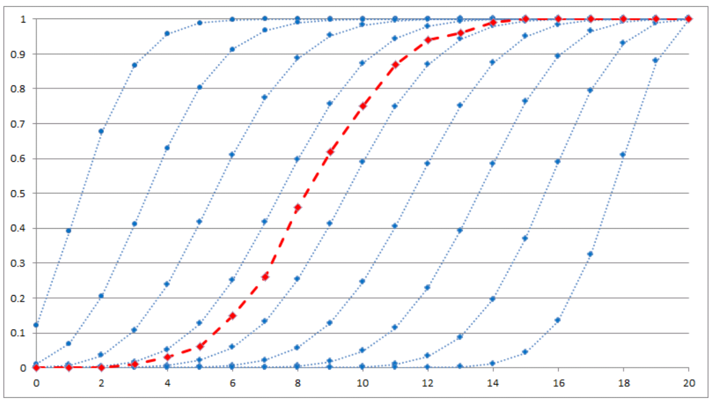
Podstawowe twierdzenie - ilustracja $n = 100$

Dystrybuanta empiryczna

Podstawowe twierdzenie - ilustracja $n = 1000$

Statystyka





Określenie
odwzorowanie $T : (\mathcal{X}, \{P_\theta,\ \theta \in \Theta\}) \to \mathbb{R}^k$
    Statystyka jest zmienną losową!

Estymator
Statystykę taką, że $T(\mathcal{X}) = \Theta$ nazywamy estymatorem parametru θ
    Oznaczenie: $\hat{\theta}$

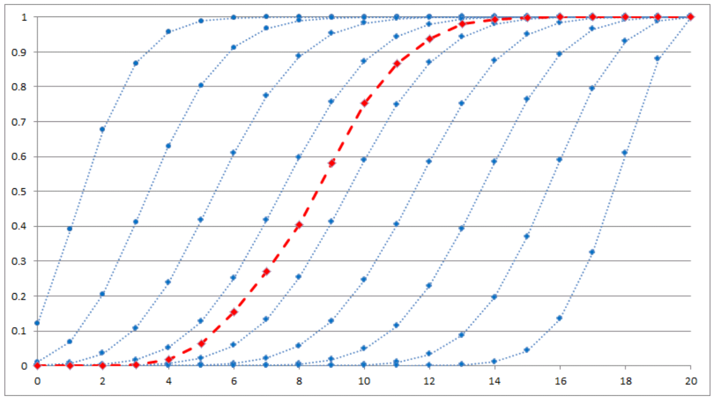
Funkcja straty

Określenie
Funkcja

$$L : \Theta \times \Theta \to \mathbb{R}_+$$

taka, że

$$(\forall \theta \in \Theta) \qquad L(\theta, \theta) = 0$$



Zadanie
Znaleźć taki estymator $\hat{\theta}$, który „minimalizuje” stratę $L(\hat{\theta}(x), \theta)$ dla wszystkich $\theta \in \Theta$

Typowo: minimalizacja $E_\theta L(\hat{\theta}(X), \theta)$

Kwadratowa funkcja straty

Określenie
Funkcja

$$\Theta \times \Theta \ni (\vartheta, \theta) \to L(\vartheta, \theta) = (\vartheta - \theta)^2 \in \mathbb{R}_+$$

Błąd średniokwadratowy estymatora $\hat{\theta}$

$$\begin{aligned} R_\theta(\hat{\theta}) &= E_\theta(\hat{\theta}(X) - \theta)^2 \\ &= E_\theta(\hat{\theta}(X) - E_\theta\hat{\theta}(X))^2 + (E_\theta\hat{\theta}(X) - \theta)^2 \\ &= D^2_\theta\hat{\theta}(X) + (E_\theta\hat{\theta}(X) - \theta)^2 \end{aligned}$$

Estymator nieobciążony

Określenie

$$E_\theta\hat{\theta} = \theta \qquad (\forall \theta \in \Theta)$$

Błąd średniokwadratowy estymatora $\hat{\theta}$
Jeżeli $E_\theta\hat{\theta} = \theta$ $(\forall \theta \in \Theta)$, to

$$R_\theta(\hat{\theta}) = D^2_\theta\hat{\theta}$$

Dostateczność

Określenie
Statystyka T jest dostateczna, jeżeli rozkład warunkowy $P_\theta\{\cdot|T = t\}$ nie jest zależny od θ dla wszystkich $\theta \in \Theta$.

Dostateczność

Przykład

$$(\{0,1\}, \{D(\theta),\ \theta \in [0,1]\})$$

Wnioskujemy o parametrze θ na podstawie próby $X_1, X_2, \ldots, X_5$.

Próba 1: $X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 0, X_5 = 0$

Próba 2: $X_1 = 1, X_2 = 0, X_3 = 1, X_4 = 0, X_5 = 1$

Próba 3: $X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 1, X_5 = 1$

Dostateczność

Przykład cd
Prawdopodobieństwa realizacji poszczególnych prób:

$$P_\theta\{\text{Próba 1}\} = \theta^3(1-\theta)^2$$

$$P_\theta\{\text{Próba 2}\} = \theta^3(1-\theta)^2$$

$$P_\theta\{\text{Próba 3}\} = \theta^3(1-\theta)^2$$

Dostateczność

Przykład cd
Istotna informacja: $X_1 + X_2 + X_3 + X_4 + X_5 (\overset{\text{ozn}}{=\!=} T)$

$$P_\theta\{X_1=x_1, X_2=x_2, X_3=x_3, X_4=x_4, X_5=x_5 | T=t\}$$
$$= \begin{cases} 1/\binom{5}{t}, & \text{jeżeli } x_1 + x_2 + x_3 + x_4 + x_5 = t, \\ 0, & \text{w przeciwnym przypadku.} \end{cases}$$

$T = \sum_{i=1}^{5} X_i$ jest **dostateczna** dla wnioskowania o θ

Dostateczność

Twierdzenie o faktoryzacji
Statystyka T jest dostateczna dla rodziny rozkładów $\{P_\theta : \theta \in \Theta\}$ wtedy i tylko wtedy, gdy gęstość $p_\theta(x)$ może być przedstawiona w postaci

$$p_\theta(x) = g_\theta\{T(x)\}h(x),$$

gdzie funkcja h nie zależy od θ.

Dostateczność

Minimalna statystyka dostateczna
Statystykę dostateczną T nazywamy minimalną statystyką dostateczną, jeżeli dla każdej statystyki dostatecznej S istnieje funkcja h taka, że $T = h(S)$

Zupełność
Statystyka T jest zupełna, jeżeli dla wszystkich $\theta \in \Theta$ zachodzi $E_\theta h(T) = 0$, to $h \equiv 0$

Rodziny wykładnicze

Określenie
Rodzina rozkładów prawdopodobieństwa $\{P_\theta : \theta \in \Theta\}$ nazywa się rodziną wykładniczą, jeżeli każdy rozkład P_θ ma gęstość p_θ o postaci

$$p_\theta(x) = \exp\left\{\sum_{j=1}^{k} c_j(\theta)T_j(x) - b(\theta)\right\} \cdot h(x),$$

gdzie $T_1(x), T_2(x), \ldots, T_k(x)$ są liniowo niezależnymi funkcjami oraz $\{(c_1(\theta), c_2(\theta), \ldots, c_k(\theta)) : \theta \in \Theta\}$ jest pewnym k wymiarowym zbiorem w $\mathbb{R}^k$.

Rodziny wykładnicze

Twierdzenie
Jeżeli $\{P_\theta : \theta \in \Theta\}$ oraz $\Theta \subset \mathbb{R}^k$ jest rodziną wykładniczą rozkładów z gęstościami

$$p_\theta(x) = \exp\left\{\sum_{j=1}^{k} c_j(\theta)T_j(x) - b(\theta)\right\} \cdot h(x),$$

to $(T_1(x), T_2(x), \ldots, T_k(x))$ jest (k–wymiarową) minimalną statystyką dostateczną zupełną.

