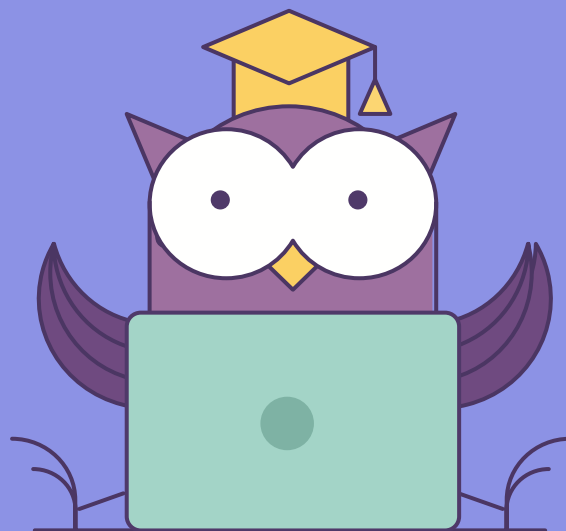




ОНЛАЙН-ОБРАЗОВАНИЕ

Меня хорошо слышно && видно?



Напишите в чат, если есть проблемы!

Ставьте если все хорошо

ММП

Метод максимального правдоподобия.





- Заканчиваю механико-математический факультет МГУ им. Ломоносова
- Учился в Техносфере от Mail.Ru Group
- Являюсь ментором в Техносфере
- Работаю программистом-исследователем в Mail.Ru Group
- Веду лекции открытого курса mlcourse.ai



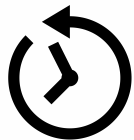
Активно участвуем



Задаем вопросы в чат



Off-topic обсуждаем в Slack
#mathfords-2019-07 или **#general**



Вопросы вижу в чате, могу ответить не сразу

Будем пытаться оценивать неизвестные параметры с помощью математической статистики, но в общем задача остается такой же.

Задача. Вы зашли в метро на незнакомой станции метро. И неожиданно для вас на этой станции метро сейчас мало людей и вы комфортно садитесь в вагон.

Гипотезы

- **Гипотеза 1.** Каждый день в это время на этой станции метро мало людей.
- **Гипотеза 2.** Раз в неделю есть день, когда на этой станции метро мало людей и вы попали именно в этот день.
- **Гипотеза 3.** Раз в месяц есть день, когда на этой станции метро мало людей и вы попали именно в этот день.
- **Гипотеза 4.** Раз в год есть день, когда на этой станции метро мало людей и вы попали именно в этот день.

Какую гипотезу выбрали бы вы?

Теперь будем пытаться оценивать неизвестные параметры с помощью математической статистики, но в общем задача остается такой же.

Задача. Вы зашли в метро на незнакомой станции метро. И неожиданно для вас на этой станции метро сейчас мало людей и вы комфортно садитесь в вагон.

Гипотезы

- **Гипотеза 1.** Каждый день в это время на этой станции метро мало людей.
- **Гипотеза 2.** Раз в неделю есть день, когда на этой станции метро мало людей и вы попали именно в этот день.
- **Гипотеза 3.** Раз в месяц есть день, когда на этой станции метро мало людей и вы попали именно в этот день.
- **Гипотеза 4.** Раз в год есть день, когда на этой станции метро мало людей и вы попали именно в этот день.

Метод максимального правдоподобия выберет гипотезу с максимальной вероятностью, а значит гипотезу 1.

Определение. Пусть есть выборка X_1, \dots, X_n из распределения P_θ , где $\theta \in \Theta$ - неизвестные параметры.

Определение. Назовем $L(\mathbf{x} \mid \theta): \Theta \rightarrow \mathbb{R}$ функцией правдоподобия, где $\mathbf{x} \in \mathbb{R}^n$

$L = \prod p(x_i \mid \theta)$ в случае дискретного распределения

$L = \prod f(x_i \mid \theta)$ в случае непрерывного распределения

Будем искать точечную оценку для параметров

Определение. Точечную оценку $\hat{\theta}_{\text{МП}} = \hat{\theta}_{\text{МП}}(X_1, \dots, X_n) = \operatorname{argmax}_{\theta \in \Theta} L(X_1, \dots, X_n \mid \theta)$

будем называть оценкой максимального правдоподобия параметра θ .

То есть оценка ММП - это такая точечная оценка, при которой функция правдоподобия достигает своего максимума при заданных параметрах.

