# Квадратично-оптимальні оцінки в класі незміщених. Умови регулярності, нерівність Крамера-Рао.

## Теоретичні відомості та приклади

### Функція вірогідності кратної вибірки

Для простоти розглянемо $(S, \Sigma)$ – такий вибірковий простір, де $S \subset \mathbb{R}^n$ та $\Sigma = \mathcal{B}(S)$ – борелева $\sigma$-алгебра підмножин $S$. Відповідно вибірка $X(\omega) : \Omega \to S$ є випадковим вектором на $S$.

Розглянемо кратну вибірку $X = (X_1, \ldots, X_n) \in S$ з розподілом спостережень $F(t; \theta) = \mathbf{P}_\theta (X_1 < t)$, $\theta \in \Theta \subset \mathbb{R}^d$. Отже сумісний розподіл $X$ має вигляд:

$$F(x_1, \ldots, x_n; \theta) = \mathbf{P}_\theta (X_1 < x_1, \ldots, X_n < x_n) = \prod_{j=1}^{n} \mathbf{P}_\theta (X_1 < x_j) = \prod_{j=1}^{n} F(x_j), x = (x_1, \ldots, x_n)^T \in S$$

Припустимо, що існує така $\sigma$-скінченна міра $\mu$ на просторі значень вибірки $X$, відносно якої міра, породжена функцією розподілу $F(x_1, \ldots, x_n; \theta)$, має щільність $f(x_1, \ldots, x_n; \theta)$ (похідна Радона-Никодима якщо формально).

Наприклад,

1. Якщо розподіл $X$ є дискретним, то $\mu$ – деяка точкова міра. Тоді
$$f(x_1, \ldots, x_n; \theta) = \mathbf{P}_\theta (X_1 = x_1, \ldots, X_n = x_n)$$
2. Якщо розподіл $X$ є абсолютно неперервним, то $\mu$ – міра Лебега. Тоді
$$f(x_1, \ldots, x_n; \theta) = \frac{\partial^n}{\partial x_1 \ldots \partial x_n} F(x_1, \ldots, x_n; \theta)$$

Функцією вірогідності вибірки $X$ називають сумісну щільність її розподілу, тобто

$$L(x_1, \ldots, x_n; \theta) = f(x_1, \ldots, x_n; \theta).$$

Для спрощення $L(x; \theta) = L(x_1, \ldots, x_n; \theta)$, $x = (x_1, \ldots, x_n)^T \in S$.

Емпіричною функцією вірогідності вибірки $X$ називають значення функції вірогідності для заданої вибірки, тобто

$$L(X; \theta) = L(x; \theta)\Big|_{x=X} = L(X_1, \ldots, X_n; \theta)$$

Якщо $X$ – кратна вибірка, то спостереження $X_j$ незалежні в сукупності та мають однаковий розподіл $F(x; \theta) = \mathbf{P}_\theta (X_1 < x)$. Відповідно функція вірогідності $L(X, \theta)$ є добутком щільності розподілу спостереження $f(x; \theta)$:

$$L(x_1, \ldots, x_n, \theta) = \prod_{j=1}^{n} f(x_j; \theta)$$

Наведемо деякі приклади функції вірогідності в залежності від розподілу спостережень.

**Приклад 1.** Нехай $X_1 \sim N(\mu, \sigma^2)$ та припустимо що $\theta = (\mu, \sigma^2) \in \Theta = \mathbb{R} \times (0, \infty)$. Щільність розподілу спостереження має вигляд

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\ x \in \mathbb{R}$$

Отже, $L(x;\theta)$ має вигляд

$$L(x_1,\ldots,x_n,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_j-\mu)^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\right)$$

Ситуація та сама, якщо припускати що невідомим є один із параметрів розподілу.