2 ENMW

2.1 Określenie

Estymatory nieobciążone o minimalnej wariancji

Problem
Model statystyczny $(\mathcal{X}, \mathcal{B}, \{P_\theta,\ \theta \in \Theta\})$
Niech $g : \Theta \to \mathbb{R}^1$ będzie znaną funkcją
Zadanie: oszacować nieznaną wartość $g(\theta)$

Estymatory nieobciążone o minimalnej wariancji

Estymator wielkości $g(\theta)$
Wybrać takie $\delta(X_1, X_2, \ldots, X_n)$ by $(\forall \theta \in \Theta)$

$$E_\theta \delta(X_1, X_2, \ldots, X_n) = g(\theta)$$

$$D^2_\theta(\delta) = E_\theta \left(\delta(X_1, X_2, \ldots, X_n) - g(\theta)\right)^2 = \min$$

2.2 Twierdzenie Rao-Blackwella

Estymatory nieobciążone o minimalnej wariancji

Twierdzenie (*Rao–Blackwella)*
Jeżeli $\hat{g}$ jest estymatorem nieobciążonym i jeżeli T jest statystyką dostateczną zupełną, to $E_\theta(\hat{g}|T)$ jest również estymatorem nieobciążonym o jednostajnie nie większej wariancji niż wariancja estymatora $\hat{g}$.

Estymatory nieobciążone o minimalnej wariancji

Twierdzenie
Jeżeli T jest statystyką dostateczną zupełną i jeżeli dla danej funkcji g istnieje funkcja $\hat{g}$ taka, że

$$(\forall \theta \in \Theta)\, E_\theta \hat{g}(T) = g(\theta),$$

to $\hat{g}(T)$ jest $ENMW[g(\theta)]$.

2.3 Model dwupunktowy

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja parametru θ
Model pojedynczej obserwacji X:

$$(\{0,1\}, \{D(\theta),\ \theta \in [0,1]\})$$

Rodzina $\{D(\theta),\ \theta \in [0,1]\}$ jest rodziną wykładniczą:

$$\begin{aligned} p_\theta(x) &= \theta^x(1-\theta)^{1-x} \\ &= \exp\left\{x \log\frac{\theta}{1-\theta} + \log(1-\theta)\right\},\ x = 0,1 \end{aligned}$$

Statystyka dostateczna: $T(x) = x$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja parametru θ
Model dla próby $X_1, X_2, \ldots, X_n$:

$$(\{0,1\}, \{D(\theta),\ \theta \in [0,1]\})^n$$

$$\begin{aligned} p_\theta(x_1,\ldots,x_n) &= \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i} \\ &= \exp\left\{\log\frac{\theta}{1-\theta}\sum_{i=1}^n x_i + \log(1-\theta)\right\} \end{aligned}$$

Statystyka dostateczna: $T = \sum X_i$.

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja parametru θ
Model dla statystyki T

$$(\{0,1,\ldots,n\}, \{Bin\,(n,\theta)\,,\ \theta \in [0,1]\})$$

Funkcja g: $g(\theta) = \theta$

Ponieważ

$$E_\theta T = n\theta,$$

więc

$$ENMW[\theta] = \frac{T}{n}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja wariancji $\theta(1-\theta)$
Funkcja g: $g(\theta) = \theta(1-\theta)$
Wyznaczyć $\hat{g}$ taką, że $\forall\theta \in \Theta$

$$E_\theta \hat{g}(T) = \theta(1-\theta)$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja wariancji $\theta(1-\theta)$

$$E_\theta \hat{g}(T) = \sum_{t=0}^{n} \hat{g}(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} = \theta(1-\theta)$$

$$(1-\theta)^n \sum_{t=0}^{n} \hat{g}(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t = \theta(1-\theta)$$

$$\sum_{t=0}^{n} \hat{g}(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t = \left(\frac{\theta}{1-\theta}\right) \left(1 + \left(\frac{\theta}{1-\theta}\right)\right)^{n-2}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model dwupunktowy: estymacja wariancji $\theta(1-\theta)$
Podstawienie: $v = \theta/(1-\theta)$

$$\sum_{t=0}^{n} \hat{g}(t) \binom{n}{t} v^t = v\,(1+v)^{n-2}$$

$$\sum_{t=0}^{n} \hat{g}(t) \binom{n}{t} v^t = \sum_{t=0}^{n-2} \binom{n-2}{t} v^{t+1}$$

$$ENMW[\theta(1-\theta)] = \frac{T(n-T)}{n(n-1)}$$

2.4 Model Poissona

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X = 0\} = e^{-\theta}$
Model pojedynczej obserwacji X:

$$(\{0, 1, 2, \ldots\}, \{Po(\theta),\ \theta \in \mathbb{R}_+\})$$

Rodzina $\{Po(\theta),\ \theta \in \mathbb{R}_+\}$ jest rodziną wykładniczą:

$$\begin{aligned} p_\theta(x) &= \frac{\theta^x}{x!} e^{-\theta} \\ &= \exp\left\{x \log\theta - \theta\right\} \frac{1}{x!},\ x = 0, 1, 2, \ldots \end{aligned}$$

Statystyka dostateczna: $T(x) = x$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X = 0\} = e^{-\theta}$
Model dla próby $X_1, X_2, \ldots, X_n$:

$$(\{0, 1, 2, \ldots\}, \{Po(\theta),\ \theta \in \mathbb{R}_+\})^n$$

$$\begin{aligned} p_\theta(x_1, \ldots, x_n) &= \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta} \\ &= \exp\left\{\log\theta \sum_{i=1}^{n} x_i - n\theta\right\} \frac{1}{\prod x_i!} \end{aligned}$$

Statystyka dostateczna: $T = \sum X_i$.

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X = 0\} = e^{-\theta}$
Model dla statystyki T

$$(\{0, 1, 2, \ldots\}, \{Po\,(n\theta)\,,\ \theta \in \mathbb{R}_+\})$$

Funkcja g: $g(\theta) = e^{-\theta} \overset{\text{ozn}}{=\!=} \lambda$

- estymator nieobciążony λ^*
- $\hat{\lambda} = E(\lambda^* | T)$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X=0\}=e^{-\theta}$

$$Y_j=\begin{cases}1, & \text{jeżeli } X_j=0,\\ 0, & \text{jeżeli } X_j>0.\end{cases}$$

$$\lambda^*=\frac{1}{n}\sum_{j=1}^{n}Y_j$$

Estymator λ^* jest nieobciążony

$$E_\theta Y_j=1\cdot P_\theta\{X=0\}+0\cdot P_\theta\{X>0\}=e^{-\theta}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X=0\}=e^{-\theta}$
Wyznaczamy $\hat{\lambda}=E_\theta(\lambda^*|T)$

$$\begin{aligned}&E_\theta(\lambda^*|T=t)=E_\theta\left(\frac{1}{n}\sum_{j=1}^{n}Y_j|T=t\right)\\&=E_\theta\left(Y_1|T=t\right)=P_\theta\{X_1=0|T=t\}\\&=\sum_{x_2+\cdots+x_n=t}P_\theta\{X_1=0,X_2=x_2,\ldots,X_n=x_n|T=t\}\\&\overset{\text{ozn}}{=\!=}\heartsuit\end{aligned}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X=0\}=e^{-\theta}$

$$\begin{aligned}&P_\theta\{X_1=x_1,X_2=x_2,\ldots,X_n=x_n|T=t\}\\&\quad=\frac{P_\theta\{X_1=x_1,X_2=x_2,\ldots,X_n=x_n,T=t\}}{P_\theta\{T=t\}}\\&\quad=\begin{cases}0, & \text{jeżeli } \sum x_i\neq t\\ \frac{\theta^{x_1+\cdots x_n}e^{-n\theta}}{x_1!\ldots x_n!}\cdot\frac{t!}{(n\theta)^te^{-n\theta}}, & \text{jeżeli } \sum x_i=t\end{cases}\\&\quad=\begin{cases}0, & \text{jeżeli } \sum x_i\neq t\\ \frac{t!}{n^tx_1!\ldots x_n!}, & \text{jeżeli } \sum x_i=t\end{cases}\end{aligned}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X=0\}=e^{-\theta}$
Ponieważ

$$(\alpha_1+\cdots+\alpha_n)^t=\sum_{x_1+\cdots+x_n=t}\frac{t!}{x_1!\ldots x_n!}\alpha_1^{x_1}\cdots\alpha_n^{x_n}$$

więc

$$\sum_{x_1+\cdots+x_n=t}\frac{t!}{x_1!\ldots x_n!}=n^t$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X=0\}=e^{-\theta}$
Zatem

$$\heartsuit=\sum_{x_2+\cdots+x_n=t}\frac{t!}{n^t x_2!\ldots x_n!}=\frac{(n-1)^t}{n^t}$$