Задача. К вам на работе приставили стажера. Он задает вам вопросы каждый день. Вы заметили такую зависимость:

$$y_1 = 0, y_2 = 1, y_3 = 2, y_4 = 1$$

Введем такую модель поведения стажера

| | | | |
|----------|-----|------|--------|
| y_i | 0 | 1 | 2 |
| $P(y_i)$ | p | $2p$ | $1-3p$ |

Оценим точечную оценку с помощью метода максимального правдоподобия, считая поведения стажера в разные дни независимо

Решение. Для начала напишем чему будет равна вероятность наших наблюдений с учетом модели поведения стажера:

$$\begin{aligned} P(y_1 = 0; y_2 = 1; y_3 = 2; y_3 = 0) &= \\ P(y_1 = 0)P(y_2 = 1)P(y_3 = 2)P(y_3 = 0) &= \\ p * 2p * (1 - 3p) * p &= 2p^3 (1 - 3p) \end{aligned}$$

Продифференцируем получившееся уравнение и найдем максимум

$$\begin{aligned} (2p^3 (1 - 3p))' &= 6p^2 (1 - 3p) - 6p^3 = \\ 6p^2 - 24p^3 &= 6p^2 (1 - 4p) = 0 \end{aligned}$$

$$\text{Получаем ответ: } p = \frac{1}{4}$$

Решение. Для начала напишем чему будет равна вероятность наших наблюдений с учетом модели поведения стажера:

$$\begin{aligned} L &= P(y_1 = 0; y_2 = 1; y_3 = 2; y_3 = 0) = \\ &P(y_1 = 0)P(y_2 = 1)P(y_3 = 2)P(y_3 = 0) = \\ &p * 2p * (1 - 3p) * p = 2p^3(1 - 3p) \end{aligned}$$

Часто вместо взятия производной от произведения напрямую выражение немного преобразовывают

$$\ln(L) = \ln(2) + 3\ln(p) + \ln(1 - 3p)$$

$$\ln'(L) = \frac{3}{p} - \frac{3}{1-3p} = \frac{3-9p-3p}{p(1-3p)} = \frac{3-12p}{p(1-3p)} = 0$$

$$\text{Получаем ответ: } p = \frac{1}{4}$$

1. Пусть вам задали выборку

$$y_0 = 0, y_1 = 3, y_2 = 2, y_3 = 3, y_4 = 1, y_5 = 1, y_6 = 0$$

И задали модель

| | | | | |
|----------|-----|------|--------|------|
| y_i | 0 | 1 | 2 | 3 |
| $P(y_i)$ | p | $2p$ | $1-5p$ | $2p$ |

$$\hat{p} = \frac{1}{7}$$

Найдите с помощью ММП точечную оценку \hat{p}

2. Пусть вам задана выборка из непрерывного распределения

$$y_1 = 1.1, \dots, y_{100} = 1.5$$

Также известно, что $\sum y_i = 200$

Модель задана распределением $f(y) = \lambda e^{-\lambda y}$

$$\hat{\lambda} = \frac{1}{2}$$

Найдите с помощью ММП точечную оценку $\hat{\lambda}$

y_i **независимы в обеих выборках**

1. Пусть у вас есть выборка X_1, \dots, X_n из распределения Пуассона, при $\lambda > 0$.
Найти оценку ММП $\hat{\lambda}$.

$$f(y) = \frac{\lambda^y}{y!} e^{-\lambda} \quad \hat{\lambda} = \bar{X}$$

Метод максимального правдоподобия обладает несколькими очень полезными свойствами, которые выделяют его на фоне остальных

- Оценки ММП состоятельны, то есть $\hat{\theta}_{ML} \rightarrow \theta$ при $n \rightarrow \infty$
- Оценки ММП асимптотически несмещенные, то есть $M(\hat{\theta}_{ML}) \rightarrow \theta$ при $n \rightarrow \infty$
- Оценки ММП асимптотически эффективны, то есть дисперсия $D(\hat{\theta}_{ML})$ будет наименьшей среди асимптотически несмещенных оценок
- Оценки ММП асимптотически нормальны, то есть $\hat{\theta}_{ML} \sim N(\theta, I^{-1})$ при $n \rightarrow \infty$, где I - информация Фишера, $I = -ln(L''(\theta))$

Есть вопросы или замечания?



Напишите в чат свои вопросы и замечания!

Ставьте если все понятно



АНТОН ЛОСКУТОВ

Mail: antonloskutov@yandex.ru

Telegram: [@LoskutovAnton](https://t.me/LoskutovAnton)

Slack: [@LoskutovAnton](#)

Пройдите опрос



Помогите нам стать лучше!
<https://otus.ru/polls/4422/>

**Спасибо
за внимание!**

