Домашняя работа

Котов Артем, МОиАД2020

17 ноября 2020 г.

Task

Условие: Про эффективность оценки для выборки из дискретного равномерного распределения.

Решение.

Будем использовать теорему Лемана–Шеффе, но для этого сначала покажем несмещенность оценки:

$$S = \frac{X_{(n)}^{n+1} - (X_{(n)} - 1)^{n+1}}{X_{(n)}^{n} - (X_{(n)} - 1)^{n}}.$$

Для этого посчитаем сначала вероятность $P(X_{(n)} = k) = \sum_{j=1}^{n} \binom{n}{j} (\frac{1}{\theta})^j (\frac{k-1}{\theta})^{n-j} = \frac{k^n - (k-1)^n}{\theta^n}$ (а-ля сумма вероятностей иметь j максимумов иметь значение k, а остальные имеют значение меньше, чем k).

Рассмотрим матожидание S

$$ES = \sum_{k=1}^{\theta} \frac{k^n - (k-1)^n}{\theta^n} \frac{k^{n+1} - (k-1)^{n+1}}{k^n - (k-1)^n} = \sum_{k=1}^{\theta} \frac{k^{n+1} - (k-1)^{n+1}}{\theta^n}$$

В последней сумме заметим, что второй вклад в числителе будет сокращать первый для предыдущего члена суммы, следовательно, выживет только первый вклад числителя для последнего члена суммы, т.е. $\frac{theta^{n+1}}{\theta^n} = \theta$. Таким образом, оценка действительно несмещенная.

Теперь разберемся со статистикой: возьмем в качестве кандидата на подходящую статистку для теоремы Лемана–Шеффе, бросающийся в глаза $T = X_{(n)}$ (вообще можно было бы получить ее из факторизации совместной плотности). Рассмотрим плотность совместного распределения

$$P(X_1 = k_1, \ldots, X_n = k_n) = \underbrace{\frac{1}{\theta^n} 1[X_{(n)} \leq \theta]}_{g_\theta} \underbrace{1[X_{(1)} \geq 1]}_{h}.$$

Т.е. T — достаточная статистика.

Теперь покажем полноту T:

$$0 = E_\theta \varphi(T) = \sum_{k=1}^{\theta} \varphi(k) \frac{k^n - (k-1)^n}{\theta^n}$$

Так как это верно для $\forall \theta \in \mathbb{N}$, то в частности верно и для $\theta = 1$:

$$\varphi(1) \cdot \text{const} = 0 \Longrightarrow \varphi(1) = 0$$

и для $\theta = 2$:

$$\underbrace{\varphi(1)}_{=0} \cdot \ldots + \varphi(2) \cdot \text{const} = 0 \Longrightarrow \varphi(2) = 0$$

и так далее по такой цепочке получим, что $\varphi(k) = 0 \; \forall k \in \mathbb{N}$, то есть T — полная статистика.

Таким образом, у нас в руках сейчас есть некоторая несмещенная оценка $S(T)$ и полная и достаточная статистика T, следовательно по теореме Лемана–Шеффе, такая статистика $S(T)$ единственная (очевидно, она существует, раз мы ее предъявили) и она эффективная. ■