**Приклад 2.** Нехай $X_1 \sim \mathrm{Pois}(\theta)$, $\theta > 0$. Щільність (відносно точкової міри на $\mathbb{Z}_+$) розподілу $X_1$ має вигляд:

$$f(k;\theta) = \frac{\theta^k}{k!}\exp(-\theta),\ k \in \mathbb{Z}_+.$$

Отже, $L(x;\theta)$ має вигляд

$$L(x_1,\ldots,x_n,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = \prod_{j=1}^{n} \frac{\theta^{x_j}}{x_j!}\exp(-\theta) = \exp(-n\theta)\cdot\theta^{\sum_{j=1}^{n} x_j}\cdot\prod_{j=1}^{n}(x_j!)^{-1}$$

**Приклад 3.** Нехай $X_1 \sim \mathrm{Exp}(\theta)$, $\theta > 0$. Щільність розподілу $X_1$ має вигляд:

$$f(x;\theta) = \mathbf{1}_{(0,\infty)}(x)\theta\exp(-\theta x),\ x \in \mathbb{R}.$$

Отже, $L(x;\theta)$ має вигляд

$$L(x_1,\ldots,x_n,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = \prod_{j=1}^{n} \mathbf{1}_{(0,\infty)}(x_j)\theta\exp(-\theta x_j) = \theta^n \exp\left(-\theta\sum_{j=1}^{n} x_j\right)\mathbf{1}_{(0,\infty)}\left(\min_{1\le k\le n} x_k\right)$$

**Приклад 4.** Нехай $X_1 \sim U[a,b]$, $\theta = (a,b)$, $a < b$. Щільність розподілу $X_1$ має вигляд:

$$f(x;\theta) = \mathbf{1}_{(a,b)}(x)\frac{1}{b-a},\ x \in \mathbb{R}.$$

Отже, $L(x;\theta)$ має вигляд

$$L(x_1,\ldots,x_n,\theta) = \prod_{j=1}^{n} f(x_j;\theta) = \prod_{j=1}^{n} \mathbf{1}_{(a,b)}(x_j)\frac{1}{b-a} = \frac{1}{(b-a)^n}\cdot\mathbf{1}_{(a,b)}\left(\min_{1\le k\le n} x_k\right)\cdot\mathbf{1}_{(a,b)}\left(\max_{1\le k\le n} x_k\right)$$

Функція вірогідності стане незабаром у нагоді.

## Вибір найкращої оцінки в класі незміщених

У рамках попередньої моделі даних припустимо, що потрібно побудувати незміщену оцінку $T = T(X)$ для функції від невідомого параметра $\tau(\theta)$. Нагадаємо, що оцінка є незміщеною для $\tau(\theta)$, якщо для всіх $\theta \in \Theta$: $\mathbf{E}_\theta [T] = \tau(\theta)$.

**Приклад (побудова незміщеної оцінки)** Нехай $X = (X_1, X_2)$ – кратна вибірка з розподілом спостережень $X_1 \sim N(a, \sigma^2)$, де $\theta = (a, \sigma^2) \in \mathbb{R} \times (0, \theta)$ вважається невідомим. Побудуємо незміщену оцінку для $\tau(\theta) = \sigma^2$, користуючись операціями над гауссовими випадковими величинами. Отже, $X_1 - X_2 \sim N(0, 2\sigma^2) \Rightarrow (X_1 - X_2)/\sqrt{2} \sim N(0, \sigma^2)$. Далі, розглянемо $|X_1 - X_2|/\sqrt{2}$. Тоді

$$\mathbf{E}_\theta \left[ \frac{|X_1 - X_2|}{\sqrt{2}} \right] = \frac{2\sigma}{\sqrt{2\pi}} \int\limits_0^{+\infty} x e^{-x^2/2} dx = |u = x^2/2| = \sqrt{2} \frac{\sigma}{\sqrt{\pi}} \int\limits_0^{+\infty} e^{-u} dx = \sqrt{2} \frac{\sigma}{\sqrt{\pi}}$$

Розглянемо $\hat{\sigma}_n = |X_1 - X_2| \cdot (\sqrt{\pi}/2)$. Внаслідок попереднього, $\hat{\sigma}_n$ є незміщеною оцінкою $\sigma$. $\square$

Позначимо через $\Gamma_\tau$ множину всіх незміщених оцінок для $\tau(\theta)$. Оцінки цього класу хороші тим, що в середньому значення будуть коливатися навколо значення невідомого параметра. Інше питання – як сильно коливаються значення тих чи інших незміщених оцінок відносно $\tau(\theta)$?

