Теория Вероятностей и Статистика

Метод максимального правдоподобия и неравенство Рао-Крамера

Потанин Богдан Станиславович

старший преподаватель, кандидат экономических наук

2021

- Продолжим изучать методы, позволяющие получать оценки, обладающие благоприятными свойствами.

Метод максимального правдоподобия

Мотивация

- Продолжим изучать методы, позволяющие получать оценки, обладающие благоприятными свойствами.
- Изученный ранее **метод моментов** хорош тем, что позволяет достаточно просто получить состоятельные оценки.

- Продолжим изучать методы, позволяющие получать оценки, обладающие благоприятными свойствами.
- Изученный ранее **метод моментов** хорош тем, что позволяет достаточно просто получить состоятельные оценки.
- Недостаток данного метода заключается в том, что эффективность его оценок, даже при большом объеме выборки, может оказаться весьма низкой.

- Продолжим изучать методы, позволяющие получать оценки, обладающие благоприятными свойствами.
- Изученный ранее **метод моментов** хорош тем, что позволяет достаточно просто получить состоятельные оценки.
- Недостаток данного метода заключается в том, что эффективность его оценок, даже при большом объеме выборки, может оказаться весьма низкой.
- В качестве альтернативы можно использовать **метод максимального правдоподобия**, оценки которого не только состоятельны, но и эффективны при достаточно больших объемах выборки (асимптотически эффективны).

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

Функция правдоподобия

Дискретный случай

- Имеется выборка $X=(X_1,...,X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x=(x_1,...,x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta;x)=P(X_1=x_1\cap,...,\cap X_n=x_n)=$$

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta; x) = P(X_1 = x_1 \cap, ..., \cap X_n = x_n) = P(X_1 = x_1) \times ... \times P(X_n = x_n) =$$

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta; x) = P(X_1 = x_1 \cap, ..., \cap X_n = x_n) = P(X_1 = x_1) \times ... \times P(X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta; x) = P(X_1 = x_1 \cap, ..., \cap X_n = x_n) = P(X_1 = x_1) \times ... \times P(X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$

Пример: имеется выборка объема n из распределения Пуассона с параметром λ. Запишем выражение для функции правдоподобия:

Функция правдоподобия

Дискретный случай

- Имеется выборка $X=(X_1,...,X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x=(x_1,...,x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta;x)=P(X_1=x_1\cap,...,\cap X_n=x_n)=P(X_1=x_1)\times...\times P(X_n=x_n)=\prod_{i=1}^{n}P(X_i=x_i)$$

Пример: имеется выборка объема n из распределения Пуассона с параметром λ. Запишем выражение для функции правдоподобия:

$$L(\lambda;x)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}=$$

Функция правдоподобия

Дискретный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из дискретного распределения $\Theta(\theta)$ с функцией вероятностей $P(t)$. Также, дана реализация выборки $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из дискретного распределения отражает вероятность получения соответствующей выборки при фиксированном значении параметра распределения:

$$L(\theta; x) = P(X_1 = x_1 \cap, ..., \cap X_n = x_n) = P(X_1 = x_1) \times ... \times P(X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$

Пример: имеется выборка объема n из распределения Пуассона с параметром λ. Запишем выражение для функции правдоподобия:

$$L(\lambda; x) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \frac{\lambda^{\sum\limits_{i=1}^{n} x_i}}{\prod\limits_{i=1}^{n} x_i!} e^{-n\lambda}$$

Функция правдоподобия

Непрерывный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.