Czyli

$$\hat{\lambda}=\left(1-\frac{1}{n}\right)^T$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: $D^2_\lambda\hat{\lambda}$ **vs** $D^2_\lambda\lambda^*$

$$D^2_\lambda\hat{\lambda}=E_\lambda\hat{\lambda}^2-\lambda^2$$

$$E_\theta\hat{\lambda}^2=\sum_{t=0}^{\infty}\left[\left(1-\frac{1}{n}\right)^t\right]^2\frac{(n\theta)^t}{t!}e^{-n\theta}=e^{-n\theta}\sum_{t=0}^{\infty}\frac{\left[\frac{(n-1)^2}{n}\theta\right]^t}{t!}$$
$$=e^{-n\theta}e^{\frac{(n-1)^2}{n}\theta}=e^{-\left(2+\frac{1}{n}\right)\theta}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: $D^2_\lambda\hat{\lambda}$

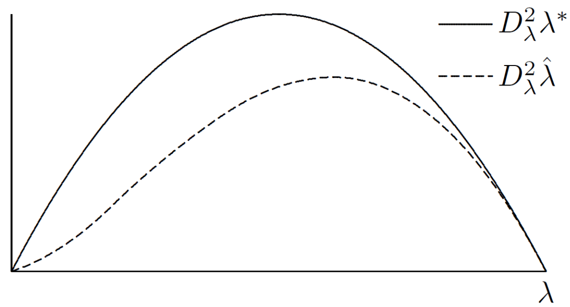
$$D^2_\lambda\hat{\lambda}=\lambda^{\left(2-\frac{1}{n}\right)}-\lambda^2$$

Model Poissona: $D^2_\lambda\lambda^*$

$$D^2_\lambda\lambda^*=\frac{\lambda(1-\lambda)}{n}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: $D^2_\lambda \hat{\lambda}$ **vs** $D^2_\lambda \lambda^*$



Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X = 0\} = e^{-\theta}$

Niech $\varepsilon > 0$

$$P_\lambda\{|\lambda^* - \lambda| < \varepsilon\} = ? \qquad P_\lambda\{|\hat{\lambda} - \lambda| < \varepsilon\} = ?$$

$$\lambda^* = \frac{1}{n}\sum_{i=1}^{n} Y_i; \quad \sum_{i=1}^{n} Y_i \sim Bin\left(n, \lambda = e^{-\theta}\right)$$

$$\hat{\lambda} = \left(1 - \frac{1}{n}\right)^T; \quad T = \sum_{i=1}^{n} X_i \sim Po\left(n\theta = -n \log \lambda\right)$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model Poissona: estymacja $P_\theta\{X = 0\} = e^{-\theta}$ **(**$\varepsilon = 0.05; n = 50$**)**

λ	$P_\lambda\{\lvert\lambda^* - \lambda\rvert < \varepsilon\}$	$P_\lambda\{\lvert\hat{\lambda} - \lambda\rvert < \varepsilon\}$
0.1	0.7661	0.9793
0.2	0.6235	0.8358
0.3	0.5593	0.7275
0.4	0.5291	0.6256
0.5	0.5201	0.6052
0.6	0.5291	0.5719
0.7	0.5593	0.5912
0.8	0.6235	0.6240
0.9	0.7661	0.8092

2.5 Model gaussowski

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów
Model pojedynczej obserwacji X:

$$\left(\mathbb{R}, \left\{N\left(\mu, \sigma^2\right),\ \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\right\}\right)$$

Rodzina $\{N(\mu, \sigma^2), \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\}$ jest wykładnicza

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów

$$\begin{aligned} f_{\mu,\sigma}(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2\sigma^2}x^2 + \frac{\mu}{\sigma^2}x - \left(\frac{\mu^2}{2\sigma^2} + \log\left(\sigma\sqrt{2\pi}\right)\right)\right\} \end{aligned}$$

Statystyka dostateczna: $(T_1(x), T_2(x)) = (x, x^2)$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów
Model dla próby $X_1, X_2, \ldots, X_n$:

$$\left(\mathbb{R}, \left\{N\left(\mu, \sigma^2\right),\ \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\right\}\right)^n =$$
$$\left(\mathbb{R}^n, \left\{N_n\left(\mu\mathbf{1}_n, \sigma^2\boldsymbol{I}_n\right),\ \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\right\}\right)$$

Statystyka dostateczna: $(T_1, T_2) = (\sum X_i, \sum X_i^2)$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów
Niech

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

$$S^2 = \begin{cases} \sum_{i=1}^{n}(X_i - \mu)^2, & \mu \text{ jest znane,} \\ \sum_{i=1}^{n}(X_i - \bar{X})^2, & \mu \text{ nie jest znane.} \end{cases}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów
Zmienna losowa $\bar{X}$ ma rozkład $N\left(\mu, \frac{\sigma^2}{n}\right)$

Zmienna losowa S^2/σ^2 ma rozkład chi–kwadrat z ν stopniami swobody:

$$\nu = \begin{cases} n, & \mu \text{ jest znane,} \\ n-1, & \mu \text{ nie jest znane.} \end{cases}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów

$$E_{\mu,\sigma} S^\alpha = \begin{cases} \dfrac{\sigma^\alpha}{K_{\nu,\alpha}}, & \nu + \alpha > 0, \\ \infty, & \nu + \alpha \le 0, \end{cases}$$

gdzie $K_{\nu,\alpha} = \Gamma\left(\frac{\nu}{2}\right) / \left(2^{\frac{\alpha}{2}}\Gamma\left(\frac{\nu+\alpha}{2}\right)\right)$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski: estymacja parametrów
Jeżeli μ oraz σ nie są znane, to

- $ENMW[\mu] = \bar{X}$
- $ENMW[\sigma^2] = \frac{1}{n-1}S^2$

- $ENMW[\sigma] = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{2}\Gamma\left(\frac{n}{2}\right)} S$
- $ENMW\left[\frac{\mu}{\sigma}\right] = \frac{\sqrt{2}\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}-1\right)} \frac{\bar{X}}{S}$

2.6 Nierówność Cramera-Rao

Estymatory nieobciążone o minimalnej wariancji

Informacja
Ilością informacji o θ zawartą w X nazywamy wielkość $I_\theta = E_\theta[\{\partial \log p_\theta(X)/\partial\theta\}^2]$

Estymatory nieobciążone o minimalnej wariancji

Nierówność Cramera-Rao
Niech $\{P_\theta : \theta \in \Theta\}$ będzie rodziną rozkładów, niech θ będzie parametrem liczbowym i niech Θ będzie przedziałem na prostej. Zakładamy, że dla każdego θ rozkład P_θ ma gęstość p_θ. Jeżeli spełnione są pewne warunki regularności, to nierówność

$$D_\theta^2 \theta^* \geq I_\theta^{-1}$$

spełniona jest dla każdego estymatora nieobciążonego θ^* parametru θ

Estymatory nieobciążone o minimalnej wariancji

Efektywność
Liczbę

$$\text{eff}(\theta^*) = \frac{I_\theta^{-1}}{D_\theta^2 \theta^*}$$

nazywamy *efektywnością* estymatora θ^*

Lemat
Jeżeli spełnione są warunki regularności, to

$$I_\theta = E_\theta\left[-\frac{\partial^2}{\partial\theta^2} \log p_\theta(X)\right]$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski

