Теория вероятностей в машинном обучение

Копылов Д. Е.

25 июля 2024 года

События и их вероятность

События

Для теории вероятности мир событиен. Событие одно из базовое понятие в теории вероятностей. События обозначают заглавными буквами английского алфавита A, B , ..., Z.

Вероятность события

Вероятность наступления события обозначается $P(A)$ и характеризуется величиной из отрезка $[0, 1]$, где 0 - событие никогда не наступит, 1 - событие обязательно наступит.

Пример событий

A - Идет дождь;
B - Ребенок вернулся с прогулки в 16:30
C - Модель предсказывает результаты с точностью выше 90%

Условная вероятность и формула Байесса

Практически всегда события осуществляются совместно с другими событиями, и часто нелязя перестать наблюдать совместные события. Например $P(\text{А и В})$, A - ребенок вернулся с прогулки в 16:30, B - дождь начался в 16:15.
Иногда бывает интересно рассмотреть как будет вести себя одна величина, при определенном значение другой. Такая вероятность называется условной и обозначается:

$$P(A|B) = \frac{P(\text{А и В})}{P(B)}$$

Замечая особенность:

$$P(A|B) \cdot P(B) = P(\text{А и В})$$

$$P(B|A) \cdot P(A) = P(\text{А и В})$$

$$P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \tag{1}$$

Совместные и несовместные события

Событие несовместные, если $P(\text{А и В}) = P(A)P(B)$, т.е. $P(A|B) = P(A)$, $P(B|A) = P(B)$

Для совместных событий операция ”или”

Выполнение хотя бы одного из событий (”или”):

$$P(\text{А или В}) = P(A) + P(B) - P(\text{А и В})$$

Для несовместных событий операция ”или”

Выполнение хотя бы одного из событий (”или”):

$$P(\text{А или В}) = P(A) + P(B)$$

Случайные величины

Не ясность

Иногда бывает, что ситуацию приходится описывать сразу большим множеством событий.
A_1 – дождь начнется в 0:00,
A_2 - дождь начнется в 0:01,
...,
A_{1440} - дождь начнется в 23:59.
Упростить себе жизнь можно введя параметр ξ. Теперь можно сказать:
"Дождь начнется через ξ минут после наступления полночи".

Случайная величина

В таком случае случайное не событие (которое наступило или нет), а величина.

Если вероятность $P(A_i) = p_i, \quad i = 1, 2, ..., 1440$,
то вероятность $P(\xi = i) = p_i, \quad i = 1, 2, ..., 1440$.
Теперь мы говорим об вероятностях случайной величены, а не о вероятносте наступления события

В реальносте же можно говорить об вещественных числах, тогда событие невозможно описать просто перечислев варианты: Температура воздуха будет ξ градусов (без округления).

Ястно, что температура *почти наверное никогда*, не будет равняться 17 градусам ровно. В случае вещественных чисел говорят об плотносте распределения величины и интервале в котором она может лежать.

Функция плотности вероятности в случае вещественной случайной величины

$$P(\xi \in (\omega_1, \omega_2)) = \int_{\omega_1}^{\omega_2} p(\xi) d\xi, \tag{2}$$

$$\int_{-\infty}^{+\infty} p(\xi) d\xi = 1, \quad p(\xi) \geq 0 \tag{3}$$

$p(\xi)$ - функция плотности распределения.

Когда говорим о дискретных величинах с индексами из I, требуется выполнения следующих условий:

$$\sum_{i \in I} p_i = 1, \quad p_i \geq 0$$

Функции от случайных величин

Со случайными величинами можно совершать привычные операции, только ответ также становится случайной величиной.

$$\eta = f(\xi)$$

Для того, чтоб можно было оценивать и случайные функции, вводятся понятие ”математическое ожидание”:

$$\mathbb{E}_{\xi \sim P} f(\xi) = \int_{-\infty}^{+\infty} p(\xi) f(\xi) d\xi \tag{4}$$

Для дискретного случая имеет место понятие среднее значение:

$$\mathbb{E}_{\xi \sim P} f(\xi) = \frac{1}{n} \sum_{i=1}^{n} p(\xi) f(\xi) \tag{5}$$

bibliography I