Вероятностные модели, условные вероятности и независимость

Антипов Денис Сергеевич

Университет ИТМО, Санкт-Петербург, Россия

УНИВЕРСИТЕТ ИТМО

11 февраля 2021 г.

О лекторе

- PhD in computer science (Университет ИТМО и École Polytechnique)
- Занимаюсь теорией эволюционных вычислений
- Раньше вел матан, курс по теорверу — впервые
- Контакты:
  - telegram: @antipovden
  - email: antipovden@yandex.ru
- Обращаться ко мне на "ты", по имени (напомню: Денис)
- Не бояться перебивать меня и задавать вопросы

О курсе

- За основу взят курс MIT
- Во многом схож с курсом, который преподавала раньше Ирина Александровна Суслина
- Немного повторяет то, что вы проходили на дискретке
  - Осторожно! Можно случайно не заметить, когда началось что-то совсем для вас новое

Что вы уже знаете

- Множества, операции с ними, законы де Моргана
- Последовательности: пределы, сходимость
- Ряды, порядок суммирования, сумма геометрического ряда
- Счетность и несчетность; почему континуум несчетен
- Мера и ее свойства
- Как интегрировать

План лекции

- Определение вероятностного пространства
- Условная вероятность
- Формула полной вероятности и формула Байеса
- Независимость событий, условная независимость (если успеем)

Часть I. Вероятностное пространство

Вероятностное пространство — это тройка $(\Omega, \Sigma, \mathrm{Pr})$

1. Ω — множество элементарных исходов
2. Σ — σ-алгебра событий
3. Pr — вероятностная мера

Множество элементарных исходов Ω

Мы рассматриваем какой-то эксперимент, у которого возможны разные исходы. Множество всех возможных исходов и есть Ω.

- Эксперимент не может закончиться сразу двумя исходами одновременно (исходы — взаимоисключающие)
- Эксперимент не может закончиться исходом не из Ω (Ω — полное)
- Множество Ω не должно быть черезчур подробным (Ω — неизбыточное)

Примеры Ω

- Подбрасывание монеты
  - $\Omega = \{$орел, решка$\}$
  - Если вы зануда: $\Omega = \{$орел, решка, ребро$\}$
- Хоккейный матч (КХЛ, NHL)
  - $\Omega = \{$В, ВО, ВБ, ПБ, ПО, П$\}$
  - Если нас интересуют только очки одной команды, то $\Omega = \{0, 1, 2\}$
- Время ожидания автобуса на остановке
  - $\Omega = [0, +\infty)$

σ-алгебра событий Σ

Событие — любое подмножество Ω ($\Leftrightarrow$ множество исходов)

На Ω должна быть задана σ-алгебра событий Σ, то есть множество подмножеств, такое, что

1. $\emptyset \in \Sigma$
2. $A \in \Sigma \Rightarrow \bar{A} \in \Sigma$
3. $\{A_1, \ldots, A_n, \ldots\} \in \Sigma \Rightarrow \bigcap\limits_{i \in \mathbb{N}} A_i \in \Sigma$ и $\bigcup\limits_{i \in \mathbb{N}} A_i \in \Sigma$

Примеры событий

- Домашняя команда выиграла $\{$В, ВО, ВБ$\}$
- Монета упала орлом $\{$орел$\}$
- Автобус приехал в течение 5 минут $[0, 5]$ (если считаем время в минутах)

Вероятность Pr

Вероятностная мера Pr — это функция, заданная на Σ, со следующими свойствами

1. Неотрицательность: $\Pr(A) \geq 0$ для любого события A
2. Нормализация: $\Pr(\Omega) = 1$
3. Счетная аддитивнотсь: $A_1, \ldots, A_n, \ldots$ — последовательность попарно непересекающихся событий, тогда

$$\Pr\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \Pr(A_i).$$

NB: В литературе встречаются обозначения $P, p, \mathbb{P}$, их можно использовать

Свойства Pr, следующие из аксиом

- $\Pr(A) \leq 1$
- $\Pr(\emptyset) = 0$
- $\Pr(A) + \Pr(\bar{A}) = 1$
- $A \subset B \Rightarrow \Pr(A) \leq \Pr(B)$
- $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$
- $\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$ — Union bound

