Statystyka i rachunek prawdopodobieństwa

Wykład 9: Estymacja punktowa

Mariusz Tarnopolski

Instytut Astronomii UMK

Estymacja punktowa

- **Teorią estymacji** nazywamy metody szacowania parametrów charakteryzujących populację generalną na podstawie próby.
- W **estymacji punktowej** na podstawie próby szacujemy nieznaną wartość pewnego parametru. Takim nieznanym parametrem mogą być parametry rozkładów zmiennych losowych populacji (np. średnia, wariancja itp.) lub ich bardziej złożone funkcje, np. współczynnik korelacji (zob. dalsze wykłady). Ze względu na **losowy charakter próby** możemy o tych parametrach wypowiadać się jedynie w sposób probabilistyczny, a nie definitywny. Estymacja dostarcza jedynie **oszacowanie parametru populacyjnego**, którą często nazywa się **parametrem empirycznym**.
- Estymatorem parametru Q nazywamy każdą statystykę $\hat{Q}(X_1, X_2, \ldots, X_n)$ zależną od tego parametru, którą przyjmujemy za oszacowanie parametru Q.
- Odchylenie standardowe rozkładu parametru $\hat{Q}$ stanowi błąd estymatora.

Estymacja punktowa

- **Teorią estymacji** nazywamy metody szacowania parametrów charakteryzujących populację generalną na podstawie próby.
- W **estymacji punktowej** na podstawie próby szacujemy nieznaną wartość pewnego parametru. Takim nieznanym parametrem mogą być parametry rozkładów zmiennych losowych populacji (np. średnia, wariancja itp.) lub ich bardziej złożone funkcje, np. współczynnik korelacji (zob. dalsze wykłady). Ze względu na **losowy charakter próby** możemy o tych parametrach wypowiadać się jedynie w sposób probabilistyczny, a nie definitywny. Estymacja dostarcza jedynie **oszacowanie parametru populacyjnego**, którą często nazywa się **parametrem empirycznym**.
- Estymatorem parametru Q nazywamy każdą statystykę $\hat{Q}(X_1, X_2, \ldots, X_n)$ zależną od tego parametru, którą przyjmujemy za oszacowanie parametru Q.
- Odchylenie standardowe rozkładu parametru $\hat{Q}$ stanowi błąd estymatora.

Estymacja punktowa

- **Teorią estymacji** nazywamy metody szacowania parametrów charakteryzujących populację generalną na podstawie próby.
- W **estymacji punktowej** na podstawie próby szacujemy nieznaną wartość pewnego parametru. Takim nieznanym parametrem mogą być parametry rozkładów zmiennych losowych populacji (np. średnia, wariancja itp.) lub ich bardziej złożone funkcje, np. współczynnik korelacji (zob. dalsze wykłady). Ze względu na **losowy charakter próby** możemy o tych parametrach wypowiadać się jedynie w sposób probabilistyczny, a nie definitywny. Estymacja dostarcza jedynie **oszacowanie parametru populacyjnego**, którą często nazywa się **parametrem empirycznym**.
- Estymatorem parametru Q nazywamy każdą statystykę $\hat{Q}(X_1, X_2, \dots, X_n)$ zależną od tego parametru, którą przyjmujemy za oszacowanie parametru Q.
- Odchylenie standardowe rozkładu parametru $\hat{Q}$ stanowi błąd estymatora.

