Содержание

1 День 0*. Прелюдия, повторяем комбинаторику

2 День 1. Случайные события и элементарное определение вероятности

Назовем *множеством элементарных исходов* $\Omega = \{\omega_1\omega_2, ...\omega_n\}$ такое конечное множеством, что ω_i и ω_j несовместны

Тогда *Событие* - это любое множество элементарных исходов.

Пусть $A \subset \Omega$ - событие, тогда

- $A \cup B = \{\omega \in \Omega : \omega \in A$ или $\omega \in B\}$
- $A \cap B = \{\omega \in \Omega : \omega \in A$ и $\omega \in B\}$
- $\overline{A} = \{\omega \in \Omega : \omega \notin A\}$

Примечание. Рассмотрим эксперимент бросок кубика d6

$\Omega = \{1, 2, 3, 4, 5, 6\}$

Событие, что выпала грань с четным числом - {2,4,6} Событие, что выпала грань с числом меньше 3 - {1,2}

Определим некоторую функцию $P : \Omega \to [0, 1]$, которую назовем *распределением вероятностей* Такую, что

$$\sum_{\omega \in \Omega} P(\omega) = 1$$

Вероятность P(A) *события* A тогда определим как

$$P(A) = \sum_{\omega \in A} P(\omega)$$

Пару (Ω, P) будем называть *Дискретным вероятностным пространством*

Примитивные свойства:

1. $P(\varnothing) = 0$
2. $P(\Omega) = 1$
3. $P(\overline{A}) = 1 - P(A)$
4. Если A и B *не совместны* то есть $A \cap B = \varnothing$ то $P(A \cup B) = P(A) + P(B)$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Вопрос аудитории: Верно ли что нулевая вероятность может быть только у пустого события ($\varnothing$) ?

Теорема 1 (Формула включений-исключений)**.**

$$P(\bigcup_{i=1}^{m} A_i) = \sum_{1}^{m} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) ... + (-1)^{m-1} P(\bigcap_{i=1}^{m} A_i)$$

Доказательство. Будем вести доказательство индукцией по m

База при m = 1 очевидна, при m = 2 сошлемся на пятый пункт, который также очевиден.

Переход от m к m+1:

Пусть $B = \bigcup\limits_{i=1}^{m} A_i$

$$P(\bigcup_{i=1}^{m+1} A_i) = P(B \cup A_{m+1}) = P(B) + P(A_{m+1}) - P(B \cap A_{m+1})$$

Пусть $B_i = A_i \cap A_{m+1} \Rightarrow P(B \cap A_{m+1}) = P(\bigcup\limits_{i=1}^{m} B_i)$

$$\Rightarrow P(B) + P(A_{m+1}) - P(B \cap A_{m+1}) = P(B) + P(A_{m+1}) - P(\bigcup_{i=1}^{m} B_i)$$

$$= (\sum_{1}^{m} P(A_i) - \sum_{i<j\leqslant m} P(A_i \cap A_j) + \sum_{i<j<k\leqslant m} P(A_i \cap A_j \cap A_k) - ...) + P(A_{m+1})$$

$$-(\sum_{1}^{m} P(A_i \cap A_{m+1}) - \sum_{i<j\leqslant m} P(A_i \cap A_j \cap A_{m+1}) + ...)$$

Остается сгруппировать и понять, что это именно то, что нас интересует.

Q.E.D.

Рассмотрим важный частный случай, когда все элементарные исходы равновозможны, то есть $P(\omega_i) = \frac{1}{|\Omega|}$

Тогда вероятность события A равна

$$P(A) = \frac{|A|}{|\Omega|}$$

3 День 2. Условная вероятность и независимые события

Вопрос аудитории: Предположим Степан ученик 11 класса и сдает 3 предмета в этом году, Степан имеет дома обширную коллекцию книг, а в 8 классе районной библиотеке он даже получил награду "Чтец года как вы думаете какой предмет будет более вероятно сдавать Степан в этом году информатику (примечание для преподавателя 21% на 2024) или литературу (примечание для преподавателя 8% на 2024)?

В этом разделе я предлагаю отойти от чуть более общепринятых обозначений A и B для событий и использовать обозначения H (hypothesis) и E (evidence)

3.1 Условная вероятность

Пусть $P(E) > 0$

Тогда Вероятность события H *при условии* что наступило событие E определим как

$$P(H|E) = \frac{P(H \cap E)}{P(E)}$$

То есть мы сужаем наше вероятностное пространство до тех событий которые удовлетворяют некоторому факту

Свойства

1. $P(H|H) = 1$ и если $E \subset H$ то $P(H|E) = 1$
2. $P(\varnothing|E) = 0$
3. Если H_1 и H_2 не совместны, то $P(H_1 \cup H_2|E) = P(H_1|E) + P(H_2|E)$
4. $P(\overline{H}|E) + P(H|E) = 1$

Вопрос аудитории: Будет ли равняться 1 выражение $P(H|E)+P(H|\overline{E})$

3.2 Теорема Байеса

Лемма 1 (Формула Байеса)**.**

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)}$$

Доказательство.

$$P(H|E)P(E) = P(E \cap H) = P(H)P(E|H)$$

Q.E.D.