Obliczamy I_μ

$$p_\mu(\boldsymbol{x}) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2\right\}$$

$$\log p_\mu(\boldsymbol{x}) = -n\log\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^n (X_i-\mu)^2$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Model gaussowski

$$\frac{\partial}{\partial\mu}\log p_\mu(\boldsymbol{x}) = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i-\mu)$$

$$\frac{\partial^2}{\partial\mu^2}\log p_\mu(\boldsymbol{x}) = -\frac{n}{\sigma^2}$$

Zatem

$$I_\mu = \frac{n}{\sigma^2}$$

Ponieważ $D^2\bar{X} = \sigma^2/n$, więc

$$\text{eff}(\bar{X}) = 1$$

2.7 Antyprzykłady

Estymatory nieobciążone o minimalnej wariancji - przykład

Zły estymator - ucięty rozkład Poissona

Model pojedynczej obserwacji X:

$$(\{1,2,\ldots\}, \{P_\theta : \theta > 0\})$$

$$P_\theta\{X = x\} = \frac{\theta^x e^{-\theta}}{x!(1-e^{-\theta})}$$

Zadanie: oszacować $e^{-\theta}$ na podstawie jednej obserwacji

Estymatory nieobciążone o minimalnej wariancji - przykład

Zły estymator - ucięty rozkład Poissona
$T(X)$ jest $ENMW[e^{-\theta}]$, jeżeli

$$\sum_{x=1}^{\infty} T(x) \frac{\theta^x e^{-\theta}}{x!(1-e^{-\theta})} = e^{-\theta}$$

Rozwiązanie:

$$T(x) = (-1)^{x+1}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Nie istnieje ENMW
Model pojedynczej obserwacji X:

$$(\mathbb{Z}, \{P_\theta : \theta \in \mathbb{Z}\})$$

$$P_\theta\{X = \theta - 1\} = P_\theta\{X = \theta\} = P_\theta\{X = \theta + 1\} = \frac{1}{3}$$

Estymator nieobciążony: $\hat{\theta}(X) = X$
Wariancja: $D^2\hat{\theta} = 2/3$

Estymatory nieobciążone o minimalnej wariancji - przykład

Nie istnieje ENMW
Niech $a_0 + a_1 + a_2 = 0$ oraz

$$\delta(x) = \begin{cases} a_0, & \text{mod}(x;3) = 0 \\ a_1, & \text{mod}(x;3) = 1 \\ a_2, & \text{mod}(x;3) = 2 \end{cases}$$

Niech $\theta^*(X) = X + \delta(X)$

Estymatory nieobciążone o minimalnej wariancji - przykład

Nie istnieje ENMW

$$E_\theta \theta^* = \theta \quad (\forall \theta \in \Theta)$$

$$D^2_\theta \theta^* = \begin{cases} ((a_2 - 1)^2 + a_0^2 + (a_1 + 1)^2)/3, & \text{mod}(\theta;3) = 0 \\ ((a_0 - 1)^2 + a_1^2 + (a_2 + 1)^2)/3, & \text{mod}(\theta;3) = 1 \\ ((a_1 - 1)^2 + a_2^2 + (a_0 + 1)^2)/3, & \text{mod}(\theta;3) = 2 \end{cases}$$

Estymatory nieobciążone o minimalnej wariancji - przykład

Nie istnieje ENMW

θ	$D^2_\theta\hat{\theta} = 2/3$	$D^2_\theta\theta^*$ $a_0 = 0.5, a_1 = 0.5, a_2 = -1$
0	0.6667	2.1667
1	0.6667	0.1667
2	0.6667	1.1667
3	0.6667	2.1667
4	0.6667	0.1667
5	0.6667	1.1667

3 ENW

3.1 Przykład wstępny

ENW: przykład wstępny

Wykonano 50 rzutów nieznaną monetą i otrzymano 32 orły.

Jakie jest prawdopodobieństwo wyrzucenia orła w pojedynczym rzucie?

ENW: przykład wstępny

Model statystyczny rzutów monetą

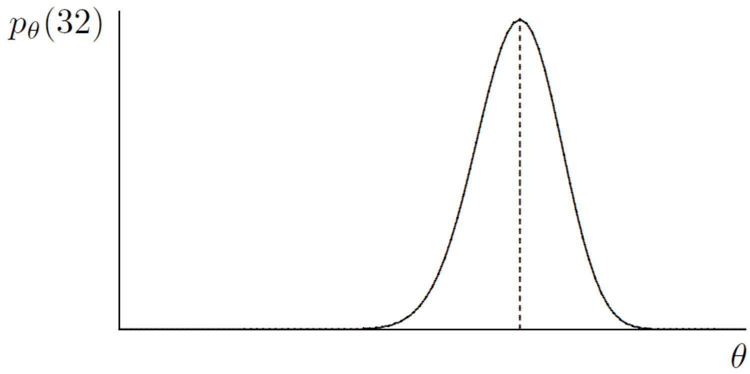
$$(\{0, 1, \ldots, 50\}, \{Bin\,(50, \theta)\,,\ \theta \in [0, 1]\})$$

Prawdopodobieństwo uzyskania 32 orłów w 50 rzutach

$$p_\theta(32) = \binom{50}{32}\theta^{32}(1-\theta)^{50-32}$$

Dla jakiego θ uzyskanie 32 orłów jest najbardziej prawdopodobne?

ENW: przykład wstępny



ENW**: przykład wstępny**

Zasada wiarogodności
Wybrać tę wartość parametru θ, dla której uzyskanie 32 orłów w 50 rzutach jest najbardziej prawdopodobne

tzn. wartość θ maksymalizującą $p_\theta(32)$

3.2 Określenie

Estymatory największej wiarogodności

Wiarogodność
Model statystyczny $(\mathcal{X}, \mathcal{B}, \{P_\theta,\ \theta \in \Theta\})$

Dla ustalonego $x \in \mathcal{X}$ wielkość

$$L(\theta; x) = p_\theta(x)$$

nazywamy wiarogodnością parametru θ.

Estymatory największej wiarogodności

Estymator Największej Wiarogodności
Jeżeli przy każdym ustalonym $x \in \mathcal{X}$ istnieje $\hat{\theta} = \hat{\theta}(x) \in \Theta$ takie, że

$$L(\hat{\theta}; x) \geq L(\theta; x) \quad (\forall \theta \in \Theta),$$

to odwzorowanie $\hat{\theta} : \mathcal{X} \to \Theta$ nazywamy estymatorem największej wiarogodności.

Estymatory największej wiarogodności

Twierdzenie
Jeżeli $h : \Theta \to \Theta$ jest funkcją różnowartościową oraz $\hat{\theta}$ jest $ENW[\theta]$, to $h(\hat{\theta})$ jest $ENW[h(\theta)]$.

ENW funkcji parametrycznej
Jeżeli $h : \Theta \to \Theta$, to $ENW[h(\theta)]$ określamy jako $h(\hat{\theta})$, gdzie $\hat{\theta}$ jest $ENW[\theta]$.

