Посещаемость семинаров

Опубликовал
sobodv
Автор или источник
2
Предмет
Математическая Статистика (/Subjects/Details?id=5)
Тема
Метод максимального правдоподобия (/Topics/Details?id=31)
Раздел
Введение в ММП (/SubTopics/Details?id=109)
Дата публикации
12.12.2021
Дата последней правки
18.12.2021
Последний вносивший правки
sobodv
Рейтинг
★★★★★

Условие

Вероятность посещения семинара случайно выбранным студентом равняется $p \in (0,1)$. У вас имеется выборка из записей о факте посещения или пропуска семинаров n студентами.

1. При помощи метода максимального правдоподобия оцените параметр p.

2. Проверьте, является ли найденная вами оценка эффективной в классе несмещенных оценок.

Решение

1. Речь идет о выборке из распределения Бернулли $X = (X_1, \ldots, X_n)$, где $X_1 \sim Ber(p)$. Обратим внимание, что при $t \in \{0, 1\}$ функция вероятностей Бернуллиевской случайной величины может быть записана как:

$$P(X_1 = t) = p^t (1-p)^{1-t}$$

В результате получаем функцию правдоподобия:

$$L(p; x) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$

Логарифмируя имеем:

$$\ln L(p;x) = \sum_{i=1}^{n} x_i \ln(p) + (1-x_i)\ln(1-p)$$

Рассмотрим условия первого порядка:

$$\frac{d\ln L(p;x)}{dp} = \frac{1}{p}\sum_{i=1}^{n} x_i - \frac{1}{1-p}\sum_{i=1}^{n}(1-x_i) = 0$$

Решая соответствующее равенство, получаем точку, подозреваемую на максимум:

$$p^* = \overline{x}_n$$

Рассматривая условия второго порядка покажем, что функция правдоподобия является вогнутой, откуда будет следовать, что мы нашли точку максимума:

$$\frac{d^2\ln L(p;x)}{d^2p} = -\frac{1}{p^2}\sum_{i=1}^{n} x_i - \frac{1}{(1-p)^2}\left(n - \sum_{i=1}^{n} x_i\right) < 0$$

В результате получаем ММП оценку:

$$\hat{p}_n = \overline{X}_n$$

2. Убедимся, что найденная оценка является несмещенной:

$$E(\hat{p}_n) = E(\overline{X}_n) = E(X_1) = p$$

Найдем информацию Фишера:

$$i_n(p) = -E\left(-\frac{X_1}{p^2} - \frac{1-X_1}{(1-p)^2}\right) = \frac{E(X_1)}{p^2} + \frac{1-E(X_1)}{(1-p)^2} = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p(1-p)}$$

$$I_n(p) = n i_n(p) = \frac{n}{p(1-p)}$$

Отсюда получаем границу для неравенства Рао-Крамера:

$$\frac{1}{I_n(p)} = \frac{p(1-p)}{n}$$

Убедимся, что дисперсия найденной нами ММП оценки совпадает с соответствующей границей, что докажет эффективность найденной оценки в классе несмещенных оценок:

$$Var(\hat{p}_n) = Var(\overline{X}_n) = \frac{Var(X_1)}{n} = \frac{p(1-p)}{n} = \frac{1}{I_n(p)}$$

Показать решение