Estymacja punktowa

- **Teorią estymacji** nazywamy metody szacowania parametrów charakteryzujących populację generalną na podstawie próby.
- W **estymacji punktowej** na podstawie próby szacujemy nieznaną wartość pewnego parametru. Takim nieznanym parametrem mogą być parametry rozkładów zmiennych losowych populacji (np. średnia, wariancja itp.) lub ich bardziej złożone funkcje, np. współczynnik korelacji (zob. dalsze wykłady). Ze względu na **losowy charakter próby** możemy o tych parametrach wypowiadać się jedynie w sposób probabilistyczny, a nie definitywny. Estymacja dostarcza jedynie **oszacowanie parametru populacyjnego**, którą często nazywa się **parametrem empirycznym**.
- Estymatorem parametru Q nazywamy każdą statystykę $\hat{Q}(X_1, X_2, \ldots, X_n)$ zależną od tego parametru, którą przyjmujemy za oszacowanie parametru Q.
- Odchylenie standardowe rozkładu parametru $\hat{Q}$ stanowi błąd estymatora.

Np. średnia z próbki $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (która jest zmienną losową) jest estymatorem średniej μ populacji (która nie jest zmienną losową tylko nieznanym parametrem rozkładu; jest to nieznany parametr Q).

Np. średnia z próbki $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (która jest zmienną losową) jest estymatorem średniej μ populacji (która nie jest zmienną losową tylko nieznanym parametrem rozkładu; jest to nieznany parametr Q). Przypuśćmy, że $f(x;\mu)$ jest rozkładem gęstości prawdopodobieństwa zmiennej losowej $\bar{X}$ w populacji. Niech będzie to rozkład normalny $\mathcal{N}(\mu,\sigma)$. Wariancja średniej to zatem

$$V(\bar{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

Np. średnia z próbki $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (która jest zmienną losową) jest estymatorem średniej μ populacji (która nie jest zmienną losową tylko nieznanym parametrem rozkładu; jest to nieznany parametr Q). Przypuśćmy, że $f(x;\mu)$ jest rozkładem gęstości prawdopodobieństwa zmiennej losowej $\bar{X}$ w populacji. Niech będzie to rozkład normalny $\mathcal{N}(\mu,\sigma)$. Wariancja średniej to zatem

$$V(\bar{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

czyli odchylenie standardowe (tj. błąd/niepewność średniej) wynosi

$$\Delta\bar{X} = \sqrt{V(\bar{X})} = \frac{\sigma}{\sqrt{n}}$$

Za oszacowanie parametru Q można przyjąć różne wielkości. **Dobroć** estymatora $\hat{Q}$ określają pewne kryteria:

- Estymator $\hat{Q}_n(X_1, X_2, \ldots, X_n)$ parametru Q nazywany jest **zgodnym** gdy jest zbieżny wg prawdopodobieństwa do wartości Q, czyli

$$\forall \varepsilon > 0 : \quad \lim_{n \to \infty} P\left(\left|\hat{Q}_n - Q\right| < \varepsilon\right) = 1$$

- Estymator $\hat{Q}_n$ parametru Q nazywany jest **nieobciążonym** jeśli jego wartość przeciętna równa się parametrowi estymowanemu dla każdego n:

$$\mathbb{E}\left(\hat{Q}_n\right) = Q$$

- Jeśli natomiast $\mathbb{E}\left(\hat{Q}_n\right) \neq Q$ to wtedy $\hat{Q}_n$ nazywany jest **estymatorem obciążonym**, a różnicę $\mathbb{E}\left(\hat{Q}_n\right) - Q$ **obciążeniem estymatora**.
- W przypadku gdy $\lim\limits_{n \to \infty} \mathbb{E}\left(\hat{Q}_n\right) = Q$ to estymator nazywany jest **asymptotycznie nieobciążonym**.

