Estymacja parametrów

Krzysztof Topolski

Wykład 3

Wrocław, 29 października 2018

Definicja.

Estymatorem punktowym nazywamy dowolną funkcję, która zależy jedynie od próby losowej $X_1, X_2, ..., X_n$.

Uwaga 1.

Przy tak przyjętej definicji każda statystyka jest estymatorem.

Definicja.

Estymatorem punktowym nazywamy dowolną funkcję, która zależy jedynie od próby losowej $X_1, X_2, ..., X_n$.

Uwaga 1.

Przy tak przyjętej definicji każda statystyka jest estymatorem.

Uwaga 2.

Należy odróżnić estymator od wartości estymatora.

Jeśli $X_1, X_2, ..., X_n$ jest próba losowa, a $x_1, x_2, ..., x_n$ jest realizacja próby losowej to $W(X_1, X_2, ..., X_n)$ jest estymatorem, a $W(x_1, x_2, ..., x_n)$ jest wartością estymatora.

Metody wyznaczania estymatorów.

Przegląd metod wyznaczania estymatorów zaczniemy od **metody momentów**. Przy konstrukcji estymatorów możemy skorzystać z dobrze znanych estymatorów. Jednym z nich jest średnia próbkowa

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

która jest dobrym estymatorem wartości oczekiwanej $\mathbf{E}X_1$ zmiennej losowej X_1.

Podobnie

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2,$$

jest dobrym estymatorem momentu rzędu dwa, $\mathbf{E}X_1^2$, zmiennej losowej X_1.

Metody wyznaczania estymatorów.

Przegląd metod wyznaczania estymatorów zaczniemy od **metody momentów**. Przy konstrukcji estymatorów możemy skorzystać z dobrze znanych estymatorów. Jednym z nich jest średnia próbkowa

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

która jest dobrym estymatorem wartości oczekiwanej $\mathbf{E}X_1$ zmiennej losowej X_1.

Podobnie

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2,$$

jest dobrym estymatorem momentu rzędu dwa, $\mathbf{E}X_1^2$, zmiennej losowej X_1.

Ogólnie dla dowolnego $k = 1, 2, \ldots$

$$\frac{1}{n} \sum_{i=1}^{n} X_i^k,$$

jest dobrym estymatorem momentu rzędu k, $\mathbf{E} X_1^k$, zmiennej losowej X_1.

Ta obserwacja jest punktem wyjścia następującej konstrukcji.

Ogólnie dla dowolnego $k = 1, 2, ...$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k,$$

jest dobrym estymatorem momentu rzędu k, $\mathbf{E}X_1^k$, zmiennej losowej X_1.

Ta obserwacja jest punktem wyjścia następującej konstrukcji.

Metoda momentów wyznaczania estymatorów.

Niech $X_1, X_2, ..., X_n$ będzie próbą losową z populacji o rozkładzie z gęstością $f(x|\theta_1, \theta_2, ..., \theta_k)$.

Moment	**Estymator momentu**
$\mu_1 = \mathbf{E}X_1$	$m_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$
$\mu_2 = \mathbf{E}X_1^2$	$m_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$
$\vdots$	$\vdots$
$\mu_k = \mathbf{E}X_1^k$	$m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

Metoda momentów

Zwykle momenty μ_j są funkcjami parametrów $\theta_1, ..., \theta_k$ i wtedy

$$\mu_j = g_j(\theta_1, ..., \theta_k).$$

Estymator $(\hat{\theta}_1, ..., \hat{\theta}_k)$ wektora parametrów $(\theta_1, ..., \theta_k)$ wyznaczony metodą momentów powstaje jako rozwiązanie układu równań

$$\begin{array}{rcl} m_1 & = & g_1(\theta_1, ..., \theta_k) \equiv \mu_1 \\ m_2 & = & g_2(\theta_1, ..., \theta_k) \equiv \mu_2 \\ \vdots & \vdots & \vdots \\ m_k & = & g_k(\theta_1, ..., \theta_k) \equiv \mu_k \end{array}$$

względem $\theta_1, ..., \theta_k$.

Przykład 1. (Rozkład normalny)

Rozpatrzmy próbę losową $(X_1, X_2, ..., X_n)$ z rozkładu normalnego $N(\mu, \sigma^2)$ o nieznanej wartości oczekiwanej μ i nieznanej wariancji σ^2. W przyjętej notacji szukamy estymatora wektora parametrów modelu $(\theta_1, \theta_2) = (\mu, \sigma^2)$.