3.3 Model dwupunktowy

Estymator Największej Wiarogodności - przykład

Model dwupunktowy: $ENW[\theta]$
Zaobserwowano $X_1 = x_1, \ldots, X_n = x_n$.
Wiarogodność parametru θ:

$$L(\theta; x_1, \ldots, x_n) = \binom{n}{t} \theta^t (1-\theta)^{n-t} \qquad \left(t = \sum_{i=1}^{n} x_i \right)$$

Znaleźć $\hat{\theta}$ maksymalizujące wiarogodność (t jest ustalone)

Estymator Największej Wiarogodności - przykład

Model dwupunktowy: $ENW[\theta]$
$\hat{\theta}$ jest rozwiązaniem równania

$$\frac{dL(\theta; x_1, \ldots, x_n)}{d\theta} = 0$$

Estymator Największej Wiarogodności - przykład

Model dwupunktowy: $ENW[\theta]$
Zamiast $L(\theta; x_1, \ldots, x_n) = \binom{n}{t} \theta^t (1-\theta)^{n-t}$ łatwiej

$$\begin{aligned} \mathcal{L}(\theta; x_1, \ldots, x_n) &= \log L(\theta; x_1, \ldots, x_n) \\ &= \log \binom{n}{t} + t \log \theta + (n-t) \log(1-\theta) \end{aligned}$$

Estymator Największej Wiarogodności - przykład

Model dwupunktowy: $ENW[\theta]$
Mamy

$$\frac{d\mathcal{L}}{d\theta} = \frac{t}{\theta} - \frac{n-t}{1-\theta} = 0$$

$$ENW[\theta] = \frac{T}{n}$$

Estymator Największej Wiarogodności - przykład

Model dwupunktowy: $ENW[\theta(1-\theta)]$

$$ENW[\theta(1-\theta)] = \hat{\theta}(1-\hat{\theta})$$

gdzie

$$\hat{\theta} = ENW[\theta]$$

3.4 Model Poissona

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\theta]$
Zaobserwowano $X_1 = x_1, \ldots, X_n = x_n$.

Wiarogodność parametru θ:

$$L(\theta; x_1, \ldots, x_n) = \frac{(n\theta)^t}{x_1! \cdots x_n!} e^{-(n\theta)} \qquad \left(t = \sum_{i=1}^{n} x_i\right)$$

Znaleźć $\hat{\theta}$ maksymalizujące wiarogodność (t jest ustalone)

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\theta]$
$\hat{\theta}$ jest rozwiązaniem równania

$$\frac{dL(\theta; x_1, \ldots, x_n)}{d\theta} = 0$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\theta]$
Zamiast $L(\theta; x_1, \dots, x_n) = \frac{(n\theta)^t}{x_1! \cdots x_n!} e^{-(n\theta)}$ łatwiej

$$\begin{aligned}\mathcal{L}(\theta; x_1, \dots, x_n) &= \log L(\theta; x_1, \dots, x_n) \\ &= t \log n + t \log \theta - \sum_{i=1}^{n} \log(x_i!) - n\theta\end{aligned}$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\theta]$
Mamy

$$\frac{d\mathcal{L}}{d\theta} = \frac{t}{\theta} - n = 0$$

$$ENW[\theta] = \frac{T}{n}$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\lambda = e^{-\theta}]$

$$ENW[\lambda] = e^{-\tilde{\theta}} (\overset{\text{ozn}}{=\!=} \tilde{\lambda})$$

gdzie

$$\tilde{\theta} = ENW[\theta]$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $ENW[\lambda]$ **vs** $ENMW[\lambda]$
Ryzyko $\tilde{\lambda} = ENW[\lambda]$

$$R_{\tilde{\lambda}}(\lambda) = E_\theta(\tilde{\lambda} - \lambda)^2 = E_\theta \tilde{\lambda}^2 - 2\lambda E_\theta \tilde{\lambda} + \lambda^2$$

Ryzyko $\hat{\lambda} = ENMW[\lambda]$

$$R_{\hat{\lambda}}(\lambda) = E_\theta(\hat{\lambda} - \lambda)^2 = D_\theta^2 \hat{\lambda}^2$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $R_{\tilde{\lambda}}(\lambda)$

$$E_\theta \tilde{\lambda}^2 = \sum_{t=0}^{\infty} \left(e^{-\frac{t}{n}}\right)^2 \frac{(n\theta)^t}{t!} e^{-n\theta} = e^{-n\theta} \sum_{t=0}^{\infty} \frac{(n\theta e^{-\frac{2}{n}})^t}{t!}$$
$$= \exp\left\{-n\theta\left(1 - e^{-\frac{2}{n}}\right)\right\} = \lambda^{n(1-e^{-\frac{2}{n}})}$$
$$E_\theta \tilde{\lambda} = \sum_{t=0}^{\infty} e^{-\frac{t}{n}} \frac{(n\theta)^t}{t!} e^{-n\theta} = e^{-n\theta} \sum_{t=0}^{\infty} \frac{(n\theta e^{-\frac{1}{n}})^t}{t!}$$
$$= \exp\left\{-n\theta\left(1 - e^{-\frac{1}{n}}\right)\right\} = \lambda^{n(1-e^{-\frac{1}{n}})}$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $R_{\tilde{\lambda}}(\lambda)$

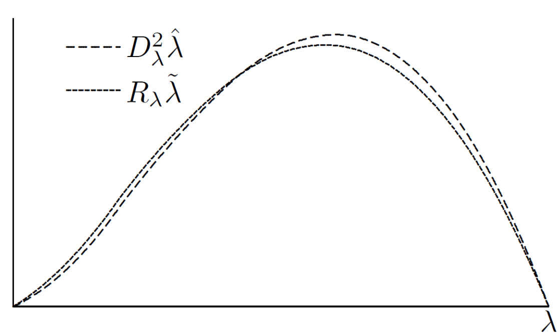
$$R_{\tilde{\lambda}}(\lambda) = \lambda^{n(1-e^{-\frac{2}{n}})} - 2\lambda^{1+n(1-e^{-\frac{1}{n}})} + \lambda^2$$

Model Poissona: $R_{\hat{\lambda}}(\lambda)$

$$R_{\hat{\lambda}}(\lambda) = D^2_\lambda \hat{\lambda} = \lambda^{\left(2-\frac{1}{n}\right)} - \lambda^2$$

Estymator Największej Wiarogodności - przykład

Model Poissona: $R_{\tilde{\lambda}}(\lambda)$ **vs** $R_{\hat{\lambda}}(\lambda)$



3.5 Model gaussowski

Estymator Największej Wiarogodności - przykład

Model gaussowski: $ENW[\mu]$ **i** $ENW[\sigma^2]$
Zaobserwowano $X_1 = x_1, \ldots, X_n = x_n$.
Wiarogodność parametru (μ, σ^2):

$$L(\mu, \sigma^2; x_1, \ldots, x_n) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2\right\}$$

Znaleźć $\hat{\mu}$ i $\hat{\sigma^2}$ maksymalizujące wiarogodność

Estymator Największej Wiarogodności - przykład

Model gaussowski:$ENW[\mu]$ **i** $ENW[\sigma^2]$
$(\hat{\mu}, \hat{\sigma^2})$ jest rozwiązaniem równania

$$\begin{cases} \frac{\partial L(\mu,\sigma^2;x_1,\ldots,x_n)}{\partial \mu} = 0 \\ \frac{\partial L(\mu,\sigma^2;x_1,\ldots,x_n)}{\partial \sigma^2} = 0 \end{cases}$$

Estymator Największej Wiarogodności - przykład

Model gaussowski: $ENW[\mu]$ **i** $ENW[\sigma^2]$
Zamiast $L(\theta; x_1, \ldots, x_n)$ łatwiej

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2; x_1, \ldots, x_n) &= \log L(\mu, \sigma^2; x_1, \ldots, x_n) \\ &= -n \log\sqrt{2\pi} - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \end{aligned}$$

Estymator Największej Wiarogodności - przykład

Model gaussowski: $ENW[\mu]$ **i** $ENW[\sigma^2]$
Mamy

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \mu} = -\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \\ \frac{\partial \mathcal{L}}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0 \end{cases}$$

Estymator Największej Wiarogodności - przykład

Model gaussowski: $ENW[\mu]$ **i** $ENW[\sigma^2]$
Otrzymujemy

$$ENW[\mu] = \bar{X}$$

$$ENW[\sigma^2] = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$$

Estymator Największej Wiarogodności - przykład

Model gaussowski: estymacja wariancji
Niech $c > 0$ oraz

$$S^2(c) = c \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$$

Ryzyko estymatora $S^2(c)$:

$$R_{\mu,\sigma^2}\left(c \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2\right)$$

Dla jakiego c ryzyko jest jednostajnie najmniejsze?