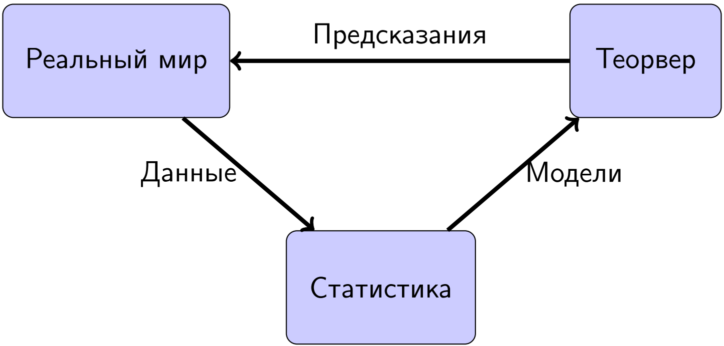
Объект теории вероятности

- Мы будем работать с вероятнсотным пространством $(\Omega, \Sigma, \mathsf{Pr})$:
  - Описывать множество элементарных исходов и определять события
  - Задавать вероятностную меру на этих событиях
  - Делать какие-то выводы о построенной модели
- Будем делать это уже на ближайшей практике ☺

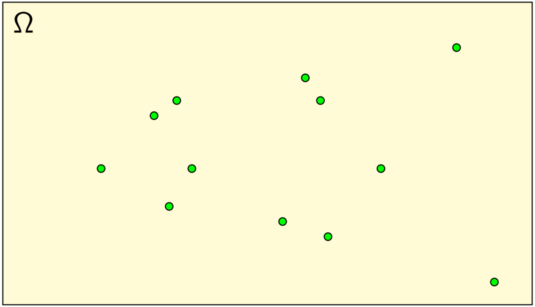
Как воспринимать нашу деятельность

- Теорвер — просто раздел математики
  - На основе определенных аксиом выводим какие-то более сложные утверждения и называем их теоремами
  - Теорема Вероятность события A есть p.
- Как интерпретировать вероятность?
  - Вероятность = частота события (Пример: если много раз кидать монетку, то примерно половина результатов будет “орел”)
  - Вероятность = наша вера в событие (Пример: с вероятностью 0.125 выиграет в этом году кубок Гагарина)

Взаимодействие с реальным миром

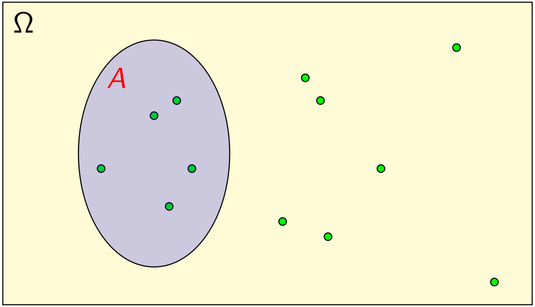


Часть II. Условная вероятность



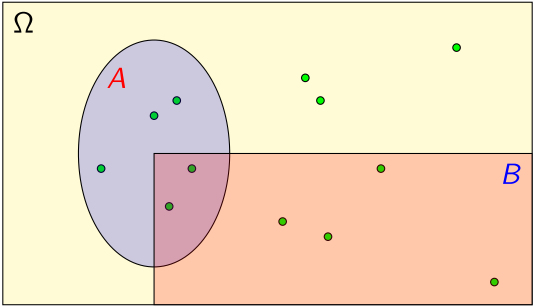
- Пусть все исходы равновероятны (вероятность каждого $\frac{1}{12}$)

Часть II. Условная вероятность



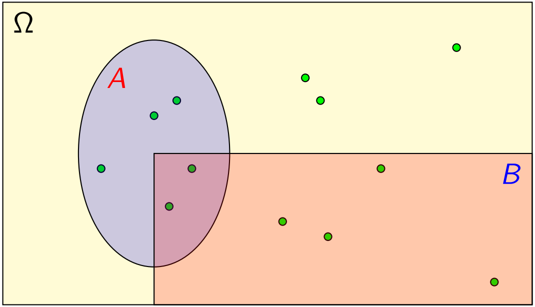
- Пусть все исходы равновероятны (вероятность каждого $\frac{1}{12}$)
- Вероятность события A есть $\Pr(A) = 5 \cdot \frac{1}{12} = \frac{5}{12}$

Часть II. Условная вероятность



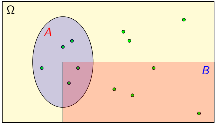
- Пусть все исходы равновероятны (вероятность каждого $\frac{1}{12}$)
- Вероятность события A есть $\Pr(A) = 5 \cdot \frac{1}{12} = \frac{5}{12}$
- Нам стало известно, что произошло событие B. Какой стала вероятность события A?