Було б добре мати таку оцінку, яка мала б найменший розкид відсносно параметра, що оцінюється. Для порівняння характеру розкиду будемо використовувати середньоквадратичну похибку

$$\mathrm{MSE}(T; \tau(\theta)) = \mathbf{E}_\theta \left[ (T - \tau(\theta))^2 \right].$$

Очевидно, що для $T \in \Gamma_\tau$, $\mathrm{MSE}(T; \tau(\theta)) = \mathbf{D}_\theta [T]$.

Якщо існує така оцінка для $\tau(\theta)$, що має найменше середньоквадратичне відхилення, то таку оцінку називають оптимальною. Сформулюємо це формально:

**Означення.** Оцінка $\tilde{T}$ називається оптимальною оцінкою параметра $\tau(\theta)$ в деякому класі оцінок $\Gamma$, якщо для всіх інших оцінок $T \in \Gamma$ для $\tau(\theta)$ виконується

$$\mathrm{MSE}(\tilde{T}; \tau(\theta)) \leq \mathrm{MSE}(T; \tau(\theta)), \forall \theta \in \Theta,$$

причому $\mathrm{MSE}(\tilde{T}; \tau(\theta_*)) < \mathrm{MSE}(T; \tau(\theta_*))$ для деякого $\theta_* \in \Theta$.

Чи можна підібрати таку оцінку в класі незміщених, яка має найменшу дисперсію? Взагалі кажучи, яка 'нижня' межа для дисперсій незміщених оцінок? Відповідь на це питання дає теорема Крамера-Рао, яка дає дві речі: нижню межу на дисперсію незміщених оцінок та критерій існування такої оцінки, дисперсія якої дорівнює цій межі.

Теорема Крамера-Рао має місце за виконання умов регулярності, накладених на функцію вірогідності вибірки. Тому варто розпочати саме з постановки цих умов.

Надалі під $\frac{\partial}{\partial\theta}$ розуміється похідна (у разі скалярного $\theta$) або градієнт (у разі векторного $\theta$) деякої функції по $\theta$.

Введемо функцію впливу $U(X,\theta) = \frac{\partial}{\partial\theta}\ln L(X,\theta)$. Назву можна інтерпретувати так: якщо параметр $\theta$ не впливає на розподіл спостережень, тоді $U(X,\theta) = 0$.

Для кратної вибірки $U(X,\theta) = \sum_{j=1}^n u(X_j,\theta)$, де $u(X_j,\theta) = \frac{\partial}{\partial\theta}\ln f(X_j,\theta)$ – функція впливу за вибіркою з одного спостерження (а розподіл спостережень однаковий).

**Приклад (обчислення функції впливу)** Продовжимо на прикладі $X_j \sim \mathrm{Exp}(\theta)$, інші випадки перевірте самостійно. Оскільки

$$L(x_1,\ldots,x_n,\theta) = \prod_{j=1}^n f(x_j;\theta) = \prod_{j=1}^n \mathbf{1}_{(0,\infty)}(x_j)\theta\exp(-\theta x_j) = \theta^n \exp\left(-\theta\sum_{j=1}^n x_j\right)\mathbf{1}_{(0,\infty)}\left(\min_{1\le k\le n} x_k\right)$$

Відомо, що $S = (0,\infty)^n$. Тобто $X_j > 0$, а тому $\mathbf{1}_{(0,\infty)}\left(\min_{1\le k\le n} X_k\right) = 1$ майже напевно. Тому можемо нехтувати індикатором у функції вірогідності надалі.

Запишемо логарифм функції вірогідності на $S$:

$$\ln L(x,\theta) = n\ln\theta - \theta\sum_{j=1}^n x_j$$

Тоді похідна по $\theta$ має вигляд:

$$\frac{\partial}{\partial\theta}\ln L(x,\theta) = \frac{n}{\theta} - \sum_{j=1}^n x_j$$

В результаті $U(X,\theta) = \dfrac{n}{\theta} - \sum_{j=1}^n X_j$, отримавши підстановкою в аргумент спостережуваних значень у вибірці. $\square$

Перейдемо до визначення інформації за Фішером:

1. Якщо $\theta$ – скалярний параметр, то
$$I(\theta) = \mathbf{D}_\theta\left[U(X,\theta)\right]$$
2. Якщо $\theta$ – векторний параметр, то
$$I(\theta) = \mathrm{Cov}_\theta\left[U(X,\theta)\right],$$
тобто коваріаційна матриця векторної функції впливу $U(X\theta)$.