Функция правдоподобия

Непрерывный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

Функция правдоподобия
Непрерывный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) =$$

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) = f(x_1) \times ... \times f(x_n) =$$

Функция правдоподобия

Непрерывный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) = f(x_1) \times ... \times f(x_n) = \prod_{i=1}^{n} f(x_i)$$

Функция правдоподобия

Непрерывный случай

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) = f(x_1) \times ... \times f(x_n) = \prod_{i=1}^{n} f(x_i)$$

Пример: имеется выборка объема n из экспоненциального распределения с параметром λ. Запишем выражение для функции правдоподобия:

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) = f(x_1) \times ... \times f(x_n) = \prod_{i=1}^{n} f(x_i)$$

Пример: имеется выборка объема n из экспоненциального распределения с параметром λ. Запишем выражение для функции правдоподобия:

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

- Имеется выборка $X = (X_1, ..., X_n)$ из непрерывного распределения $\Theta(\theta)$ с функцией плотности $f(t)$ и реализацией $x = (x_1, ..., x_n)$.
- **Функция правдоподобия** выборки из непрерывного распределения отражает совместную плотность выборки при фиксированном значении параметра:

$$L(\theta; x) = f_X(x_1, ..., x_n) = f(x_1) \times ... \times f(x_n) = \prod_{i=1}^{n} f(x_i)$$

Пример: имеется выборка объема n из экспоненциального распределения с параметром λ. Запишем выражение для функции правдоподобия:

$$L(\lambda; x) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}$$

Функция правдоподобия

Дополнительные примеры

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t) = (1-p)^{t-1}p$ при $t \in \{1, 2, 3, ...\}$. Запишите функцию правдоподобия данной выборки.

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t) = (1-p)^{t-1}p$ при $t \in \{1, 2, 3, ...\}$. Запишите функцию правдоподобия данной выборки.

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

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t)=(1-p)^{t-1}p$ при $t\in\{1,2,3,...\}$. Запишите функцию правдоподобия данной выборки.

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

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t) = (1-p)^{t-1}p$ при $t \in \{1, 2, 3, ...\}$. Запишите функцию правдоподобия данной выборки.

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

- Имеется выборка объема n из равномерного распределения $U(0, b)$. Сперва для $b \geq \max(x_1, .., x_n)$, а затем для $b < \max(x_1, .., x_n)$ запишите функция правдоподобия данной выборки:

Функция правдоподобия

Дополнительные примеры

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t) = (1-p)^{t-1}p$ при $t \in \{1, 2, 3, ...\}$. Запишите функцию правдоподобия данной выборки.

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

- Имеется выборка объема n из равномерного распределения $U(0, b)$. Сперва для $b \geq \max(x_1, .., x_n)$, а затем для $b < \max(x_1, .., x_n)$ запишите функция правдоподобия данной выборки:

$$L(b; x) = \prod_{i=1}^{n}\frac{1}{b} = \frac{1}{b^n}, \text{ при } b \geq \max(x_1, .., x_n)$$

Функция правдоподобия
Дополнительные примеры

- Имеется выборка объема n из геометрического распределения, то есть с функцией вероятностей $P(t) = (1-p)^{t-1}p$ при $t \in \{1, 2, 3, ...\}$. Запишите функцию правдоподобия данной выборки.

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

- Имеется выборка объема n из равномерного распределения $U(0, b)$. Сперва для $b \geq \max(x_1, .., x_n)$, а затем для $b < \max(x_1, .., x_n)$ запишите функция правдоподобия данной выборки:

$$L(b; x) = \prod_{i=1}^{n}\frac{1}{b} = \frac{1}{b^n}, \text{ при } b \geq \max(x_1, .., x_n)$$

Если по крайней мере в одной точке $b < \max(x_1, .., x_n)$, то существует x_i, такой, что $f(x_i) = 0$, откуда $L(b; x) = 0$.