Часть II. Условная вероятность



- Пусть все исходы равновероятны (вероятность каждого $\frac{1}{12}$)
- Вероятность события A есть $\Pr(A) = 5 \cdot \frac{1}{12} = \frac{5}{12}$
- Нам стало известно, что произошло событие B. Какой стала вероятность события A?
- В событие B входят 6 равновероятных исходов, из которых только 2 входят в событие A. То есть теперь вероятность события A есть $2 \cdot \frac{1}{6} = \frac{1}{3}$.

Поступление новой информации



- При появлении новой информации мы хотим изменить свою модель, чтобы она соответствовала новым условиям
- Пусть мы знаем, что произошло событие B. Тогда мы хотим:
  - $A : A \cap B = \emptyset \Rightarrow \Pr(A \mid B) = 0$
  - $A \subset B \Rightarrow \Pr(A \mid B) = \frac{\Pr(A)}{\Pr(B)}$ (сохраняем нормализацию)
- Любое событие можно представить как $A = (A \cap B) \cup (A \cap \bar{B})$, тогда по аддитивности вероятности

$$\Pr(A \mid B) = \Pr(A \cap B \mid B) + \Pr(A \cap \bar{B} \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)} + 0.$$

Определение условной вероятности

$$\boxed{\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}}$$

$\Pr(A \mid B)$ — Вероятность события A при условии события B

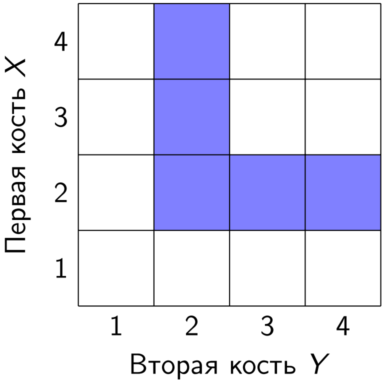
NB: $\Pr(\cdot \mid B)$ — это новая вероятностная мера, заданная на той же самой σ-алгебре, то есть для нее выполняются все аксиомы меры и следствия из них