**Приклад (обчислення інформації за Фішером)** Продовжуємо розглядати приклад з $X_j \sim \mathrm{Exp}(\theta)$. Маємо

$$I(\theta) = \mathbf{D}_\theta\left[U(X,\theta)\right] = \mathbf{D}_\theta\left[n/\theta - \sum_{j=1}^n X_j\right] = \mathbf{D}_\theta\left[\sum_{j=1}^n X_j\right] = \sum_{j=1}^n \mathbf{D}_\theta\left[X_j\right] = \sum_{j=1}^n \theta^{-2} = n\theta^{-2},$$

де скористалися незалежністю та однаковою розподіленістю спостержень, далі обчислили дисперсію експоненційного розподілу. $\square$

**Умови регулярності:**

1. Множина тих значень, коли $L(x,\theta) > 0$ не залежить від $\theta$.
2. $L(x,\theta)$ є двічі неперервно диференційовною за $\theta$.
3. Функція впливу $U(X,\theta)$ є квадратично інтегровною, тобто $\mathbf{E}_\theta\left[(U(X,\theta))^2\right] < \infty$.
4. Порядок інтегрування та взяття похідної по $\theta$ допустимий в інтегралах виду

$$\int\limits_S g(x,\theta)L(x,\theta)\mu(dx),$$

де $g(x,\theta) = 1$ або $g(x,\theta) = \frac{\partial}{\partial\theta}\ln L(x,\theta)$ (вибір таких $g(x,\theta)$ спричинено використання для доведення теореми Крамера-Рао та деяких властивостей функції впливу).

Умови потрібно перевіряти. Це як переходити дорогу без світлофора: не перевірив що коїться, то може збити автобус.

Для дослідження четвертої умови інколи стає у нагоді теорема про зміну порядку інтегрування та диференціювання за параметром в інтегралі Лебега, яку нагадємо прямо тут.

Нехай $(X,\mathcal{F},\lambda)$ – вимірний простір з мірою, $T$ – відкрита підмножина $\mathbb{R}$ та розглянемо функцію $f(x,t) : X \times T \to \overline{\mathbb{R}}$. Покладемо $I(t) = \int_X f(x,t)\lambda(dx)$.
**Теорема.** Нехай справджуються умови:

1. $f(\cdot,t) \in L(X,\lambda)$ для кожного $t \in T$ (тобто $f(x,t)$ інтегровна по $x$),
2. $\frac{\partial f}{\partial t}$ визначена на $X \times T$ (існування похідної),
3. $\left|\frac{\partial f(x,t)}{\partial t}\right| \le g(x)$ для деякого $g \in L(X,\lambda)$ (похідна мажорується інтегровною функцією).

Тоді

$$\frac{dI(t)}{dt} = \int\limits_X \frac{\partial f(x,t)}{\partial t}\lambda(dx).$$

**Приклад (перевірка умов регулярності)** Продовжуємо розглядати приклад з експоненційним розподілом, де припускаємо що $\theta \in [a, b]$, $0 < a < b$. Перевіряємо умови поетапно:

1. $L(x, \theta) = \theta^n \exp\left(-\theta \sum_{j=1}^n x_j\right)$ для $x = (x_1, \ldots, x_n) \in S$. За умовою $\theta \geq a > 0$ та $\exp(\cdot) > 0$ з властивостей експоненційної функції. Отже, $L(x, \theta) > 0$, незалежно від вибору $\theta$.
2. $\mathbf{E}_\theta\left[U(X, \theta)^2\right] < \infty$ – наслідок з обчислень $I(\theta) = \mathbf{D}_\theta\left[U(X, \theta)\right] < \infty$.
3. Для спрощення $n \geq 2$. Тоді
$$\frac{\partial}{\partial \theta} L(x, \theta) = \left(n\theta^{-1} - \sum_{j=1}^n x_j\right) \cdot L(x, \theta) \in C[a, b] \text{ по } \theta$$
$$\frac{\partial^2}{\partial \theta^2} L(x, \theta) = \left(-n\theta^{-2}\right) \cdot L(x, \theta) + \left(n\theta^{-1} - \sum_{j=1}^n x_j\right) \cdot \frac{\partial}{\partial \theta} L(x, \theta) \in C[a, b] \text{ по } \theta$$
Умова на похідні за параметром $\theta$ виконується.
4. Для дослідження останньої умови використаємо теорему про заміну порядку інтегрування та диференціювання.