- Оценка параметра θ может быть получена при помощи метода максимального правдоподобия (ММП оценка) за счет максимизации функции правдоподобия по θ:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ L(\theta; X)$$

- Оценка параметра θ может быть получена при помощи метода максимального правдоподобия (ММП оценка) за счет максимизации функции правдоподобия по θ:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ L(\theta; X)$$

- В случае, если на рассматриваемом множестве параметров функция правдоподобия не может обращаться в ноль (как в равномерном распределении), то для ее максимизации удобно взять логарифм:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ \ln\left(L(\theta; X)\right) = \underset{\theta}{\operatorname{argmax}}\ \ln L(\theta; X)$$

- Оценка параметра θ может быть получена при помощи метода максимального правдоподобия (ММП оценка) за счет максимизации функции правдоподобия по θ:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ L(\theta; X)$$

- В случае, если на рассматриваемом множестве параметров функция правдоподобия не может обращаться в ноль (как в равномерном распределении), то для ее максимизации удобно взять логарифм:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ \ln\left(L(\theta; X)\right) = \underset{\theta}{\operatorname{argmax}}\ \ln L(\theta; X)$$

- При соблюдении некоторых (достаточных) **условий регулярности** последовательность ММП оценок $\hat{\theta}_n$ будет состоятельной.

- Оценка параметра θ может быть получена при помощи метода максимального правдоподобия (ММП оценка) за счет максимизации функции правдоподобия по θ:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ L(\theta; X)$$

- В случае, если на рассматриваемом множестве параметров функция правдоподобия не может обращаться в ноль (как в равномерном распределении), то для ее максимизации удобно взять логарифм:

$$\hat{\theta}_n = \underset{\theta}{\operatorname{argmax}}\ \ln\left(L(\theta; X)\right) = \underset{\theta}{\operatorname{argmax}}\ \ln L(\theta; X)$$

- При соблюдении некоторых (достаточных) **условий регулярности** последовательность ММП оценок $\hat{\theta}_n$ будет состоятельной.
- Условия регулярности довольно сложны с технической точки зрения, поэтому мы их не оговариваем и предполагаем, что они всегда соблюдаются.

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}x_i - n = 0 \implies$$

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}x_i - n = 0 \implies \lambda^* = \frac{\sum\limits_{i=1}^{n}x_i}{n} = \overline{x}_n \implies$$

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}x_i - n = 0 \implies \lambda^* = \frac{\sum\limits_{i=1}^{n}x_i}{n} = \overline{x}_n \implies \hat{\lambda}_n = \overline{X}_n$$

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}x_i - n = 0 \implies \lambda^* = \frac{\sum\limits_{i=1}^{n}x_i}{n} = \overline{x}_n \implies \hat{\lambda}_n = \overline{X}_n$$

- Чтобы досказать, что $\hat{\lambda}_n$ является ММП оценкой, нужно убедиться, что мы нашли максимум, для чего покажем, что функция правдоподобия является вогнутой по оцениваемому параметру:

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) = \sum_{i=1}^{n}\ln\left(\frac{\lambda^{x_i}}{x_i!}e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n}(x_i\ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda)\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}x_i - n = 0 \implies \lambda^* = \frac{\sum\limits_{i=1}^{n}x_i}{n} = \overline{x}_n \implies \hat{\lambda}_n = \overline{X}_n$$

- Чтобы досказать, что $\hat{\lambda}_n$ является ММП оценкой, нужно убедиться, что мы нашли максимум, для чего покажем, что функция правдоподобия является вогнутой по оцениваемому параметру:

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

Метод максимального правдоподобия (ММП)

Пример с распределением Пуассона

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из распределения Пуассона. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\right) = \sum_{i=1}^{n} \ln\left(\frac{\lambda^{x_i}}{x_i!} e^{-\lambda}\right) =$$

$$= \sum_{i=1}^{n} (x_i \ln(\lambda) - \ln(x_i!) - \lambda) = \ln(\lambda) \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln(x_i!) - n\lambda$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\,(lnL(\lambda; x))}{d\lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0 \implies \lambda^* = \frac{\sum\limits_{i=1}^{n} x_i}{n} = \overline{x}_n \implies \hat{\lambda}_n = \overline{X}_n$$

- Чтобы досказать, что $\hat{\lambda}_n$ является ММП оценкой, нужно убедиться, что мы нашли максимум, для чего покажем, что функция правдоподобия является вогнутой по оцениваемому параметру:

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

Примечание: строго говоря случай $x_1 = \ldots = x_n = 0$ следует рассмотреть отдельно: в данной ситуации правдоподобие примет вид $L(\lambda; x) = e^{-n\lambda}$, а значит будет строго убывать по λ, поэтому вновь достигнет максимума в точке $\lambda^* = 0 = \overline{x}_n$.