Пример условной вероятности

Бросок двух тетраидальных костей

Пример условной вероятности

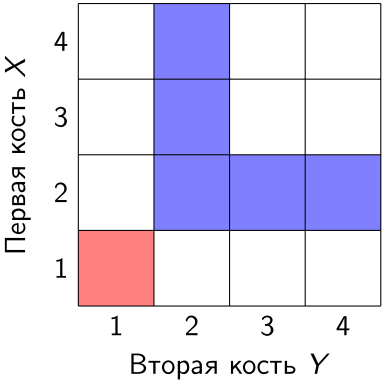
Бросок двух тетраидальных костей



Событие B: $\min\{X, Y\} = 2$

Пример условной вероятности

Бросок двух тетраидальных костей

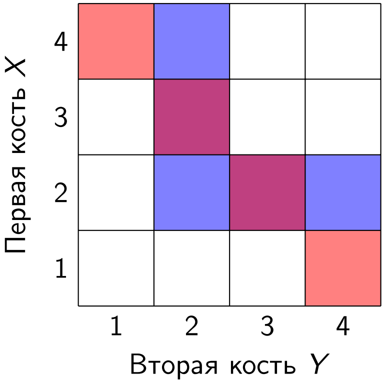


Событие B: $\min\{X, Y\} = 2$

- $\Pr(\max\{X, Y\} = 1 \mid B) = 0$

Пример условной вероятности

Бросок двух тетраидальных костей

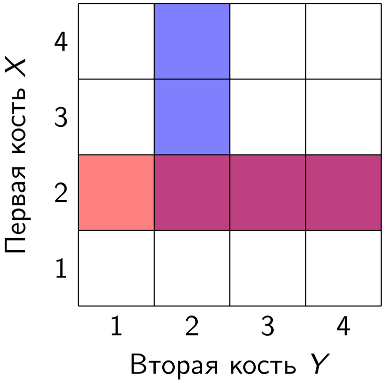


Событие B: $\min\{X, Y\} = 2$

- $\Pr(\max\{X, Y\} = 1 \mid B) = 0$
- $\Pr(X + Y = 5 \mid B) = \frac{2}{5}$

Пример условной вероятности

Бросок двух тетраидальных костей

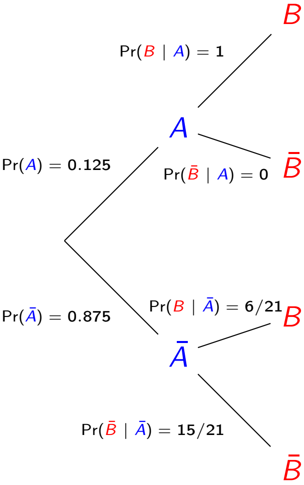


Событие B: $\min\{X, Y\} = 2$

- $\Pr(\max\{X, Y\} = 1 \mid B) = 0$
- $\Pr(X + Y = 5 \mid B) = \frac{2}{5}$
- $\Pr(X = 2 \mid B) = \frac{3}{5}$

Модели с условной вероятностью

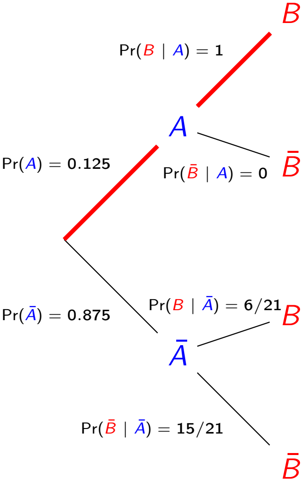
- В HoMM3 у вашего героя экспертный навык удачи 
- Ваш титан  атакует дендройда 
  - Урон титана 40-60 (выбирается равновероятно)
  - Здоровье дендройда 55
  - Пусть навыки атаки и защиты одинаковы
  - Удача срабатывает с вероятностью 0.125 и удваивает урон
- Событие A: сработала удача
- Событие B: дендройд пал



Модели с условной вероятностью

$$\Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)}$$

- $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A) = 0.125$

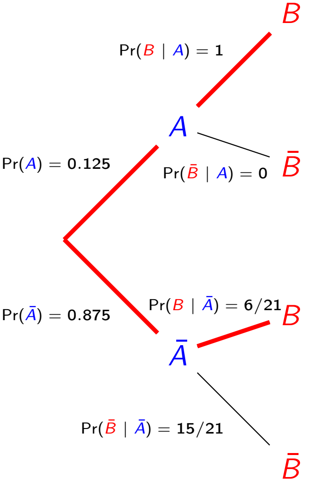


Модели с условной вероятностью

$$\Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)}$$

- $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A) = 0.125$
- $\Pr(B) = \quad \Pr(B \cap A) + \Pr(B \cap \bar{A}) =$

  $\Pr(A)\Pr(B \mid A) + \Pr(\bar{A})\Pr(B \mid \bar{A}) =$
  $0.125 \cdot 1 + 0.875 \cdot \frac{6}{21} = 0.375$

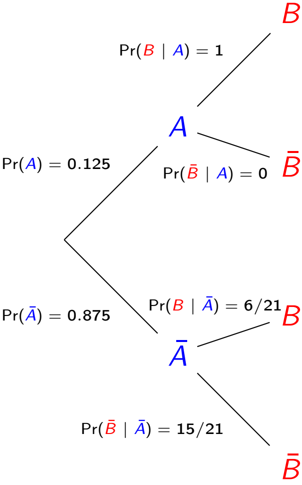


Модели с условной вероятностью

$$\Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)}$$

- $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A) = 0.125$
- $\Pr(B) = \quad \Pr(B \cap A) + \Pr(B \cap \bar{A}) =$