Za oszacowanie parametru Q można przyjąć różne wielkości. **Dobroć** estymatora $\hat{Q}$ określają pewne kryteria:

- Estymator $\hat{Q}_n(X_1, X_2, \ldots, X_n)$ parametru Q nazywany jest **zgodnym** gdy jest zbieżny wg prawdopodobieństwa do wartości Q, czyli

$$\forall \varepsilon > 0 : \quad \lim_{n\to\infty} P\left(\left|\hat{Q}_n - Q\right| < \varepsilon\right) = 1$$

- Estymator $\hat{Q}_n$ parametru Q nazywany jest **nieobciążonym** jeśli jego wartość przeciętna równa się parametrowi estymowanemu dla każdego n:

$$\mathbb{E}\left(\hat{Q}_n\right) = Q$$

- Jeśli natomiast $\mathbb{E}\left(\hat{Q}_n\right) \neq Q$ to wtedy $\hat{Q}_n$ nazywany jest **estymatorem obciążonym**, a różnicę $\mathbb{E}\left(\hat{Q}_n\right) - Q$ **obciążeniem estymatora**.
- W przypadku gdy $\lim_{n\to\infty} \mathbb{E}\left(\hat{Q}_n\right) = Q$ to estymator nazywany jest **asymptotycznie nieobciążonym**.

Za oszacowanie parametru Q można przyjąć różne wielkości. **Dobroć** estymatora $\hat{Q}$ określają pewne kryteria:

- Estymator $\hat{Q}_n(X_1, X_2, \ldots, X_n)$ parametru Q nazywany jest **zgodnym** gdy jest zbieżny wg prawdopodobieństwa do wartości Q, czyli

$$\forall \varepsilon > 0 : \quad \lim_{n\to\infty} P\left(\left|\hat{Q}_n - Q\right| < \varepsilon\right) = 1$$

- Estymator $\hat{Q}_n$ parametru Q nazywany jest **nieobciążonym** jeśli jego wartość przeciętna równa się parametrowi estymowanemu dla każdego n:

$$\mathbb{E}\left(\hat{Q}_n\right) = Q$$

- Jeśli natomiast $\mathbb{E}\left(\hat{Q}_n\right) \neq Q$ to wtedy $\hat{Q}_n$ nazywany jest **estymatorem obciążonym**, a różnicę $\mathbb{E}\left(\hat{Q}_n\right) - Q$ **obciążeniem estymatora**.
- W przypadku gdy $\lim_{n\to\infty} \mathbb{E}\left(\hat{Q}_n\right) = Q$ to estymator nazywany jest asymptotycznie nieobciążonym.

Za oszacowanie parametru Q można przyjąć różne wielkości. **Dobroć** estymatora $\hat{Q}$ określają pewne kryteria:

- Estymator $\hat{Q}_n(X_1, X_2, \ldots, X_n)$ parametru Q nazywany jest **zgodnym** gdy jest zbieżny wg prawdopodobieństwa do wartości Q, czyli

$$\forall \varepsilon > 0 : \quad \lim_{n\to\infty} P\left(\left|\hat{Q}_n - Q\right| < \varepsilon\right) = 1$$

- Estymator $\hat{Q}_n$ parametru Q nazywany jest **nieobciążonym** jeśli jego wartość przeciętna równa się parametrowi estymowanemu dla każdego n:

$$\mathbb{E}\left(\hat{Q}_n\right) = Q$$

- Jeśli natomiast $\mathbb{E}\left(\hat{Q}_n\right) \neq Q$ to wtedy $\hat{Q}_n$ nazywany jest **estymatorem obciążonym**, a różnicę $\mathbb{E}\left(\hat{Q}_n\right) - Q$ **obciążeniem estymatora**.
- W przypadku gdy $\lim_{n\to\infty} \mathbb{E}\left(\hat{Q}_n\right) = Q$ to estymator nazywany jest **asymptotycznie nieobciążonym**.

- Estymator nieobciążony $\hat{Q}_n$, który ma najmniejszą wariancję spośród wszystkich estymatorów danego parametru Q z prób n-elementowych: $V\left(\hat{Q}_n\right) = \min$, nazywany jest **estymatorem najefektywniejszym**. Oznaczmy go symbolem $\hat{Q}_0$.
- **Efektywność** estymatora to stosunek

$$e\left(\hat{Q}\right) = \frac{V\left(\hat{Q}_0\right)}{V\left(\hat{Q}_n\right)}$$

- **Estymator asymptotycznie najefektywniejszy** to taki estymator, który spełnia

$$\lim_{n\to\infty} e\left(\hat{Q}_n\right) = 1$$

Najbardziej przydatne są estymatory najefektywniejsze; w przypadkach, gdy nie da się ich skonstruować — estymatory asymptotycznie najefektywniejsze.