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n} \lambda e^{-x_i\lambda}\right) =$$

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n}\ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n}x_i$$

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n}\ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n}x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n}\ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n}x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}x_i = 0 \implies$$

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n} \lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n} \ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \implies \lambda^* = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\overline{x}_n} \implies$$

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n} \lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n} \ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \implies \lambda^* = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\overline{x}_n} \implies \hat{\lambda}_n = \frac{1}{\overline{X}_n}$$

Метод максимального правдоподобия (ММП)

Пример с экспоненциальным распределением

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n}\ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n}x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}x_i = 0 \implies \lambda^* = \frac{n}{\sum\limits_{i=1}^{n}x_i} = \frac{1}{\overline{x}_n} \implies \hat{\lambda}_n = \frac{1}{\overline{X}_n}$$

- Чтобы досказать, что $\hat{\lambda}_n$ является ММП оценкой, покажем, что функция правдоподобия является вогнутой по оцениваемому параметру, вследствие чего λ^* – точка максимума:

- При помощи метода максимального правдоподобия найдем оценку параметра λ по выборке из экспоненциального распределения. Логарифм функции правдоподобия будет иметь вид:

$$\ln L(\lambda; x) = \ln\left(\prod_{i=1}^{n}\lambda e^{-x_i\lambda}\right) = \sum_{i=1}^{n}\ln\left(\lambda e^{-\lambda x_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n}x_i$$

- Максимизируем данную функцию, сперва рассмотрев условия первого порядка:

$$\frac{d\left(lnL(\lambda; x)\right)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n}x_i = 0 \implies \lambda^* = \frac{n}{\sum\limits_{i=1}^{n}x_i} = \frac{1}{\overline{x}_n} \implies \hat{\lambda}_n = \frac{1}{\overline{X}_n}$$

- Чтобы досказать, что $\hat{\lambda}_n$ является ММП оценкой, покажем, что функция правдоподобия является вогнутой по оцениваемому параметру, вследствие чего λ^* – точка максимума:

$$\frac{d^2\left(lnL(\lambda; x)\right)}{d^2\lambda} = -\frac{n}{\lambda^2} < 0$$

- При помощи метода максимального правдоподобия найдем оценку параметра b по выборке из равномерного распределения $U(0, b)$, где $b > 0$.

Метод максимального правдоподобия (ММП)

Пример с равномерным распределением

- При помощи метода максимального правдоподобия найдем оценку параметра b по выборке из равномерного распределения $U(0, b)$, где $b > 0$.
- Если $b < \max(x_1, ..., x_n)$, то, как мы ранее показали, функция правдоподобия обращается в ноль, что мотивирует рассматривать лишь случай $b \geq \max(x_1, ..., x_n)$.

- При помощи метода максимального правдоподобия найдем оценку параметра b по выборке из равномерного распределения $U(0, b)$, где $b > 0$.
- Если $b < \max(x_1, ..., x_n)$, то, как мы ранее показали, функция правдоподобия обращается в ноль, что мотивирует рассматривать лишь случай $b \geq \max(x_1, ..., x_n)$.
- Тогда функция правдоподобия принимает вид:

$$L(b; x) = \frac{1}{b^n} = b^{-n}$$

- При помощи метода максимального правдоподобия найдем оценку параметра b по выборке из равномерного распределения $U(0, b)$, где $b > 0$.
- Если $b < \max(x_1, ..., x_n)$, то, как мы ранее показали, функция правдоподобия обращается в ноль, что мотивирует рассматривать лишь случай $b \geq \max(x_1, ..., x_n)$.
- Тогда функция правдоподобия принимает вид:

$$L(b; x) = \frac{1}{b^n} = b^{-n}$$

- Поскольку правдоподобие убывает по параметру b, то с учетом обозначенного ограничения функция достигает максимума при $b^* = \max(x_1, ..., x_n)$, а значит ММП оценка примет вид:

$$\hat{b}_n = \max(X_1, ..., X_n)$$

Информация Фишера

Определение

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta; X_1)\right)}{d\theta}\right)^2\right) =$$