Układ równań

$$\begin{aligned} \mu_1 &= g_1(\theta_1, \theta_2) = g_1(\mu, \sigma^2) = \mu \\ \mu_2 &= g_2(\theta_1, \theta_2) = g_2(\mu, \sigma^2) = \mu^2 + \sigma^2 \end{aligned}$$

Przykład 1. (Rozkład normalny)

Rozpatrzmy próbę losową $(X_1, X_2, ..., X_n)$ z rozkładu normalnego $N(\mu, \sigma^2)$ o nieznanej wartości oczekiwanej μ i nieznanej wariancji σ^2. W przyjętej notacji szukamy estymatora wektora parametrów modelu $(\theta_1, \theta_2) = (\mu, \sigma^2)$.

Układ równań

$$\begin{aligned} \mu_1 &= g_1(\theta_1, \theta_2) = g_1(\mu, \sigma^2) = \mu \\ \mu_2 &= g_2(\theta_1, \theta_2) = g_2(\mu, \sigma^2) = \mu^2 + \sigma^2 \end{aligned}$$

Przykład 1, cd.

Stąd otrzymujemy równości

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

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

Przykład 1, cd.

Rozwiązując ten układ ze względu na μ i σ^2, otrzymujemy $\hat{\mu}$ estymator wartości oczekiwanej μ oraz $\hat{\sigma^2}$ estymator wariancji σ^2, wyznaczone metodą momentów.

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

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]^2$$

lub w zwartej postaci

$$\hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X})^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

Przykład 2.

Rozpatrzmy próbę losową $(X_1, X_2, ..., X_n)$ z rozkładu bernoulliego $b(k, p)$ o gęstości postaci

$$P(X_1 = i \,|\, k, p) = \binom{k}{i} p^i (1-p)^{k-i}, \quad i = 0, 1, ..., k.$$

Zakładamy, że zarówno k jak i p są nieznane.

W przyjętej w opisie metody momentów notacji szukamy estymatora wektora parametrów modelu $(\theta_1, \theta_2) = (k, p)$.

W tym przypadku odpowiedni układ równań ma postć:

$$\mu_1 = g_1(\theta_1, \theta_2) = g_1(k, p) = kp$$

$$\mu_2 = g_2(\theta_1, \theta_2) = g_2(k, p) = kp(1-p) + k^2 p^2$$

Przykład 2.

Rozpatrzmy próbę losową $(X_1, X_2, ..., X_n)$ z rozkładu bernoulliego $b(k, p)$ o gęstości postaci

$$P(X_1 = i \,|\, k, p) = \binom{k}{i} p^i (1-p)^{k-i}, \quad i = 0, 1, ..., k.$$

Zakładamy, że zarówno k jak i p są nieznane.

W przyjętej w opisie metody momentów notacji szukamy estymatora wektora parametrów modelu $(\theta_1, \theta_2) = (k, p)$.

W tym przypadku odpowiedni układ równań ma postć:

$$\begin{aligned} \mu_1 &= g_1(\theta_1, \theta_2) = g_1(k, p) = kp \\ \mu_2 &= g_2(\theta_1, \theta_2) = g_2(k, p) = kp(1-p) + k^2 p^2 \end{aligned}$$

Przykład 2 cd.

Otrzymujemy stąd równości

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} X_i &= kp \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 &= kp(1-p) + k^2 p^2\end{aligned}$$

Rozwiązując ten układ ze względu na k i p otrzymujemy $\hat{k}$, estymator k, oraz $\hat{p}$, estymator p, wyznaczone metodą momentów.

$$\begin{aligned}\hat{k} &= \frac{\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{i=1}^{n} X_i\right)^2} \\ \hat{p} &= \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\hat{k}}\end{aligned}$$

Przykład 2 cd.

Otrzymujemy stąd równości

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} X_i &= kp \\ \frac{1}{n}\sum_{i=1}^{n} X_i^2 &= kp(1-p) + k^2 p^2\end{aligned}$$

Rozwiązując ten układ ze względu na k i p otrzymujemy $\hat{k}$, estymator k, oraz $\hat{p}$, estymator p, wyznaczone metodą momentów.

$$\begin{aligned}\hat{k} &= \frac{\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2}{\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{i=1}^{n} X_i\right)^2} \\ \hat{p} &= \frac{\frac{1}{n}\sum_{i=1}^{n} X_i}{\hat{k}}\end{aligned}$$

Przykład 2 cd.