- Estymator nieobciążony $\hat{Q}_n$, który ma najmniejszą wariancję spośród wszystkich estymatorów danego parametru Q z prób n-elementowych: $V\left(\hat{Q}_n\right) = \min$, nazywany jest **estymatorem najefektywniejszym**. Oznaczmy go symbolem $\hat{Q}_0$.
- **Efektywność** estymatora to stosunek

$$e\left(\hat{Q}\right) = \frac{V\left(\hat{Q}_0\right)}{V\left(\hat{Q}_n\right)}$$

- **Estymator asymptotycznie najefektywniejszy** to taki estymator, który spełnia

$$\lim_{n\to\infty} e\left(\hat{Q}_n\right) = 1$$

Najbardziej przydatne są estymatory najefektywniejsze; w przypadkach, gdy nie da się ich skonstruować — estymatory asymptotycznie najefektywniejsze.

- Estymator nieobciążony $\hat{Q}_n$, który ma najmniejszą wariancję spośród wszystkich estymatorów danego parametru Q z prób n-elementowych: $V\left(\hat{Q}_n\right) = \min$, nazywany jest **estymatorem najefektywniejszym**. Oznaczmy go symbolem $\hat{Q}_0$.
- **Efektywność** estymatora to stosunek

$$e\left(\hat{Q}\right) = \frac{V\left(\hat{Q}_0\right)}{V\left(\hat{Q}_n\right)}$$

- **Estymator asymptotycznie najefektywniejszy** to taki estymator, który spełnia

$$\lim_{n\to\infty} e\left(\hat{Q}_n\right) = 1$$

Najbardziej przydatne są estymatory najefektywniejsze; w przypadkach, gdy nie da się ich skonstruować — estymatory asymptotycznie najefektywniejsze.

- Estymator nieobciążony $\hat{Q}_n$, który ma najmniejszą wariancję spośród wszystkich estymatorów danego parametru Q z prób n-elementowych: $V\left(\hat{Q}_n\right) = \min$, nazywany jest **estymatorem najefektywniejszym**. Oznaczmy go symbolem $\hat{Q}_0$.
- **Efektywność** estymatora to stosunek

$$e\left(\hat{Q}\right) = \frac{V\left(\hat{Q}_0\right)}{V\left(\hat{Q}_n\right)}$$

- **Estymator asymptotycznie najefektywniejszy** to taki estymator, który spełnia

$$\lim_{n\to\infty} e\left(\hat{Q}_n\right) = 1$$

Najbardziej przydatne są estymatory najefektywniejsze; w przypadkach, gdy nie da się ich skonstruować — estymatory asymptotycznie najefektywniejsze.

- Estymator nieobciążony $\hat{Q}_n$, który ma najmniejszą wariancję spośród wszystkich estymatorów danego parametru Q z prób n-elementowych: $V\left(\hat{Q}_n\right) = \min$, nazywany jest **estymatorem najefektywniejszym**. Oznaczmy go symbolem $\hat{Q}_0$.
- **Efektywność** estymatora to stosunek

$$e\left(\hat{Q}\right) = \frac{V\left(\hat{Q}_0\right)}{V\left(\hat{Q}_n\right)}$$

- **Estymator asymptotycznie najefektywniejszy** to taki estymator, który spełnia

$$\lim_{n\to\infty} e\left(\hat{Q}_n\right) = 1$$

Najbardziej przydatne są estymatory najefektywniejsze; w przypadkach, gdy nie da się ich skonstruować — estymatory asymptotycznie najefektywniejsze.