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta; X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta; X_1)\right)}{d^2\theta}\right)$$

Информация Фишера

Определение

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta; X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta; X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta;X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta;X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

Пример: найдем информацию Фишера о параметре λ, содержащуюся в выборке из распределения Пуассона:

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta;X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta;X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

Пример: найдем информацию Фишера о параметре λ, содержащуюся в выборке из распределения Пуассона:

$$lnL(\lambda;X_1) = \ln\left(\frac{\lambda^{X_1}}{X_1!}e^{-\lambda}\right) = X_1\ln(\lambda) - X_1 - \lambda$$

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta;X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta;X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

Пример: найдем информацию Фишера о параметре λ, содержащуюся в выборке из распределения Пуассона:

$$lnL(\lambda;X_1) = \ln\left(\frac{\lambda^{X_1}}{X_1!}e^{-\lambda}\right) = X_1\ln(\lambda) - X_1 - \lambda$$

$$\frac{d\left(lnL(\lambda;X_1)\right)}{d\lambda} = \frac{X_1}{\lambda} - 1 \implies \frac{d^2\left(lnL(\lambda;X_1)\right)}{d^2\lambda} = -\frac{X_1}{\lambda^2}$$

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta; X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta; X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

Пример: найдем информацию Фишера о параметре λ, содержащуюся в выборке из распределения Пуассона:

$$lnL(\lambda; X_1) = \ln\left(\frac{\lambda^{X_1}}{X_1!}e^{-\lambda}\right) = X_1\ln(\lambda) - X_1 - \lambda$$

$$\frac{d\left(lnL(\lambda; X_1)\right)}{d\lambda} = \frac{X_1}{\lambda} - 1 \implies \frac{d^2\left(lnL(\lambda; X_1)\right)}{d^2\lambda} = -\frac{X_1}{\lambda^2}$$

$$i_n(\lambda) = -E\left(-\frac{X_1}{\lambda^2}\right) = \frac{1}{\lambda^2}E(X_1) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} \implies I_n(\lambda) = \frac{n}{\lambda}$$

Информация Фишера

Определение

- **Информация Фишера** о параметре θ, содержащаяся в одном наблюдении, может быть записана двумя эквивалентными способами:

$$i(\theta) = E\left(\left(\frac{d\left(lnL(\theta;X_1)\right)}{d\theta}\right)^2\right) = -E\left(\frac{d^2\left(lnL(\theta;X_1)\right)}{d^2\theta}\right)$$

- Информация Фишера, содержащаяся во всей выборке, определяется как:

$$I_n(\theta) = n \times i(\theta)$$

Пример: найдем информацию Фишера о параметре λ, содержащуюся в выборке из распределения Пуассона:

$$lnL(\lambda;X_1) = \ln\left(\frac{\lambda^{X_1}}{X_1!}e^{-\lambda}\right) = X_1\ln(\lambda) - X_1 - \lambda$$

$$\frac{d\left(lnL(\lambda;X_1)\right)}{d\lambda} = \frac{X_1}{\lambda} - 1 \implies \frac{d^2\left(lnL(\lambda;X_1)\right)}{d^2\lambda} = -\frac{X_1}{\lambda^2}$$

$$i_n(\lambda) = -E\left(-\frac{X_1}{\lambda^2}\right) = \frac{1}{\lambda^2}E(X_1) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda} \implies I_n(\lambda) = \frac{n}{\lambda}$$

Примечание: информация Фишера существует лишь для распределений, у которых носитель не зависит от параметров. Например, для равномерного распределения информация Фишера не определена.