Korzystajac z oznaczenia $\frac{1}{n}\sum_{i=1}^{n} X_i \equiv \bar{X}$ można zapisać otrzymane estymatory bardziej zwartej postaci

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

$$\hat{p} \;=\; \frac{\bar{X}}{\hat{k}}$$

Zdecydowanie nie są to najlepsze estymatory gdyż na ich podstawie możemy otrzymać ujemne wartości jako oszacowanie k i p co jest niemożliwe gdyż z definicji oba te parametry są liczbami dodatnimi.

Przykład 2 cd.

Korzystajac z oznaczenia $\frac{1}{n}\sum_{i=1}^{n} X_i \equiv \bar{X}$ można zapisać otrzymane estymatory bardziej zwartej postaci

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

$$\hat{p} = \frac{\bar{X}}{\hat{k}}$$

Zdecydowanie nie są to najlepsze estymatory gdyż na ich podstawie możemy otrzymać ujemne wartości jako oszacowanie k i p co jest niemożliwe gdyż z definicji oba te parametry są liczbami dodatnimi.

Metoda podstawiania częstości

Za oszacowanie nieznanych prawdopodobieństw pojawiania się zdarzeń przyjmujemy czestości ich wystąpienia w próbie losowej.

Przykład 3.

Załóżmy, że n obiektów wybranych w sposób niezależny klasyfikujemy (ze względu na wybraną cechę) do k rozłącznych klas. Niech

- N_i, oznacza liczbę obiektów w i-tej klasie,
- p_i, oznacza prawdopodobieństwo należenia do $i-$tej klasy.

Metoda podstawiania częstości

Za oszacowanie nieznanych prawdopodobieństw pojawiania się zdarzeń przyjmujemy czestości ich wystąpienia w próbie losowej.

Przykład 3.

Załóżmy, że n obiektów wybranych w sposób niezależny klasyfikujemy (ze względu na wybraną cechę) do k rozłącznych klas. Niech

- N_i, oznacza liczbę obiektów w i-tej klasie,
- p_i, oznacza prawdopodobieństwo należenia do $i-$tej klasy.

Metoda podstawiania częstości

Za oszacowanie nieznanych prawdopodobieństw pojawiania się zdarzeń przyjmujemy czestości ich wystąpienia w próbie losowej.

Przykład 3.

Załóżmy, że n obiektów wybranych w sposób niezależny klasyfikujemy (ze względu na wybraną cechę) do k rozłącznych klas. Niech

- N_i, oznacza liczbę obiektów w i-tej klasie,
- p_i, oznacza prawdopodobieństwo należenia do $i-$tej klasy.

Przykład 3, cd.

Wektor obserwacji $(N_1, N_2, ..., N_k)$ ma rozkład wielomianowy $M(x \mid n, p_1, p_2, ..., p_k)$ o gęstości

$$P\left(N_1 = n_1, N_2 = n_2, ..., N_k = n_k\right) = \frac{n\,!}{\prod_{i=1}^{k}\ n_i\,!} \prod_{i=1}^{k} p_i^{n_i},$$

gdzie $\sum_{i=1}^{k}\ n_i = n$, oraz $\sum_{i=1}^{k} p_i = 1$.

Przykład 3 cd.

W tej sytuacji naturalnym oszacowaniem wektora nieznanych prawdopodobieństw

$$(p_1,\ p_2,\ \ldots,\ p_k)$$

jest zastąpienie ich przez obserwowane częstości

$$(\hat{p}_1,\ \hat{p}_2,\ \ldots,\ \hat{p}_k) = \left(\frac{N_1}{n},\ \frac{N_2}{n},\ \ldots,\ \frac{N_k}{n}\right).$$

Przykład 3 cd.

W tej sytuacji naturalnym oszacowaniem wektora nieznanych prawdopodobieństw

$$(p_1,\ p_2,\ \ldots,\ p_k)$$

jest zastąpienie ich przez obserwowane częstości

$$(\hat{p}_1,\ \hat{p}_2,\ \ldots,\ \hat{p}_k) = \left(\frac{N_1}{n},\ \frac{N_2}{n},\ \ldots,\ \frac{N_k}{n}\right).$$