Estymator Największej Wiarogodności - przykład

Model gaussowski: estymacja wariancji

$$\begin{aligned} & R_{\mu,\sigma^2}\left(c \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2\right) \\ &= E_{\mu,\sigma^2}\left[c \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2 - \sigma^2\right]^2 \\ &= \sigma^4 \left[c^2(n^2 - 1) - 2c(n-1) + 1\right] \end{aligned}$$

Estymator Największej Wiarogodności - przykład

Model gaussowski: estymacja wariancji

$$R_{\mu,\sigma^2}(ENMW[\sigma^2]) = \frac{2}{n-1}\sigma^4$$

$$R_{\mu,\sigma^2}(ENW[\sigma^2]) = \frac{2n-1}{n^2}\sigma^4$$

Estymator o jednostajnie najmniejszym ryzyku

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

3.6 Asymptotyka

Estymator Największej Wiarogodności

Twierdzenie
Niech $p_\theta(x)$ spełniają pewne warunki regularności.

Jeżeli $X_1, \ldots, X_n$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie o gęstości $p_\theta(x)$, to równanie wiarogodności

$$\frac{\partial}{\partial\theta}[p_\theta(x_1)\cdots p_\theta(x_n)] = 0$$

ma rozwiązanie $\hat{\theta}_n(x_1, \ldots, x_n)$ takie, że

Estymator Największej Wiarogodności

Twierdzenie

$$\forall\varepsilon > 0 \quad \lim_{n\to\infty} P_\theta\left\{\left|\hat{\theta}_n(X_1, \ldots, X_n) - \theta\right| < \varepsilon\right\} = 1$$

$$(*) \qquad \sqrt{n}\left(\hat{\theta}_n(X_1, \ldots, X_n) - \theta\right) \to N\left(0, \frac{1}{I_\theta}\right).$$

Definicja
Estymator $\hat{\theta}_n$ spełniający warunek $(*)$ nazywamy *asymptotycznie efektywnym*

3.7 Inne przykłady

Inne przykłady

Rodziny wykładnicze

Estymujemy parametr θ jednoparametrowej rodziny wykładniczej $p_\theta(x) = \exp\{\theta T(x) - b(\theta)\}$.

Wiarogodność

$$L(\theta; x_1, \ldots, x_n) = \exp\left\{\theta \sum_{i=1}^{n} T(x_i) - nb(\theta)\right\}$$

Inne przykłady

Rodziny wykładnicze

Estymator Największej Wiarogodności $\hat{\theta}_n = ENW[\theta]$ istnieje i jest rozwiązaniem równania

$$\frac{1}{n}\sum_{i=1}^{n} T(x_i) = \frac{db(\theta)}{d\theta}$$

Inne przykłady

Rodziny wykładnicze

Ponadto

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} P_\theta\left\{\left|\hat{\theta}_n - \theta\right| < \varepsilon\right\} = 1$$

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \to N\left(0, \frac{1}{I_\theta = D_\theta^2 T}\right)$$

Inne przykłady

ENW wyznaczony tylko numerycznie

Model statystyczny

$$(\mathbb{R}_+, \{P_\theta : \theta \in \mathbb{R}_+\})$$

$$p_\theta(x) = \theta x^{\theta-1} \exp\left\{-x^\theta\right\}$$

Inne przykłady

ENW wyznaczony tylko numerycznie
Rozkład próby

$$p_\theta(x_1, \ldots, x_n) = \theta \left(\prod_{i=1}^{n} x_i \right)^{\theta - 1} \exp \left\{ - \sum_{i=1}^{n} x_i^\theta \right\}$$

Logarytm wiarogodności

$$\mathcal{L}(\theta; x_1, \ldots, x_n) = n \log \theta + (\theta - 1) \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \exp \{\theta \log x_i\}$$

Inne przykłady

ENW wyznaczony tylko numerycznie
Pochodna

$$\frac{d\mathcal{L}(\theta; x_1, \ldots, x_n)}{d\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log x_i \exp \{\theta \log x_i\}$$

$$\frac{d\mathcal{L}(\theta; x_1, \ldots, x_n)}{d\theta} \text{ jest ciągła}$$

$$\lim_{\theta \to 0} \frac{d\mathcal{L}(\theta; x_1, \ldots, x_n)}{d\theta} = +\infty, \quad \lim_{\theta \to +\infty} \frac{d\mathcal{L}(\theta; x_1, \ldots, x_n)}{d\theta} = -\infty$$

Inne przykłady

ENW wyznaczony tylko numerycznie
Równanie wiarogodności

$$\frac{d\mathcal{L}(\theta; x_1, \ldots, x_n)}{d\theta} = 0$$

ma rozwiązanie.

Istnieje $ENW[\theta]$

Inne przykłady

ENW wyznaczony niejednoznacznie
Model statystyczny

$$\left(\mathbb{R}, \left\{U\left(\theta - \frac{1}{2}, \theta + \frac{1}{2}\right) : \theta \in \mathbb{R}\right\}\right)$$

Rozkład próby

$$p_\theta(x_1, \ldots, x_n) = \begin{cases} 1, & \text{gdy } \theta - \frac{1}{2} \leq x_1, \ldots, x_n \leq \theta + \frac{1}{2}, \\ 0, & \text{poza tym.} \end{cases}$$

Inne przykłady

ENW wyznaczony niejednoznacznie
Wiarogodność

$$L(\theta; x_1, \ldots, x_n) = \begin{cases} 1, & \text{gdy } x_{n:n} - \frac{1}{2} \leq \theta \leq x_{1:n} + \frac{1}{2}, \\ 0, & \text{poza tym.} \end{cases}$$

$$x_{1:n} = \min\{x_1, \ldots, x_n\}, \; x_{n:n} = \max\{x_1, \ldots, x_n\}$$

Każda liczba z przedziału $\left[X_{n:n} - \frac{1}{2}, X_{1:n} + \frac{1}{2}\right]$ jest $ENW[\theta]$.

Inne przykłady

ENW nie istnieje (model błędów grubych)
Model statystyczny

$$\left(\mathbb{R}, \left\{P_{\mu,\sigma^2} : (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\right\}\right)$$

$$\begin{aligned} p_{\mu,\sigma^2}(x) = (1 - \varepsilon) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x - \mu)^2\right\} \\ + \varepsilon \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right\} \end{aligned}$$

Inne przykłady

ENW nie istnieje (model błędów grubych)
Wiarogodność

$$L(\mu, \sigma^2; x_1, \ldots, x_n) = \prod_{i=1}^{n} p_{\mu,\sigma^2}(x_i)$$

Dla każdego $M > 0$ istnieją (μ, σ^2) takie, że

$$L(\mu, \sigma^2; x_1, \ldots, x_n) > M$$

4 EMNK

4.1 Przykład wstępny

$EMNK$**: przykład wstępny**

Funkcja Cobba–Douglasa
W ekonomicznej teorii produkcji rozważa się funkcję produkcji Cobba–Douglasa:

$$z = AL^{\alpha}K^{\beta}$$

gdzie z oznacza wielkość produkcji, L jest nakładem pracy, K nakładem kapitału. Liczby A, α, β są pewnymi stałymi. Ekonomiści interesują się wielkościami A, α i β.