Неравенство Рао-Крамера

Формулировка для несмещенных оценок

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

Неравенство Рао-Крамера

Формулировка для несмещенных оценок

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

- Если данное неравенство соблюдается как равенство, то соответствующая оценка является эффективной в классе несмещенных оценок (полученных по выборке объема n).

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

- Если данное неравенство соблюдается как равенство, то соответствующая оценка является эффективной в классе несмещенных оценок (полученных по выборке объема n).
  Пример: для выборки из распределения Пуассона рассмотрим несмещенную оценку $\hat{\lambda}_n = \overline{X}_n$ и рассчитаем ее дисперсию:

Неравенство Рао-Крамера

Формулировка для несмещенных оценок

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

- Если данное неравенство соблюдается как равенство, то соответствующая оценка является эффективной в классе несмещенных оценок (полученных по выборке объема n).

  Пример: для выборки из распределения Пуассона рассмотрим несмещенную оценку $\hat{\lambda}_n = \overline{X}_n$ и рассчитаем ее дисперсию:

$$Var(\hat{\lambda}_n) = Var(\overline{X}_n) = \frac{Var(X_1)}{n} =$$

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

- Если данное неравенство соблюдается как равенство, то соответствующая оценка является эффективной в классе несмещенных оценок (полученных по выборке объема n).

  Пример: для выборки из распределения Пуассона рассмотрим несмещенную оценку $\hat{\lambda}_n = \overline{X}_n$ и рассчитаем ее дисперсию:

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

Неравенство Рао-Крамера

Формулировка для несмещенных оценок

- Для любой несмещенной оценки $\hat{\theta}$ соблюдается **неравенство Рао-Крамера**:

$$Var(\hat{\theta}_n) \geq \frac{1}{I_n(\theta)}$$

- Если данное неравенство соблюдается как равенство, то соответствующая оценка является эффективной в классе несмещенных оценок (полученных по выборке объема n).

  Пример: для выборки из распределения Пуассона рассмотрим несмещенную оценку $\hat{\lambda}_n = \overline{X}_n$ и рассчитаем ее дисперсию:

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

  Из полученного равенства следует, что оценка $\hat{\lambda}_n = \overline{X}_n$ является эффективной в классе несмещенных оценок.

- Если оценка $\hat{\theta}_n$ была получена при помощи метода максимального правдоподобия, то ее асимптотическое распределение является нормальным (**асимптотическая нормальность**):

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{i(\theta)}\right)$$

- Если оценка $\hat{\theta}_n$ была получена при помощи метода максимального правдоподобия, то ее асимптотическое распределение является нормальным (**асимптотическая нормальность**):

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{i(\theta)}\right)$$

- Поскольку на практике $i(\theta)$ нам неизвестна, то, применив теорему Манна-Вальда, мы можем воспользоваться ее состоятельной оценкой $i(\hat{\theta}_n)$, откуда, в силу теоремы Слуцкого получаем:

$$\sqrt{ni(\hat{\theta}_n)}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, 1\right)$$

Асимптотические нормальность и эффективность ММП оценок

Формулировка

- Если оценка $\hat{\theta}_n$ была получена при помощи метода максимального правдоподобия, то ее асимптотическое распределение является нормальным (**асимптотическая нормальность**):

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{i(\theta)}\right)$$

- Поскольку на практике $i(\theta)$ нам неизвестна, то, применив теорему Манна-Вальда, мы можем воспользоваться ее состоятельной оценкой $i(\hat{\theta}_n)$, откуда, в силу теоремы Слуцкого получаем:

$$\sqrt{ni(\hat{\theta}_n)}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, 1\right)$$

- На практике, при больших $n \geq 100$ как правило предполагают, что:

$$\hat{\theta}_n \dot{\sim} \mathcal{N}\left(\theta, \frac{1}{I_n(\theta)}\right) \qquad \text{или} \qquad \hat{\theta}_n \dot{\sim} \mathcal{N}\left(\theta, \frac{1}{I_n(\hat{\theta}_n)}\right)$$