Wybór najefektywniejszego (nieobciążonego) estymatora umożliwia **nierówność Rao-Craméra**:

$$V\left(\hat{Q}_n\right) \geqslant \frac{1}{n\mathbb{E}\left[\left(\frac{\partial \ln f(x;Q)}{\partial Q}\right)^2\right]}$$

która mówi, że precyzja estymatora jest ograniczona informacją zawartą w danych. [Dowód poprzez informację Fishera oraz nierówność Cauchy'ego-Schwarza — pomijamy.]

Wybór najefektywniejszego (nieobciążonego) estymatora umożliwia **nierówność Rao-Craméra**:

$$V\left(\hat{Q}_n\right) \geqslant \frac{1}{n\mathbb{E}\left[\left(\frac{\partial \ln f(x;Q)}{\partial Q}\right)^2\right]}$$

która mówi, że precyzja estymatora jest ograniczona informacją zawartą w danych. [Dowód poprzez informację Fishera oraz nierówność Cauchy'ego-Schwarza — pomijamy.]

Przykład. Z twierdzenia Chinczyna: estymator zgodny wartości oczekiwanej μ to wartość średnia w próbce: $\bar{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$.

Wartość średnia jest estymatorem nieobciążonym:

$$\mathbb{E}\left(\bar{X}\right) = \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\mathbb{E}\left(\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\left(\sum_{k=1}^{n}\mathbb{E}(X_k)\right) = \frac{1}{n}n\mu = \mu$$

Aby sprawdzić, czy jest to estymator najefektywniejszy, przyjmijmy rozkład normalny $\mathcal{N}(\mu, \sigma)$. Wtedy:

$$\ln f(x;\mu,\sigma) = -\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{(x-\mu)^2}{2\sigma^2}, \qquad \frac{\partial \ln f(x;\mu,\sigma)}{\partial\mu} = \frac{x-\mu}{\sigma^2}$$

Aby sprawdzić, czy jest to estymator najefektywniejszy, przyjmijmy rozkład normalny $\mathcal{N}(\mu, \sigma)$. Wtedy:

$$\ln f(x; \mu, \sigma) = -\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{(x-\mu)^2}{2\sigma^2}, \qquad \frac{\partial \ln f(x; \mu, \sigma)}{\partial \mu} = \frac{x-\mu}{\sigma^2}$$

Wg nierówności Rao-Craméra:

$$V\left(\bar{X}\right) \geqslant \frac{1}{n\mathbb{E}\left[\left(\frac{x-\mu}{\sigma^2}\right)^2\right]} = \frac{1}{\frac{n}{\sigma^4}\mathbb{E}\left[(x-\mu)^2\right]} = \frac{1}{\frac{n}{\sigma^4}\sigma^2} = \frac{\sigma^2}{n}$$

Aby sprawdzić, czy jest to estymator najefektywniejszy, przyjmijmy rozkład normalny $\mathcal{N}(\mu, \sigma)$. Wtedy:

$$\ln f(x; \mu, \sigma) = -\ln\left(\sqrt{2\pi\sigma^2}\right) - \frac{(x-\mu)^2}{2\sigma^2}, \qquad \frac{\partial \ln f(x; \mu, \sigma)}{\partial \mu} = \frac{x-\mu}{\sigma^2}$$

Wg nierówności Rao-Craméra:

$$V\left(\bar{X}\right) \geqslant \frac{1}{n\mathbb{E}\left[\left(\frac{x-\mu}{\sigma^2}\right)^2\right]} = \frac{1}{\frac{n}{\sigma^4}\mathbb{E}\left[(x-\mu)^2\right]} = \frac{1}{\frac{n}{\sigma^4}\sigma^2} = \frac{\sigma^2}{n}$$

Ale wariancja wartości średniej dla dowolnego rozkładu (o skończonej wariancji) wynosi $V\left(\bar{X}\right) = \dfrac{\sigma^2}{n}$, zatem **wartość średnia jest najefektywniejszym estymatorem parametru** μ w populacji o rozkładzie normalnym.

