Теория вероятности и математическая статистика—1

Винер Даниил @danya_vin

Версия от 9 сентября 2024 г.

Содержание

1 Дискретное вероятностное пространство. Базовые теоремы вероятности

Определение. $\Omega = \{\omega_1, \ldots, \omega_k, \ldots\}$ называется *пространством элементарных исходов*, где w_i — элементарный исход

Определение. A — любое подмножество Ω

Определение. Событие называется *достоверным*, если $A = \Omega$

Примечание. К A применимы те же опреации, что используются с множествами

Определение. Полная группа несовместных событий — такой набор событий, для которого выполняются такие условия:

$$A_i \cap A_j = \varnothing \ \forall i \neq j$$

$$\bigcup_i A_i = \Omega$$

Аксиома. $\forall \omega_i \ \exists p_i \geqslant 0$, при этом $\sum\limits_i p_i = 1$

Следствие. $0 \leqslant p_i \leqslant 1$

Определение. $P(A) = \sum\limits_{w_i \in A} P(w_i)$, где $P(w_i) = p_i$

(Ω, P) — вероятностное пространство в дискретном случае

Подходы к определению вероятностей

1. Априорный (предварительное знание)
2. Частотный (предел ряда частот)
3. Модельный (математическая модель)

1.1 Классическое определение вероятности

Имеет место, когда исходы равновероятны

Определение.

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

Определение. $P(A) = \sum\limits_{\omega_i \in A} p(\omega_i)$

1.2 Теорема сложения

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A_1 \cup \ldots \cup A_n) = \sum\limits_{i=1}^{n} P(A_i) - \sum\limits_{i<j} (A_i \cap A_j) + \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n)$

1.3 Условная вероятность

$P(A|B) = \dfrac{P(A \cap B)}{P(B)} \ \forall B : P(B) > 0$

1.4 Теорема умножения

$P(A_1 \cap \ldots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \ldots \cdot P(A_n|A_1 \cdot \ldots \cdot A_{n-1})$

2 Независимые события. Апостериорная и условная вероятности

Напомним классическое определение вероятности

$$P(A) = \sum_{\omega_i \in A} p(\omega_i)$$

Примечание. $P(\overline{A}) = 1 - P(A),\ P(\Omega) = 1$

2.1 Задача об авариях

Известно, что в 40% аварий виноваты пьяные водители. Утверждается, что из этого следует, что 60% виноваты трезвые водители

Пусть есть такие события: A — водитель пьян, B — водитель трезв, C — аврия случилась

Формально, задача выглядит так: $P(A|C) = 0.4 \Longrightarrow P(B|C) = 0.6$

Пусть $P(B) = 0.05$

$\dfrac{P(C|A)}{P(C|B)} = \dfrac{P(C \cap A)/P(A)}{P(C \cap B)} = \dfrac{P(A|C)P(C)}{P(B|C)P(C)} \cdot \dfrac{P(B)}{P(A)} = \dfrac{0.4}{0.6} \cdot \dfrac{0.95}{0.05} \approx 12.7$

2.2 Независимость событий

Определение. (интуитивное) События A и B **независимы**, если $P(A|B) = P(A)$

Определение. События A и B называются попарно независимыми, если:

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A|B)P(B) = P(A) \cdot P(B) \text{ — вытекает интуитивное определение}$$

Определение. События $A_1 \dots A_n$ независимы в соовокупности, если:

$$\forall i_1 < \dots < i_k < \dots < i_n\ \forall k = 1, \dots, n:$$
$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$$

Примечание. Для A_1, A_2, A_3:

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$$
$$P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$$
$$P(A_1 \cap A_3) = P(A_1) \cdot P(A_3)$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

2.3 Пример зависимых событий в совокупности

Положим, что у нас есть тетраэдр, каждая сторона которого покрашена в некоторые комбинации цветов: A — красный, B — желтый, C — зеленый, а четвертая покрашена во все три цвета, тогда вероятности:

$$\frac{1}{4} = P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C) = \frac{1}{8}$$

Таким образом, события зависимы в совокупности

Примечание. Если $A_1, \dots, A_n$ независимы в совокупности, то над любым из событий можно поставить знак отрицания, тогда система останется независимой

