1. Przykład – wyznaczanie głębokości studni

wiedza wstępna (a priori)

według praw Newtona, związek pomiędzy prawdziwą głębokością studni θ^*, a czasem x opadania kamienia na dno studni ma postać

$$\theta^* = \frac{gx^2}{2} \text{ – wiedza parametryczna} \tag{1}$$

pomija się opór powietrza i skończoną prędkość dźwięku

wiedza z pomiarów (eksperymentów)

wrzucamy kolejno N kamieni i mierzymy czas spadania

uzyskujemy ciąg losowy pomiarów czasu spadania x (różne wyniki)

$$x_1, x_2, x_3, ..., x_N \text{ – zależą od } \theta^*$$

cel: korzystając z obydwu typów wiedzy, wyznaczyć możliwie "dokładnie" (wyestymować) θ^*

$$\theta_N(x_1, ..., x_N; \text{wiedza wstępna}) \text{ – estymator}$$

Pomysł 1 – $\widehat{\theta}_N^{(1)}$

uśrednić pomiary czasów i wynik wstawić do wzoru (1)

$$\widehat{\theta}_N^{(1)} = \frac{g\overline{x}^2}{2}, \text{ gdzie } \overline{x} = \frac{x_1 + ... + x_N}{N}$$

Pomysł 2 – $\widehat{\theta}_N^{(2)}$

po kolejnych pomiarach czasów spadania, na podstawie (1) wyznaczyć oszacowania głębokości, a następnie uśrednić je

$$\widehat{\theta}_N^{(2)} = \frac{\frac{gx_1^2}{2} + ... + \frac{gx_N^2}{2}}{N}$$

2. Własności estymatorów i kryteria ich oceny

Praktyczne – dla ustalonej liczby pomiarów $N = N_0$

– **obciążenie**

$$bias\widehat{\theta}_{N_0} = \mathbf{E}\widehat{\theta}_{N_0} - \theta^*$$

gdy $bias\widehat{\theta}_{N_0} = 0$, tj. $\mathbf{E}\widehat{\theta}_{N_0} = \theta^*$ wtedy estymator nazywamy *nieobciążonym* (może takich istnieć więcej niż jeden – klasa)

– **wariancja**

$$var\widehat{\theta}_{N_0} = \mathbf{E}(\widehat{\theta}_{N_0} - \mathbf{E}\widehat{\theta}_{N_0})^2$$

dla estymatorów nieobciążonych jest ona tożsama z tzw. błędem średniokwadratowym

$$var\widehat{\theta}_{N_0} = \mathbf{E}(\widehat{\theta}_{N_0} - \theta^*)^2 = MSE(\widehat{\theta}_{N_0}) \text{ – chcemy by był jak najmniejszy}$$

Asymptotyczne – graniczne, gdy $N \to \infty$

– **estymator słabo zgodny**

$$\widehat{\theta}_N \xrightarrow{p} \theta^*, \text{ gdy } N \to \infty$$

– **estymator mocno zgodny**

$$\widehat{\theta}_N \xrightarrow{p.1} \theta^*, \text{ gdy } N \to \infty$$

– **estymator średniokwadratowo zgodny**

$$\widehat{\theta}_N \xrightarrow{L_2} \theta^*, \text{ gdy } N \to \infty, \text{ lub równoważnie } \lim_{N\to\infty} MSE(\widehat{\theta}_N) = 0$$

ang. mean squared error

zgodność średniokwadratowa – zbieżność do prawdziwej wartości θ^* wg średniej rzędu $r = 2$

$$\begin{aligned} MSE(\widehat{\theta}_N) &= \mathbf{E}(\widehat{\theta}_N - \theta^*)^2 = \mathbf{E}\left\{(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N) + (\mathbf{E}\widehat{\theta}_N - \theta^*)\right\}^2 = \\ &= \mathbf{E}\left\{(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N)^2 + (\mathbf{E}\widehat{\theta}_N - \theta^*)^2 + 2(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N)(\mathbf{E}\widehat{\theta}_N - \theta^*)\right\} = \\ &= \mathbf{E}(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N)^2 + (\mathbf{E}\widehat{\theta}_N - \theta^*)^2 + 0 = var\widehat{\theta}_N + bias^2\widehat{\theta}_N \end{aligned}$$