Odchylenie standardowe średniej:

$$\sigma\left(\bar{X}\right) = \frac{\sigma}{\sqrt{n}}$$

stanowi błąd estymatora. Im jest on mniejszy (tj. rozkład zmiennej losowej jest bardziej skoncentrowany wokół nieznanej wartości μ w populacji) tym mniejsze jest prawdopodobieństwo odchyleń od prawdziwej wartości μ.

Odchylenie standardowe średniej:

$$\sigma\left(\bar{X}\right) = \frac{\sigma}{\sqrt{n}}$$

stanowi błąd estymatora. Im jest on mniejszy (tj. rozkład zmiennej losowej jest bardziej skoncentrowany wokół nieznanej wartości μ w populacji) tym mniejsze jest prawdopodobieństwo odchyleń od prawdziwej wartości μ.
Dla populacji skończonej o N elementach przy wyborze próby bez zwracania elementy próby są zależne (próbka nie jest próbką prostą).

Odchylenie standardowe średniej:

$$\sigma\left(\bar{X}\right) = \frac{\sigma}{\sqrt{n}}$$

stanowi błąd estymatora. Im jest on mniejszy (tj. rozkład zmiennej losowej jest bardziej skoncentrowany wokół nieznanej wartości μ w populacji) tym mniejsze jest prawdopodobieństwo odchyleń od prawdziwej wartości μ.
Dla populacji skończonej o N elementach przy wyborze próby bez zwracania elementy próby są zależne (próbka nie jest próbką prostą).
Średnia jest nadal dobrym estymatorem wartości przeciętnej natomiast jej wariancja wynosi (pozostawiamy bez dowodu)

$$V\left(\bar{X}\right) = \frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)$$

Estymator wariancji.

Przykład 1. Niech populacja generalna ma dowolny rozkład o znanej wartości przeciętnej μ i nieznanej wariancji $V(X)$. Oznaczmy zmienne losowe Z_i, $i = 1, 2, \ldots, n$, oraz ich średnią jako

$$Z_i = (X_i - \mu)^2; \qquad S_0^2 = \frac{1}{n}\sum_{i=1}^{n} Z_i = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

Z tw. Chinczyna: S_0^2 jako średnia zmiennych Z_i jest stochastycznie zbieżne do $\mathbb{E}(Z_i) = \mathbb{E}\left[(X_i - \mu)^2\right] = \sigma^2$.

Zatem S_0^2 jest estymatorem zgodnym wariancji (przy znanej wartości przeciętnej μ).

S_0^2 jest też estymatorem nieobciążonym:

$$\mathbb{E}\left(S_0^2\right) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[(X_i - \mu)^2\right] = \frac{1}{n} nV(X) = \sigma^2$$

Przykład 2. Niech populacja generalna ma dowolny rozkład o nieznanej wartości przeciętnej μ oraz nieznanej wariancji $V(X)$. Wtedy

$$S_1^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$$

jest estymatorem zgodnym wariancji (dowód podobny do powyższego).

Przykład 2. Niech populacja generalna ma dowolny rozkład o nieznanej wartości przeciętnej μ oraz nieznanej wariancji $V(X)$. Wtedy

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

jest estymatorem zgodnym wariancji (dowód podobny do powyższego). Przekształćmy estymator do postaci:

$$\begin{aligned} S_1^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[(X_i-\mu)+(\mu-\bar{X})\right]^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 + \frac{2}{n}(\mu-\bar{X})\sum_{i=1}^{n}(X_i-\mu) + \frac{1}{n}\sum_{i=1}^{n}(\mu-\bar{X})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - (\bar{X}-\mu)^2 \end{aligned}$$

gdzie $\mu = \mathbb{E}(X)$. Zatem:

$$\begin{aligned}\mathbb{E}(S_1^2) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \mathbb{E}\left[(\bar{X}-\mu)^2\right] \\ &= \frac{1}{n} nV(X) - V\left(\bar{X}\right) \\ &= V(X) - \frac{V(X)}{n} = \frac{n-1}{n}V(X) = \frac{n-1}{n}\sigma^2\end{aligned}$$

Estymator S_1^2 jest zatem obciążony z obciążeniem wynoszącym $-\frac{1}{n}\sigma^2$.