  $\Pr(A)\Pr(B \mid A) + \Pr(\bar{A})\Pr(B \mid \bar{A}) =$
  $0.125 \cdot 1 + 0.875 \cdot \frac{6}{21} = 0.375$
- $\Pr(A \mid B) = \quad \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.125}{0.375} = \frac{1}{3}$



Правило (теорема) умножения вероятностей

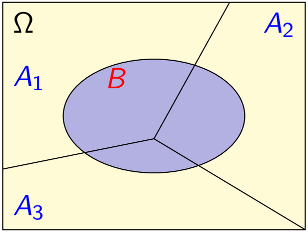
- Мы уже видели, что $\Pr(A \cap B) = \Pr(A)\Pr(B \mid A)$
- Можно обобщить:

$$\Pr\left(\bigcap_{i=1}^{n} A_i\right) = \Pr(A_1)\prod_{i=2}^{n}\Pr(A_i \mid A_1 \cap \dots \cap A_{i-1})$$

- Докажем по индукции

$$\begin{aligned}
\Pr\left(\bigcap_{i=1}^{n} A_i\right) &= \Pr\left(A_n \cap \left(\bigcap_{i=1}^{n-1} A_i\right)\right) \\
&= \Pr\left(\bigcap_{i=1}^{n-1} A_i\right)\Pr\left(A_n \mid \left(\bigcap_{i=1}^{n-1} A_i\right)\right) \\
&= \Pr(A_1)\prod_{i=2}^{n}\Pr(A_i \mid A_1 \cap \dots \cap A_{i-1})
\end{aligned}$$

Правило (теорема) полной вероятности



Есть разбиение Ω на события $A_1, A_2, A_3, \dots$

- Знаем $\Pr(A_i)$ для всех i
- Знаем $\Pr(B \mid A_i)$ для всех i

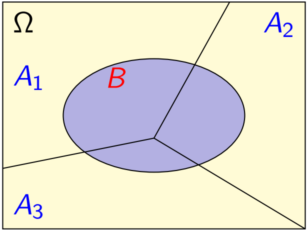
$$\Pr(B) = \sum_i \Pr(A_i) \Pr(B \mid A_i)$$

Доказательство:

- $B = \bigcup_i (B \cap A_i)$ — объединение непересекающихся множеств
- $\Pr(B \cap A_i) = \Pr(A_i) \Pr(B \mid A_i)$ — из правила умножения

NB: Это верно как для конечного, так и для счетного разбиения Ω

Формула Байеса



Есть разбиение Ω на события $A_1, A_2, A_3, \ldots$

- Знаем $\Pr(A_i)$ для всех i
- Знаем $\Pr(B \mid A_i)$ для всех i

$$\Pr(A_i \mid B) = \frac{\Pr(A_i)\Pr(B \mid A_i)}{\sum_j \Pr(A_j)\Pr(B \mid A_j)}$$

Часть III. Независимость

- Бросаем два раза нечестную монету:
  - $\Pr(P) = p \neq 0.5$
  - $\Pr(O) = (1 - p)$
- Получаем один из четырех результатов:

$$\{PP, PO, OP, OO\}$$

- Вероятности этих исходов:
  - $\Pr(PP) = \Pr(P*)\Pr(*P \mid P*) = p \cdot p = p^2$
  - $\Pr(PO) = \Pr(OP) = p(1 - p)$
  - $\Pr(OO) = (1 - p)^2$

Интуитивное понимание независимости

- Какова вероятность, что второй бросок упадет решкой?

$$\Pr(*P) = p$$

- Какова вероятность, что второй бросок упадет решкой, если первый упал орлом?

$$\Pr(*P \mid O*) = p,$$

то есть условие не влияет на вероятность

- Какова вероятность, что второй бросок упадет решкой, если есть ровно одна решка?

$$\Pr(*P \mid PO \cup OP) = \frac{\Pr(OP)}{\Pr(PO \cup OP)} = \frac{p(1-p)}{2p(1-p)} = \frac{1}{2},$$