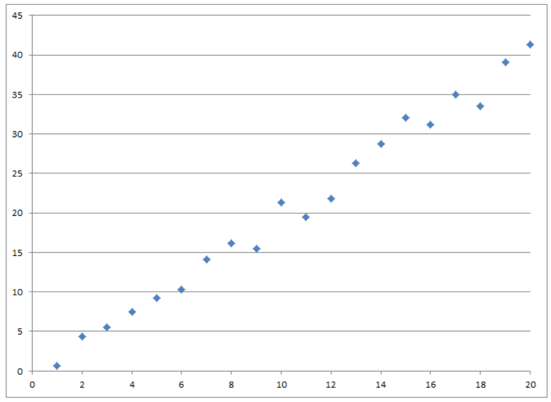
$EMNK$**: przykład wstępny**

Eksperyment
W celu oceny nieznanych wielkości A, α, β prowadzone są obserwacje wielkości produkcji Z_i przy różnych nakładach L_i oraz K_i, przy czym zakłada się, że obserwacje te obarczone są efektami losowymi ε_i

$$Z_i = AL_i^{\alpha}K_i^{\beta} + \varepsilon_i, \quad i = 1, \ldots, n$$

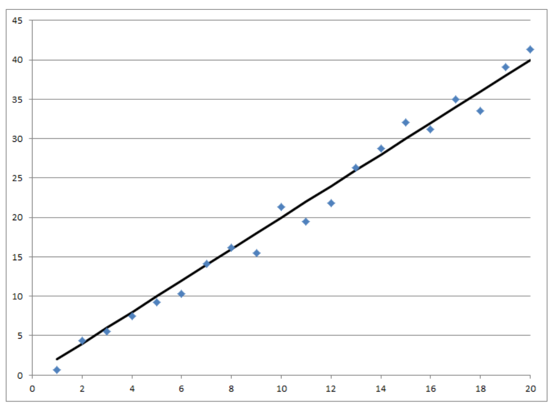
$EMNK$**: przykład wstępny**

Eksperyment



$EMNK$: **przykład wstępny**

Eksperyment

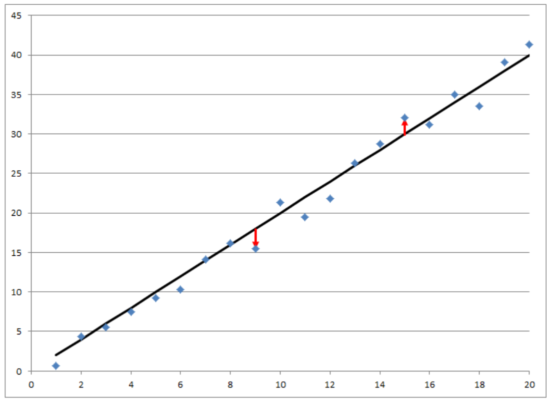


$EMNK$: **przykład wstępny**

Eksperyment

$EMNK$: **przykład wstępny**

Zasada najmniejszych kwadratów



Jako oszacowania nieznanych parametrów A, α, β przyjmuje się takie wartości, przy których błędy losowe są małe

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Z_i - A L_i^{\alpha} K_i^{\beta})^2 = \min!$$

4.2 Określenie

Estymator najmniejszych kwadratów

Model

Obserwujemy zmienne losowe $Y_1, \ldots, Y_n$ takie, że

$$EY_i = g_i(\theta), \quad i = 1, \ldots, n$$

$\theta \in \Theta \subset \mathbb{R}^k$

$g_i : \Theta \to \mathbb{R}^1$ są znanymi funkcjami

Estymator najmniejszych kwadratów

Kryterium

$$S(\theta) = \sum_{i=1}^{n} (Y_i - g_i(\theta))^2$$

Estymator najmniejszych kwadratów

Wielkość $\hat{\theta} \xlongequal{\text{ozn}} EMNK[\theta]$ minimalizująca $S(\theta)$

Resztowa suma kwadratów

$$S(\hat{\theta}) = \sum_{i=1}^{n} (Y_i - g_i(\hat{\theta}))^2$$

4.3 Model liniowy

Model liniowy

Definicja
Modelem liniowym nazywamy model statystyczny, w którym obserwacje $Y_1, \ldots, Y_n$ mają postać

$$Y_i = \beta_1 x_{1i} + \cdots + \beta_p x_{pi} + \varepsilon_i, \quad i = 1, \ldots, n$$

gdzie x_{ji} są ustalonymi liczbami, β_j są nieznanymi parametrami modelu, ε_i są niezależnymi „błędami losowymi" takimi, że $E\varepsilon_i = 0$ oraz $D^2\varepsilon_i = \sigma^2$

Model liniowy

Zapis macierzowy

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}.$$

$$\boldsymbol{Y}' = (Y_1, \ldots, Y_n) \quad \boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots \beta_p) \quad \boldsymbol{\varepsilon}' = (\varepsilon_1, \ldots, \varepsilon_n)$$

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & \cdots & x_{p1} \\ x_{12} & \cdots & x_{p2} \\ \vdots & \cdots & \vdots \\ x_{1n} & \cdots & x_{pn} \end{bmatrix}$$

Założenie: macierz $\boldsymbol{X}$ jest pełnego rzędu

Model liniowy

Resztowa suma kwadratów

$$\begin{aligned} S(\boldsymbol{\beta}) &= \sum_{i=1}^{n} \left(Y_i - \sum_{j=1}^{p} \beta_j x_{ji} \right)^2 \\ &= (\boldsymbol{Y} - \boldsymbol{X\beta})'(\boldsymbol{Y} - \boldsymbol{X\beta}) = \boldsymbol{Y}'\boldsymbol{Y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{X\beta} \end{aligned}$$

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'} = -2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X\beta}$$

Model liniowy

Układ równań normalnych

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}\boldsymbol{Y}$$

Estymator najmniejszych kwadratów

$$EMNK[\boldsymbol{\beta}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \xlongequal{\text{ozn}} \hat{\boldsymbol{\beta}}$$

Uwaga geometryczna

$\boldsymbol{X}\hat{\boldsymbol{\beta}}$ jest rzutem $\boldsymbol{Y}$ na $\{\boldsymbol{X}\boldsymbol{\beta} : \boldsymbol{\beta} \in \mathbb{R}^p\}$

$$(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{X} = \mathbf{0}$$

Model liniowy

$EMNK$ **funkcji liniowej**

Jeżeli $\boldsymbol{c} \in \mathbb{R}^p$, to

$$EMNK[\boldsymbol{c}'\boldsymbol{\beta}] = \boldsymbol{c}'\hat{\boldsymbol{\beta}}$$

Przykład

$$\beta_1 = [1, 0, \cdots, 0]\boldsymbol{\beta}$$
$$EMNK[\beta_1] = [1, 0, \cdots, 0]\hat{\boldsymbol{\beta}}$$

4.4 Przykłady

Estymacja μ

Zadanie

Na podstawie obserwacji $Y_1, \ldots, Y_n$ oszacować ich wartość oczekiwaną μ.

Estymacja μ

Model

$$Y_i = \mu + \varepsilon_i, \quad i = 1, \ldots, n$$

Zapis macierzowy

$$\boldsymbol{Y}' = [Y_1, \ldots, Y_n]',\ \beta = \mu,\ \boldsymbol{X} = \mathbf{1}_n$$

Estymator

$$EMNK[\mu] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = \bar{Y}$$

Estymacja $\mu_1 - \mu_2$

Zadanie

Na podstawie obserwacji $Y_1, \ldots, Y_{n_1}$ o wartości oczekiwanej μ_1 oraz $Y_{n_1+1}, \ldots, Y_n$ o wartości oczekiwanej μ_2 oszacować $\mu_1 - \mu_2$

Estymacja $\mu_1 - \mu_2$

Model

$$\begin{aligned} Y_i &= x_{i1}\mu_1 + x_{i2}\mu_2 + \varepsilon_i, \quad i = 1, \ldots, n \\ x_{i1} &= 1, x_{i2} = 0 \text{ dla } i = 1, \ldots, n_1 \\ x_{i1} &= 0, x_{i2} = 1 \text{ dla } i = n_1 + 1, \ldots, n \end{aligned}$$

Zapis macierzowy

$$\boldsymbol{\beta}' = [\mu_1, \mu_2], \quad \boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} \\ \mathbf{0}_{n-n_1} & \mathbf{1}_{n-n_1} \end{bmatrix}$$

Estymacja $\mu_1 - \mu_2$

Estymator wektora $\boldsymbol{\beta}' = [\mu_1, \mu_2]$

$$\begin{aligned} EMNK[\boldsymbol{\beta}] &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} \\ &= \begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{n_1} Y_i \\ \sum_{i=n_1+1}^{n} Y_i \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i \\ \frac{1}{n-n_1} \sum_{i=n_1+1}^{n} Y_i \end{bmatrix} \end{aligned}$$