Асимптотические нормальность и эффективность ММП оценок

Формулировка

- Если оценка $\hat{\theta}_n$ была получена при помощи метода максимального правдоподобия, то ее асимптотическое распределение является нормальным (**асимптотическая нормальность**):

$$\sqrt{n}\left(\hat{\theta}_n-\theta\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{1}{i(\theta)}\right)$$

- Поскольку на практике $i(\theta)$ нам неизвестна, то, применив теорему Манна-Вальда, мы можем воспользоваться ее состоятельной оценкой $i(\hat{\theta}_n)$, откуда, в силу теоремы Слуцкого получаем:

$$\sqrt{ni(\hat{\theta}_n)}\left(\hat{\theta}_n-\theta\right)\xrightarrow{d}\mathcal{N}\left(0,1\right)$$

- На практике, при больших $n\geq 100$ как правило предполагают, что:

$$\hat{\theta}_n\dot{\sim}\mathcal{N}\left(\theta,\frac{1}{I_n(\theta)}\right)\qquad \text{или} \qquad \hat{\theta}_n\dot{\sim}\mathcal{N}\left(\theta,\frac{1}{I_n(\hat{\theta}_n)}\right)$$

- **Асимптотическая дисперсия** ММП оценки записывается как $As.Var(\hat{\theta}_n)=\frac{1}{I_n(\theta)}$ и совпадает с границей неравенства Рао-Крамера, поэтому ММП оценки **асимптотически эффективны**.

Асимптотические нормальность и эффективность ММП оценок

Формулировка

- Если оценка $\hat{\theta}_n$ была получена при помощи метода максимального правдоподобия, то ее асимптотическое распределение является нормальным (**асимптотическая нормальность**):

$$\sqrt{n}\left(\hat{\theta}_n-\theta\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{1}{i(\theta)}\right)$$

- Поскольку на практике $i(\theta)$ нам неизвестна, то, применив теорему Манна-Вальда, мы можем воспользоваться ее состоятельной оценкой $i(\hat{\theta}_n)$, откуда, в силу теоремы Слуцкого получаем:

$$\sqrt{ni(\hat{\theta}_n)}\left(\hat{\theta}_n-\theta\right)\xrightarrow{d}\mathcal{N}\left(0,1\right)$$

- На практике, при больших $n\geq 100$ как правило предполагают, что:

$$\hat{\theta}_n\dot{\sim}\mathcal{N}\left(\theta,\frac{1}{I_n(\theta)}\right)\qquad \text{или} \qquad \hat{\theta}_n\dot{\sim}\mathcal{N}\left(\theta,\frac{1}{I_n(\hat{\theta}_n)}\right)$$

- **Асимптотическая дисперсия** ММП оценки записывается как $As.Var(\hat{\theta}_n)=\frac{1}{I_n(\theta)}$ и совпадает с границей неравенства Рао-Крамера, поэтому ММП оценки **асимптотически эффективны**.
- Состоятельная оценка асимптотической дисперсии ММП оценки имеет вид $\widehat{As.Var}(\hat{\theta}_n)=\frac{1}{ni(\hat{\theta}_n)}$.

Асимптотические нормальность и эффективность ММП оценок

Пример с распределением Пуассона

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.

Асимптотические нормальность и эффективность ММП оценок

Пример с распределением Пуассона

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.

Асимптотические нормальность и эффективность ММП оценок

Пример с распределением Пуассона

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.
- Ее оценкой является $\widehat{As.Var}(\hat{\lambda}_n) = \frac{\overline{X}_n}{n}$ с реализацией $\frac{\overline{x}_n}{n}$.