Формула Байеса позволяет проверять верность некоторой гипотезы при все новых поступающих фактах обрисовывающих контекст текущего эксперимента и позволяют все лучше подбирать модель под данную доменную область, будь то естественные науки или машинное обучение, в свое время байесовские методы достаточно хорошо продвинули данную область.

Можно также расширить данную формулу до полноценной теоремы, если рассуждать о некотором наборе несовместных гипотез:

Пусть есть разбиение множества элементарных исходов на несовместные события $\Omega = \bigsqcup_{i=1}^{n} H_i$ и для каждого события верно $P(H_i) > 0$

Теорема 2 (Формула полной вероянтости)**.** *Тогда иммеет место быть формула:*

$$P(E) = \sum_{i=1}^{n} P(E|H_i)P(H_i)$$

Доказательство. Пусть $E_i = E \cap H_i$ Тогда $E = E \cap \Omega = E \cap \bigsqcup_{i=1}^{n} H_i = \bigsqcup_{i=1}^{n} (E \cap H_i) = \bigsqcup_{i=1}^{n} E_i$ - Разбиение событие A на несовместные события E_i Также заметим, что из Формулы Байеса $P(E_i) = P(E \cap H_i) = P(E|H_i)P(H_i)$

$$\Rightarrow P(E) = \sum_{i=1}^{n} P(E_i) = \sum_{i=1}^{n} P(E|H_i)P(H_i)$$

Q.E.D.

Формула же полной вероятности позволяет нам действовать в обратную сторону и предполагая, что мы имеем некоторую достаточно достоверную модель делать предположения о вероятности наступления некоторого явления в будущем.

Теорема 3 (Теорема Байеса)**.** *Также имеем*

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_{j=1}^{n} P(E|H_j)P(H_j)}$$

Доказательство. Из определения имеем

$$P(E \cap H_i) = P(E|H_i)P(H_i)$$

Из формулы полной вероятности имеем

$$P(E) = \sum_{j=1}^{n} P(E|H_j)P(H_j)$$

Подставим полученные выше значения в определение условной вероятности для $P(H_i|E) = \frac{P(E \cap H_i)}{P(E)}$ И получим как раз то что нам нужно Q.E.D.

3.3 Независимые события

Назовем события A и B *независимыми* если имеет место быть $P(A \cap B) = P(A)P(B)$ И, как следует из семантики слова, верно следующее $P(A|B) = P(A)$ и $P(B|A) = P(B)$

Также заметим, что из независимости A и B следует независимость A и $\overline{B}$

Доказательство.

$$P(A \cap \overline{B}) + P(A \cap B) = P(A) \Rightarrow P(A \cap \overline{B}) = P(A) - P(A \cap B)$$
$$= P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B})$$

Q.E.D.

Вопрос аудитории: Верно ли что если $A \cap B = \varnothing$, то события независисмы?

В качестве домашнего задания подумать о Monty Hall problem на языке на котором мы сегодня говорили.

4 День 3. Случайные величины и вероятностные характеристики

4.1 Случайная величина и ее распределение

Пусть имеется дискретное вероятностное пространство (Ω, P)

Назовем *случайной величиной* некоторый функционал $\xi : \Omega \to \mathbb{R}$.

То есть случайные величины это некоторые численные характеристики, которыми обладает наша модель заданная некоторым вероятностным пространством. Например, эксперимента с подрасыванием монетки случайной величиной будет количество выпавших за эксперимент орлов.

Также пусть $X = \xi(\Omega)$ - *множество значений* случайной величины ξ В таком случае можем рассматривать события вида

$$A_x = \omega \in varOmega : \xi(\omega) = x$$

Тем самым мы получим распределение вероятнотей на множестве X. Обозначим за $P_\xi(x) = P(A_x)$

Легко заметить, что
$$\sum_{x \in X} P_\xi(x) = 1$$

Но тогда из введенных нами определений следует, что пара (Ω, P_ξ) представляет из себя ничто иное, как *дискретное вероятностное пространство.* А функция P_ξ называется *распределением случайной величины* ξ

Весьма логично также будет ввести понятие независимости для случайных величин

Пускай ξ_1, ξ_2 - случайные величины, тогда скажем, что они *независимые*, если

$$\forall x_1, x_2 \in X \Rightarrow P(\{\omega \in \Omega : \xi_1(\omega) = x_1, \xi_2(\omega) = x_2\}) = P_{\xi_1}(x_1)P_{\xi_2}(x_2)$$

5 День 4. Геометрическая вероятность и Метод Монте-Карло

6 День 5. Эпилог, что дальше?

Вероятностные методы решения комбинаторных задач и задач из теории графов? ЦПТ?

7 День 6. Зачет

Post Scriptum