Ponieważ $\lim\limits_{n\to\infty}\dfrac{\sigma^2}{n}=0$, jest to estymator asymptotycznie nieobciążony.

$$
\begin{aligned}
\mathbb{E}(S_1^2) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) - \mathbb{E}\left[(\bar{X}-\mu)^2\right] \\
&= \frac{1}{n}nV(X) - V\left(\bar{X}\right) \\
&= V(X) - \frac{V(X)}{n} = \frac{n-1}{n}V(X) = \frac{n-1}{n}\sigma^2
\end{aligned}
$$

Estymator S_1^2 jest zatem obciążony z obciążeniem wynoszącym $-\frac{1}{n}\sigma^2$.

Ponieważ $\lim\limits_{n\to\infty}\dfrac{\sigma^2}{n}=0$, jest to estymator asymptotycznie nieobciążony.

Mnożąc te estymator przez $\dfrac{n}{n-1}$ otrzymamy nowy estymator:

$$
S_2^2 = \frac{n}{n-1}S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2
$$

który jest **nieobciążonym estymatorem wariancji** — stąd mianownik postaci $n-1$ w powyższym wzorze. S_2^2 jest również estymatorem zgodnym, zaś dla rozkładu normalnego asymptotycznie najefektywniejszym.

Dla rozkładu normalnego błąd odchylenia standardowego S_2 (czyli *odchylenie standardowe odchylenia standardowego*) wynosi (dowód poprzez własności rozkładu χ^2 — pomijamy)

$$\sigma(S_2) = \frac{S_2}{\sqrt{2(n-1)}}$$

Dla rozkładu normalnego błąd odchylenia standardowego S_2 (czyli *odchylenie standardowe odchylenia standardowego*) wynosi (dowód poprzez własności rozkładu χ^2 — pomijamy)

$$\sigma(S_2) = \frac{S_2}{\sqrt{2(n-1)}}$$

Oznacza to, ze tak jak niepewność średniej $\bar{X}$ wyznaczonej z n-elementowej próbki ma niepewność statystyczną wyrażoną jako $\Delta\bar{X} = \frac{\sigma}{\sqrt{n}}$, tak samo odchylenie standardowe S_2 wyznaczone z próbki n-elementowej ma swoją niepewność statystyczną — równą $\Delta S_2 = \frac{\sigma}{\sqrt{2(n-1)}}$.

Dla rozkładu normalnego błąd odchylenia standardowego S_2 (czyli *odchylenie standardowe odchylenia standardowego*) wynosi (dowód poprzez własności rozkładu χ^2 — pomijamy)

$$\sigma(S_2) = \frac{S_2}{\sqrt{2(n-1)}}$$

Oznacza to, ze tak jak niepewność średniej $\bar{X}$ wyznaczonej z n-elementowej próbki ma niepewność statystyczną wyrażoną jako $\Delta\bar{X} = \dfrac{\sigma}{\sqrt{n}}$, tak samo odchylenie standardowe S_2 wyznaczone z próbki n-elementowej ma swoją niepewność statystyczną — równą $\Delta S_2 = \dfrac{\sigma}{\sqrt{2(n-1)}}$.
W obu wypadkach (tj. potrzeby obliczenia zarówno $\Delta\bar{X}$ jak i ΔS_2) gdy nie znamy σ, przyjmujemy $\sigma \approx S_2$ wyznaczone z próbki.