Podsumowując

$$MSE(\widehat{\theta}_N) = var\widehat{\theta}_N + bias^2\widehat{\theta}_N$$

Zarówno $var\widehat{\theta}_N$, jak i $bias\widehat{\theta}_N$ (a w konsekwencji także $MSE(\widehat{\theta}_N)$) są deterministycznymi (nie losowymi) zmiennej N.
Wniosek

$$\lim_{N\to\infty} MSE(\widehat{\theta}_N) = 0 \iff \begin{array}{l} \lim_{N\to\infty} var\widehat{\theta}_N = 0 \\ \text{i } \lim_{N\to\infty} bias\widehat{\theta}_N = 0 \end{array}$$

Szybkość zbieżności średniokwadratowej – przykład

$$\begin{aligned} var\widehat{\theta}_N &= O(\alpha_N) \\ bias^2\widehat{\theta}_N &= O(\beta_N) \end{aligned}$$

z jaką szybkością zbiega $MSE(\widehat{\theta}_N)$?

$$\begin{aligned} var\widehat{\theta}_N &\leqslant c_1\alpha_N \\ bias^2\widehat{\theta}_N &\leqslant c_2\beta_N \\ MSE(\widehat{\theta}_N) &\leqslant c_1\alpha_N + c_2\beta_N \leqslant c_3 \max(\alpha_N, \beta_N) \end{aligned}$$

a zatem

$$MSE(\widehat{\theta}_N) = O(\max(\alpha_N, \beta_N))$$

3. Przypadek wielu zmiennych – estymacja wektora

$$\theta^* = \begin{bmatrix} \theta^{*(1)} \\ \theta^{*(2)} \\ \vdots \\ \theta^{*(d)} \end{bmatrix}, \qquad \widehat{\theta}_N = \begin{bmatrix} \widehat{\theta}_N^{(1)} \\ \widehat{\theta}_N^{(2)} \\ \vdots \\ \widehat{\theta}_N^{(d)} \end{bmatrix}, \qquad bias\widehat{\theta}_N = \mathbf{E}\widehat{\theta}_N - \theta^*$$

macierz kowariancji

$$cov(\widehat{\theta}_N) = \mathbf{E}\left\{(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N)(\widehat{\theta}_N - \mathbf{E}\widehat{\theta}_N)^T\right\} = \begin{bmatrix} var\widehat{\theta}_N^{(1)} & A(1) & ... & A(d-1) \\ A(1) & var\widehat{\theta}_N^{(2)} & A(1) & ... \\ ... & A(1) & ... & A(1) \\ A(d-1) & ... & A(1) & var\widehat{\theta}_N^{(d)} \end{bmatrix}$$

4. Ocena globalna estymatora – błędy całkowe

potrzebna, gdy estymujemy nie konkretną wartość (skalar, wektor) lecz funkcję $\theta^*(x)$
dla ustalonego $x = x_0$ mamy doczynienia z estymatorem punktowym $\widehat{\theta}_N(x_0)$ parametru $\theta^*(x_0)$ (j/w)
za ocenę estymatora $\widehat{\theta}_N(x)$ funkcji $\theta^*(x)$, przyjmuje się całkę (po $x \in R$) błędu średniokwadratowego

$$IMSE(\widehat{\theta}_N(x)) = \int_{-\infty}^{\infty} MSE(\widehat{\theta}_N(x))dx = \int_{-\infty}^{\infty} \mathbf{E}(\widehat{\theta}_N(x) - \theta^*(x))^2 dx = \mathbf{E}\int_{-\infty}^{\infty} (\widehat{\theta}_N(x) - \theta^*(x))^2 dx = MISE(\widehat{\theta}_N(x))$$

ang. mean integrated square error