   Розглянемо $g(x, \theta) = 1$, тобто підінтегральною функцією виступає $L(x, \theta)$. Тоді

   (а) $L(x, \theta) > 0$ та $\int_S L(x, \theta) dx = 1$ як інтеграл від сумісної щільності розподілу вектора $X$. Тому перша умова виконується,

   (б) $\frac{\partial}{\partial \theta} L(x, \theta)$ знаходили під час дослідження другої умови регулярності. Похідна по $\theta$ існує на $S \times [a, b]$.

   (в) Знайдемо мажоранту $g(x)$: для всіх $\theta \in [a, b]$
$$\begin{aligned}\left|\frac{\partial}{\partial \theta} L(x, \theta)\right| &= \left|n\theta^{n-1} - \theta^n \sum_{j=1}^n x_j\right| \cdot \exp\left(-\theta \sum_{j=1}^n x_j\right) \leq \\ &\leq \left(n\theta^{n-1} + \theta^n \sum_{j=1}^n x_j\right) \cdot \exp\left(-\theta \sum_{j=1}^n x_j\right) \leq \\ &\leq \left(nb^{n-1} + b^n \sum_{j=1}^n x_j\right) \cdot \exp\left(-a \sum_{j=1}^n x_j\right) = \\ &= \left(nb^{n-1} + b^n \sum_{j=1}^n x_j\right) a^{-n} \cdot L(x, a) =: g(x)\end{aligned}$$
   Отримана мажоранта, $g(x)$, є інтегровною за побудовою $L(x, \theta)$:
$$\int_S g(x) dx = nb^{n-1} a^{-n} \underbrace{\int_S L(x, a) dx}_{=1} + (b/a)^n \underbrace{\int_S \sum_{j=1}^n x_j L(x, \theta) dx}_{<\infty},$$
   де $\int_S \sum_{j=1}^n x_j L(x, a) dx = \mathbf{E}_a\left[\sum_{j=1}^n X_j\right] = \sum_{j=1}^n \mathbf{E}_a\left[X_j\right] = n/a < \infty$.

Отже, умови теореми про порядок інтегрування та диференціювання мають місце для $L(x,\theta)$. Аналогічно дослідимо $\left(\frac{\partial}{\partial\theta}L(x,\theta)\right)L(x,\theta)$, тобто випадок $g(x,\theta)=\frac{\partial}{\partial\theta}L(x,\theta)$.

(а) $\int_S U(x,\theta)L(x,\theta)dx = \mathbf{E}_\theta\left[U(x,\theta)\right] = \mathbf{E}_\theta\left[n/\theta - \sum_{j=1}^n X_j\right] = n/\theta - n/\theta = 0$,

(б) $\left(\left(\frac{\partial}{\partial\theta}L(x,\theta)\right)L(x,\theta)\right)'_\theta = \left(\frac{\partial^2}{\partial\theta^2}L(x,\theta)\right)L(x,\theta) + \left(\frac{\partial}{\partial\theta}L(x,\theta)\right)^2$ – існує на $S\times[a,b]$.

(в) Можна проробити схожі кроки, що і раніше, для підбору мажоранти $g(x)$ (перевірте самостійно!).

Таким чином довели четверту умову регулярності.

Отже, ми довели виконання умов регулярності для заданої статистичної моделі. $\square$

Припустимо, що умови регулярності на функцію вірогідності виконуються. Тоді можна казати про теорему Крамера-Рао. Будемо розглядати незміщені оцінки для функції від невідомого параметра $\tau(\theta)$.

**Теорема (Крамера-Рао, випадок скалярного параметра)** Нехай $\theta\in\Theta\subset\mathbb{R}$ – невідомий параметр.

1. **Нерівність Крамера-Рао.** Якщо $T=T(X)$ – довільна незміщена оцінка $\tau(\theta)$, існує $\tau_\theta=\frac{d}{dt}\tau(\theta)$ і виконані умови регулярності, то

$$\mathbf{D}_\theta\left[T\right] = \mathbf{E}_\theta\left[(T-\tau(\theta))^2\right] \geq \frac{(\tau_\theta(\theta))^2}{I(\theta)}$$

   для всіх $\theta\in\Theta$.