Асимптотические нормальность и эффективность ММП оценок

Пример с распределением Пуассона

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.
- Ее оценкой является $\widehat{As.Var}(\hat{\lambda}_n) = \frac{\overline{X}_n}{n}$ с реализацией $\frac{\overline{x}_n}{n}$.
- Допустим, что $n = 10000$ и $\hat{\lambda}_{100}(x) = \overline{x}_{100} = 2500$, откуда $(\hat{\lambda}_n - \lambda) \dot{\sim} \mathcal{N}(0, \frac{2500}{10000})$. Тогда вероятность отклонения оценки от истинного значения более, чем на 1, можно приблизительно оценить как:

Асимптотические нормальность и эффективность ММП оценок

Пример с распределением Пуассона

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.
- Ее оценкой является $\widehat{As.Var}(\hat{\lambda}_n) = \frac{\overline{X}_n}{n}$ с реализацией $\frac{\overline{x}_n}{n}$.
- Допустим, что $n = 10000$ и $\hat{\lambda}_{100}(x) = \overline{x}_{100} = 2500$, откуда $(\hat{\lambda}_n - \lambda) \dot{\sim} \mathcal{N}(0, \frac{2500}{10000})$. Тогда вероятность отклонения оценки от истинного значения более, чем на 1, можно приблизительно оценить как:

$$P(|\hat{\lambda}_n - \lambda| > 1) = P(\hat{\lambda}_n - \lambda > 1) + P(\hat{\lambda}_n - \lambda < -1) \approx$$

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.
- Ее оценкой является $\widehat{As.Var}(\hat{\lambda}_n) = \frac{\overline{X}_n}{n}$ с реализацией $\frac{\overline{x}_n}{n}$.
- Допустим, что $n = 10000$ и $\hat{\lambda}_{100}(x) = \overline{x}_{100} = 2500$, откуда $(\hat{\lambda}_n - \lambda) \dot{\sim} \mathcal{N}(0, \frac{2500}{10000})$. Тогда вероятность отклонения оценки от истинного значения более, чем на 1, можно приблизительно оценить как:

$$P(|\hat{\lambda}_n - \lambda| > 1) = P(\hat{\lambda}_n - \lambda > 1) + P(\hat{\lambda}_n - \lambda < -1) \approx$$

$$\approx \left(1 - \Phi\left(\frac{1-0}{\sqrt{2500/10000}}\right)\right) + \left(1 - \Phi\left(\frac{1-0}{\sqrt{2500/10000}}\right)\right) \approx$$

- В случае с выборкой из распределения Пуассона получаем, что для ММП оценки параметра λ можно положить $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\lambda}{n}\right)$ или $\hat{\lambda}_n \dot{\sim} \mathcal{N}\left(\lambda, \frac{\overline{X}_n}{n}\right)$.
- Асимптотическая дисперсия ММП оценки будет иметь вид $As.Var(\hat{\lambda}_n) = \frac{\lambda}{n}$.
- Ее оценкой является $\widehat{As.Var}(\hat{\lambda}_n) = \frac{\overline{X}_n}{n}$ с реализацией $\frac{\overline{x}_n}{n}$.
- Допустим, что $n = 10000$ и $\hat{\lambda}_{100}(x) = \overline{x}_{100} = 2500$, откуда $(\hat{\lambda}_n - \lambda) \dot{\sim} \mathcal{N}(0, \frac{2500}{10000})$. Тогда вероятность отклонения оценки от истинного значения более, чем на 1, можно приблизительно оценить как:

$$P(|\hat{\lambda}_n - \lambda| > 1) = P(\hat{\lambda}_n - \lambda > 1) + P(\hat{\lambda}_n - \lambda < -1) \approx$$

$$\approx \left(1 - \Phi\left(\frac{1-0}{\sqrt{2500/10000}}\right)\right) + \left(1 - \Phi\left(\frac{1-0}{\sqrt{2500/10000}}\right)\right) \approx$$

$$\approx 2 - 2\Phi\left(\frac{1-0}{\sqrt{2500/10000}}\right) = 2 - 2\Phi(2) \approx 0.046$$