Студент: Елисей Евсеев
Группа: М 4141
Домашняя работа: №2
Дата: 16 марта 2022 г.

Задача 1. Дана выборка $X_1 ... X_n$ с распределением Бернулли с параметром p. Проверить, что
а)X_1
б)$X_1 X_2$
в)$X_1(1 - X_2)$
являются несмещенными оценками соответственно для p, p^2, $p(1-p)$. Являются ли эти оценки состоятельными?

Решение: Несмещенность относительно значения θ для оценки θ^* означает $E[\theta^*] = \theta$. Для распределения Бернулли - с.в. принимает значение 1 с вероятностью p и значение 0 с вероятностью $1-p$. Для проверки состоятельности нужно $\theta^* \xrightarrow{p} \theta$. Будем проверять по определению и используя неравенство Чебышева:

$$\lim_{n\to\infty} P|\theta^* - \theta| \geq \varepsilon = 0$$

$$P(|\theta^* - E\theta^*| \geq \varepsilon) \leq \frac{D\theta^*}{\varepsilon^2}$$

а) Несмещенная, несостоятельная.

$$E[X_1] = 1 \cdot p + 0 \cdot (1-p) = p \tag{1}$$

Т.к. предел точно не равен 0, то оценка несостоятельная

$$P(|X_1 - p| \geq \varepsilon|) \leq \frac{p(1-p)}{\varepsilon^2} > 0 \tag{2}$$

б) Несмещенная, несостоятельная.
Пользуемся независимостью с.в. в выборке X_i:

$$E[X_1 X_2] = E[X_1]E[X_2] = p \cdot p = p^2 \tag{3}$$

$$P(|X_1 X_2 - p^2| \geq \varepsilon) \leq \frac{p^2 - p^4}{\varepsilon^2} \neq 0 \tag{4}$$

в) Несмещенная, несостоятельная.
Снова пользуемся независимостью с.в. в выборке X_i с распределением B_p:

$$E[X_1(1-X_2)] = E[X_1]E[(1-X_2)] = E[X_1](E[1] - E[X_2]) = p(1-p) \tag{5}$$

$$D[X_1 - X_1 X_2] = D[X_1] + D[-X_1 X_2] + 2cov(X_1, (-X_1 X_2)) \tag{6}$$

$$cov(X_1, (-X_1 X_2)) = -E[X_1^2 X_2] + E[X_1]E[X_1 X_2] = 0 \tag{7}$$

$$D[X_1 - X_1 X_2] = p + p^2 \tag{8}$$

Таким образом, эта оценка тоже несостоятельная

Задача 2. Эмпирическая функция распределения $F_n(y)$ строится по выборке из равномерного распределения на отрезке $[0, a]$, где $a > 1$. Для какого параметра $\theta = \theta(a)$ статистика $F_n(1)$ является несмещенной оценкой? Является ли она состоятельной оценкой того же параметра?

Решение: Функция распределения равномерного распределения:

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases} \tag{9}$$

Однако у нас один параметр фиксированный - левая граница отрезка, то есть у нас один параметр θ вместо двух. Получим:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{a} & 0 \le x \le a \\ 1 & x > a \end{cases} \tag{10}$$

$$E[F_n(1)] = F(1) = \frac{1}{a} \tag{11}$$

То есть для параметра $\theta(a) = \frac{1}{a}$ оценка $F_n(1)$ является несмещенной. Проверим состоятельность, сначала узнав дисперсию (пользуемся $F(1)^2 = F(1)$).

$$D[F_n(1)] = \frac{\frac{1}{a} - \frac{1}{a^2}}{n} = \frac{a-1}{na^2} \tag{12}$$

Воспользуясь неравенством Чебышева проверим состоятельность:

$$P(|F_n(1) - \frac{1}{a}| \ge \varepsilon) \le \frac{a-1}{na^2\varepsilon} \tag{13}$$

Т.к. $n \to \infty$, а $a, \varepsilon \in R$ - конечные числа, то это справа все это ограничено 0, то есть предел вероятности равен 0 и оценка - состоятельная (не уверен, что это вышло строго, интересно как нужно строго получить это утверждение).

Задача 3. Пусть $F_n(y)$ — эмпирическая функция распределения, построенная по выборке $X_1 ... X_n$ объёма n. Является ли эмпирической функцией распределения функция (если «да», то какой выборке она соответствует)?

а) $F_n(y^3)$

б) $(F_n(y))^3$

Решение:

а) В данном случае мы видим из определения эмпирической функции распределения, что ее характер (монотонность и кусочно-линейность и пределы области значений $[0, 1]$) не изменятся, то есть сама функция может видоизмениться, «ступеньки» растянутся, но это - эмпирическая функция распределения для выборки, попробуем найти ее явный вид. По определению раскроем $F_n(y^3)$ как сумму индикаторных функций -

$$F_n(y^3) = \frac{1}{n} \sum I(\tilde{X}_i < y^3) \tag{14}$$

однако положив $\tilde{X}_i = \sqrt[3]{X_i}$ получим нужное неравенство $\sqrt[3]{X_i} < y$ для индикаторных функций (аргумент y - линейный).

Таким образом, $F_n(y^3)$ является эмпирической функцией распределения для выборки $\sqrt[3]{X_1} ... \sqrt[3]{X_n}$

б) По определению раскроем $F_n(y)^3$ как сумму индикаторных функций:

$$F_n(y)^3 = \left(\frac{1}{n} \sum I(X_i < y) \right)^3 \tag{15}$$

Общий вид индикаторной функции подходит, но нужно понять, что происходит из-за степени. Как видим всего получится слагаемых n^3, а каждое X_s встречается $s^3-(s-1)^3$ раз (из общего количества слагаемых вычитаем количество слагаемых без рассматриваемого s - ого). Отсюда получаем что это тоже эмпирическая функция распределения для выборки $X_1, X_2, ...X_n$ объема n^3 , где каждое слагаемое X_s встречается $s^3-(s-1)^3$, где $S \in [1, n]$.

Задача 4.

Ноутбук с решением на Google Colab.[1]

[1] https://colab.research.google.com/drive/1APzWO14b5mkV30WcBzGKC4k8Gr5Km9Ac?usp=sharing