условие влияет на вероятность

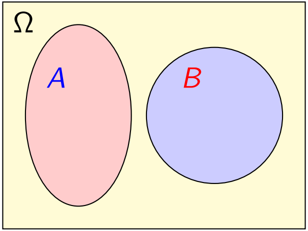
Определение независимости

- Интуитивно: события A и B независимы, если $\Pr(A \mid B) = \Pr(A)$
  - Событие B не несет никакой информации о событии A
  - Но это не работает, если $\Pr(B) = 0$
- Лучше так: события A и B независимы, если

$$\boxed{\Pr(A \cap B) = \Pr(A)\Pr(B)}$$

  - Симметрично относительно событий A и B
  - Из него следует и $\Pr(A \mid B) = \Pr(A)$, и $\Pr(B \mid A) = \Pr(B)$ (если условные вероятности определены)
  - Корректно и при $\Pr(A) = 0$, и при $\Pr(B) = 0$

Типичная ошибка в понимании независимости



События A и B — незавсимы?

Они максимально зависимы: если произошло A, то точно не произошло B, и наоборот

Независимость дополнений

- Если A и B независимы, то независимы и A и $\bar{B}$
  - Интуиция: если "событие B произошло" не дает никакой инфы про A, то и "событие B не произошло" не должно ее давать
  - Формально:

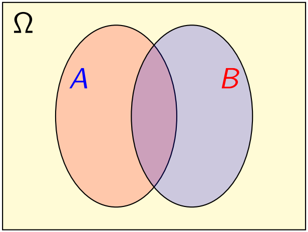
$$
\begin{aligned}
P(A) &= \Pr(A \cap B) + \Pr(A \cap \bar{B}) \\
&= \Pr(A)\Pr(B) + \Pr(A \cap \bar{B}) \\
\Rightarrow \Pr(A \cap \bar{B}) &= P(A) - \Pr(A)\Pr(B) = \Pr(A)(1 - \Pr(B)) \\
&= \Pr(A)\Pr(\bar{B})
\end{aligned}
$$

Условная независимость

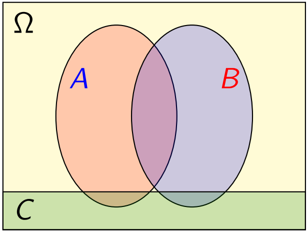
События A и B независимы при условии C, если

$$\Pr(A \cap B \mid C) = \Pr(A \mid C) \Pr(B \mid C)$$

Независимость и условная независимость не особо связаны друг с другом



Пусть A и B — независимы



$A \cap C$ и $B \cap C$ даже не пересекаются

Независимость множества событий

Пусть есть конечный набор событий $\{A_1, \ldots, A_n\}$
Они независимы, если никакой набор событий не влияет на вероятность любого другого события.

$$\Pr(A_1 \cap \bar{A}_4) = \Pr(A_1 \cap \bar{A}_4 \mid A_2 \cap (A_3 \cup \bar{A}_5))$$

разные индексы слева и справа от "|"

События $\{A_1, \ldots, A_n\}$ независимы (по совокупности), если для любого набора индексов $I \subset [1..n]$ верно

$$\Pr\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} \Pr(A_i)$$

Пример для трех событий

Есть события $\{A_1, A_2, A_3\}$. Они независимы, если

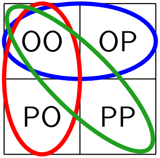
$$\left.\begin{array}{l} \Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2) \\ \Pr(A_1 \cap A_3) = \Pr(A_1)\Pr(A_3) \\ \Pr(A_2 \cap A_3) = \Pr(A_2)\Pr(A_3) \end{array}\right\} \text{Попарная независимость (слабее, чем независимость по совокупности)}$$

- $\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1)\Pr(A_2)\Pr(A_3)$

Попарная vs По совокупности

Бросаем две монеты

- A — первая монета орлом
- B — вторая монета орлом
- C — обе монеты одинаковы



Вероятности событий:

- $\Pr(A) = \frac{1}{2}$
- $\Pr(B) = \frac{1}{2}$
- $\Pr(C) = \frac{1}{2}$

Вероятности комбинаций:

- $\Pr(A \cap B) = \frac{1}{4} = \Pr(A)\Pr(B)$
- $\Pr(A \cap C) = \frac{1}{4} = \Pr(A)\Pr(C)$
- $\Pr(B \cap C) = \frac{1}{4} = \Pr(B)\Pr(C)$
- $\Pr(A \cap B \cap C) = \frac{1}{4} \neq \Pr(A)\Pr(B)\Pr(C)$