Estymacja $\mu_1 - \mu_2$

Estymator różnicy $\mu_1 - \mu_2$

Jeżeli $\boldsymbol{c}' = [1, -1]$, to $\boldsymbol{c}'\boldsymbol{\beta} = \mu_1 - \mu_2$

$$EMNK[\mu_1 - \mu_2] = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i - \frac{1}{n - n_1} \sum_{i=n_1+1}^{n} Y_i$$

Estymator średniej μ_1

Jeżeli $\boldsymbol{c}' = [1, 0]$, to $\boldsymbol{c}'\boldsymbol{\beta} = \mu_1$

$$EMNK[\mu_1] = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i$$

4.5 Własności

Własności $EMNK$

Nieobciążoność
$EMNK[\boldsymbol{\beta}]$ jest estymatorem nieobciążonym o macierzy kowariancji $\sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$

$$E_{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}} = E_{\boldsymbol{\beta}}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'E_{\boldsymbol{\beta}}\boldsymbol{Y} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{\beta}$$

Własności $EMNK$

Funkcja estymowalna
Funkcję parametryczną $\boldsymbol{c}'\boldsymbol{\beta}$ nazywamy estymowalną, jeżeli istnieje jej estymator nieobciążony postaci $\boldsymbol{b}'\boldsymbol{Y}$

Twierdzenie
Funkcja parametryczna $\boldsymbol{c}'\boldsymbol{\beta}$ jest estymowalna wtedy i tylko wtedy, gdy $\boldsymbol{c} \in \mathrm{Im}\boldsymbol{X}'$

Własności $EMNK$

Twierdzenie *Gaussa–Markowa*
Jeżeli błędy losowe $\varepsilon_1, \ldots, \varepsilon_n$ są nieskorelowanymi zmiennymi losowymi o zerowej wartości oczekiwanej i takiej samej wariancji, to dla każdej estymowalnej funkcji parametrycznej $\boldsymbol{c}'\boldsymbol{\beta}$ i dla każdego nieobciążonego estymatora liniowego $\boldsymbol{b}'\boldsymbol{Y}$ tej funkcji zachodzi

$$D^2_{\boldsymbol{\beta}}(\boldsymbol{c}'\hat{\boldsymbol{\beta}}) \leq D^2_{\boldsymbol{\beta}}(\boldsymbol{b}'\boldsymbol{Y}), \quad \forall \boldsymbol{\beta} \in \mathbb{R}^p$$

4.6 Estymacja wariancji

Estymacja wariancji σ^2

Resztowa suma kwadratów

$$\begin{aligned}
S(\hat{\boldsymbol{\beta}}) &= \sum_{i=1}^{n}\left(Y_i - \sum_{j=1}^{p}\hat{\beta}_j x_{ji}\right)^2 = \|\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}\|^2 \\
&= (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\
&= (\boldsymbol{Y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y})'(\boldsymbol{Y} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}) \\
&= \boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}'(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X})(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y} \\
&= \boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y}
\end{aligned}$$

Estymacja wariancji σ^2

Resztowa suma kwadratów

$$\begin{aligned} E_{\boldsymbol{\beta},\sigma^2}\|\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}\|^2 &= E_{\boldsymbol{\beta},\sigma^2}\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y} \\ &= \mathrm{tr}E_{\boldsymbol{\beta},\sigma^2}(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y}\boldsymbol{Y}' \\ &= \mathrm{tr}\left\{(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')E_{\boldsymbol{\beta},\sigma^2}(\boldsymbol{Y}\boldsymbol{Y}')\right\} \\ &= \sigma^2\mathrm{tr}(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}') \\ &= (n-\mathrm{rz}\boldsymbol{X})\sigma^2 \end{aligned}$$

Estymacja wariancji σ^2

$EMNK[\sigma^2]$

$$\hat{\sigma}^2 = \frac{1}{n-\mathrm{rz}\boldsymbol{X}}\boldsymbol{Y}'(\boldsymbol{I}-\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y}$$

Estymacja wariancji σ^2

Zadanie

Na podstawie obserwacji $Y_1,\ldots,Y_n$ o wartości oczekiwanej μ oszacować ich wariancję σ^2.

Estymacja wariancji σ^2

Model

$$Y_i = \mu + \varepsilon_i, \quad i = 1,\ldots,n$$

Zapis macierzowy

$$\boldsymbol{Y}' = [Y_1,\ldots,Y_n]', \ \beta = \mu, \ \boldsymbol{X} = \mathbf{1}_n$$

$$\mathrm{rz}\boldsymbol{X} = 1$$

Estymacja wariancji σ^2

Estymator

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{1}{n-1} \boldsymbol{Y}'(\boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n')\boldsymbol{Y} \\
&= \frac{1}{n-1}\left\{\boldsymbol{Y}'\boldsymbol{Y} - \frac{1}{n}\left(\boldsymbol{Y}'\mathbf{1}_n\right)\left(\mathbf{1}_n'\boldsymbol{Y}\right)\right\} \\
&= \frac{1}{n-1}\left\{\sum_{i=1}^{n} Y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Y_i\right)^2\right\} \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2
\end{aligned}
$$

Estymacja wariancji σ^2

Zadanie

Na podstawie obserwacji $Y_1, \ldots, Y_{n_1}$ o wartości oczekiwanej μ_1 i wariancji σ^2 oraz $Y_{n_1+1}, \ldots, Y_n$ o wartości oczekiwanej μ_2 i wariancji σ^2 oszacować σ^2

Estymacja wariancji σ^2

Model

$$Y_i = x_{i1}\mu_1 + x_{i2}\mu_2 + \varepsilon_i, \quad i = 1, \ldots, n$$

$$x_{i1} = 1, x_{i2} = 0 \text{ dla } i = 1, \ldots, n_1$$

$$x_{i1} = 0, x_{i2} = 1 \text{ dla } i = n_1 + 1, \ldots, n$$

Zapis macierzowy

$$\boldsymbol{\beta}' = [\mu_1, \mu_2], \quad \boldsymbol{X} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} \\ \mathbf{0}_{n-n_1} & \mathbf{1}_{n-n_1} \end{bmatrix}$$

$$\text{rz}\boldsymbol{X} = 2$$

Estymacja wariancji σ^2

Estymator

$$
\begin{aligned}
(n-2)\hat{\sigma}^2 &= \boldsymbol{Y}'\left(\boldsymbol{I} - \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} \end{bmatrix} \begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} \\ \mathbf{0}'_{n_1} & \mathbf{1}'_{n_2} \end{bmatrix}\right)\boldsymbol{Y} \\
&= \sum_{i=1}^{n_1}\left(Y_i - \frac{1}{n_1}\left(\sum_{i=1}^{n_1} Y_i\right)\right)^2 + \\
&\quad + \sum_{i=n_1+1}^{n}\left(Y_i - \frac{1}{n_2}\left(\sum_{i=n_1+1}^{n} Y_i\right)\right)^2
\end{aligned}
$$

4.7 Własności probabilistyczne

Rozkłady prawdopodobieństwa estymatorów

Model

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots \beta_p)$$

$$\boldsymbol{\varepsilon}' = (\varepsilon_1, \ldots, \varepsilon_n) \sim N_n(\mathbf{0}, \sigma^2 \boldsymbol{I}_n)$$

$EMNK$

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

$$s^2 = \hat{\sigma}^2 = \frac{1}{n - \mathrm{rz}\boldsymbol{X}}\boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}')\boldsymbol{Y}$$

Rozkłady prawdopodobieństwa estymatorów

Twierdzenie
Jeżeli $\boldsymbol{Y} \sim N_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n)$ oraz rz$\boldsymbol{X} = p$, to

- $\hat{\boldsymbol{\beta}} \sim N_p(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1})$
- $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim \sigma^2\chi^2_p$
- $\hat{\boldsymbol{\beta}}$ jest niezależne od s^2
- $(n-p)s^2 \sim \sigma^2\chi^2_{n-p}$