2. **Критерій Крамера-Рао.** Рівність у нерівності Крамера-Рао виконується тоді і тільки тоді, коли оцінка $T$ є лінійною функцією від функції впливу даних:

$$T(X) = \tau(\theta) + c(\theta)U(X,\theta) \text{ м.н. для всіх } \theta\in\Theta, \tag{1}$$

   де $c(\theta)\in\mathbb{R}$ – деяка стала, що не залежить від даних. Ця стала дорівнює $c(\theta)=\tau_\theta(\theta)/I(\theta)$.

Варто зауважити, що оцінка **не** має містити невідомий параметр, бо інакше це не оцінка. Тому треба дивитися на те, чи дійсна права частина рівності (1) не містить $\theta$.

Оцінка, що задовольняє критерій Крамера-Рао, називається ефективною оцінкою $\tau(\theta)$. Інакше кажучи, це незміщена оцінка $\tau(\theta)$, дисперсія якої дорівнює нижній межі в нерівності Крамера-Рао.

Ефективні оцінки є оптимальними в класі незміщених. Але це не працює в зворотій бік: ефективних оцінок може не бути (наприклад не виконуються умови регулярності, або критерій не дає знайти оцінку).

Наведемо приклад ефективної оцінки в рамках прикладу з експоненційним розподілом спостережень.

**Приклад (застосування критерія Крамера-Рао)** Попередньо доводили, що умови регулярності виконуються. Тому має сенс оперувати теоремою Крамера-Рао.

Перевіримо, чи існує ефективна оцінка для $\tau(\theta) = 1/\theta$. Скористаємося рівністю (1):

$$T(X) = \tau(\theta) + c(\theta)U(X, \theta) = \frac{1}{\theta} + c(\theta)\left(\frac{n}{\theta} - \sum_{j=1}^{n} X_j\right)$$

У цьому випадку $\tau_\theta(\theta) = -1/\theta^2$ та $I(\theta) = n/\theta^2$, тому $c(\theta) = -1/n$. Тоді

$$T(X) = \tau(\theta) + c(\theta)U(X, \theta) = \frac{1}{\theta} - \frac{1}{n}\left(\frac{n}{\theta} - \sum_{j=1}^{n} X_j\right) = \frac{1}{n}\sum_{j=1}^{n} X_j,$$

тобто вибіркове середнє є ефективною оцінкю математичного сподівання $1/\theta$.

Очевидно, що $T(X)$ – незміщена для $\tau(\theta)$. Перевіримо, чи дійсно $\mathbf{D}_\theta\left[T(X)\right] = (\tau_\theta(\theta))^2/I(\theta)$:

- З одного боку, $(\tau_\theta(\theta))^2 = 1/\theta^4$ та $I(\theta) = n/\theta^2$, тому $(\tau_\theta(\theta))^2/I(\theta) = 1/(n\theta^2)$.
- З іншого боку, $\mathbf{D}_\theta\left[T(X)\right] = \mathbf{D}_\theta\left[X_1\right]/n = 1/(n\theta^2)$.

Отже теорема Крамера-Рао нам не збрехала.

Тепер перевіримо, чи існує ефективна оцінка для $\tau(\theta) = \theta$. Скористаємося рівністю (1):

$$T(X) = \tau(\theta) + c(\theta)U(X, \theta) = \theta + c(\theta)\left(\frac{n}{\theta} - \sum_{j=1}^{n} X_j\right) = \left(\theta + c(\theta)\cdot\frac{n}{\theta}\right) - c(\theta)\sum_{j=1}^{n} X_j$$

Тут $\tau_\theta(\theta) = 1$, тому $c(\theta) = \theta^2/n$. Отже

$$T(X) = \left(\theta + \frac{\theta^2}{n}\cdot\frac{n}{\theta}\right) - \frac{\theta^2}{n}\sum_{j=1}^{n} X_j = 2\theta - \frac{\theta^2}{n}\sum_{j=1}^{n} X_j$$

Залежність від $\theta$ наявна, отримана величина не є оцінкою. Отже, ефективної оцінки для $\theta$ не існує.