Пример. Есть события A_1, A_2, A_3 — независимы в совокупности, то $\overline{A_1}, A_2, A_3$ — тоже независимы в совокупности

Проверим данный пример.

$$
\begin{aligned}
P(\overline{A_1} \cap A_2 \cap A_3) &= P((A_2 \cap A_1) \setminus A_1) \\
&= P(A_2)P(A_3) - P(A_1 \cap A_2 \cap A_3) \\
&= P(A_2)P(A_3) - P(A_1)P(A_2)P(A_3) \\
&= P(\overline{A_1})P(A_2)P(A_3)
\end{aligned}
$$

Теперь разберемся с двумя событиями

$$
\begin{aligned}
P(\overline{A_1} \cap A_2) &= P(\overline{A_1})P(A_2) \\
P(\overline{A_2} \cap A_1) &= P(A_2 \setminus A_1) \\
&= P(A_2) - P(A_2 \cap A_1) = P(A_2) - P(A_2)P(A_1) \\
&= P(A_2)(1 - P(A_2)) \\
&= P(A_1)P(A_2)P(A_3)
\end{aligned}
$$

2.4 Простейший варинат ЗБЧ. Неизбежность технологических катастроф

Имеется n узлов, а $A_1, \ldots, A_n$ — события, где A_i означает, что i-ый узел вышел из строя

Очевидно, что P(хотя бы один узел выйдет из строя) $= P(A_1 \cup \ldots \cup A_n)$, тогда $\forall i : 0 < \varepsilon \leqslant P(A_i) < 1$ выполняется:

$$
1 \geqslant P(A_1 \cup \ldots \cup A_n) = 1 - P(\overline{A_1} \cap \ldots \cap \overline{A_n}) = 1 - \prod_{i=1}^{n} \underbrace{P(\overline{A_i})}_{\leqslant 1-\varepsilon} \geqslant \lim_{n \mapsto \infty} 1 - (1-\varepsilon)^n = 1
$$

Это означает, что вероятность технологических катастроф при количестве узлов, стремящемся в бесконечность равна 100%

2.5 Формула полной вероятности

Пусть $\{H_i\}$ — полная группа несовместных событий (разбиение Ω)

Некоторые свойства:

- $H_i \cap H_j = \varnothing \; \forall i \neq j$ — несовместность
- $\bigcup\limits_{i=1}^{n} H_i = \Omega$ — полнота

Теорема. Тогда, $P(A) = \sum\limits_{i=1}^{n} P(A|H_i)P(H_i)$

Доказательство.

$$
\begin{aligned}
P(A) &= P\left(\bigcup_{i=1}^{n} (A \cap H_i)\right) \\
&= \sum_{i=1}^{n} P(A \cap H_i) \\
&= \sum_{i=1}^{n} P(A|H_i) \cdot P(H_i)
\end{aligned}
$$

□

2.6 Формула Байеса. Апосториорные вероятности

$$
\begin{aligned}
P(H_k|A) &= \frac{P(A|H_k) \cdot P(H_k)}{P(A)} \\
&= \frac{P(A|H_k) \cdot P(H_k)}{\sum\limits_{i=1}^{n} P(A|H_i)P(H_i)}
\end{aligned}
$$

2.7 Задача про неизличимые заболевания

Требуется вычислить, какова вероятность, что человек, который получил положительный тест на СПИД, на самом деле не болен

$P(\text{СПИД}) = 0.03$ — вероятность быть носителем СПИДа

$P(+|\text{ СПИД}) = 0.98$ — чувствительность теста, т.е. тест будет положительным, если у человека есть СПИД, с вероятностью 0.98

$P(+|\overline{\text{СПИД}}) = 0.01 = 1 -$ специфичность, т.е. тест будет положительным, если у человека нет СПИДа, с вероятностью 0.01

$$
\begin{aligned}
P(\text{СПИД}|+) &= \frac{P(+|\text{СПИД})P(\text{СПИД})}{P(+|\text{СПИД})P(\text{СПИД}) + P(+|\overline{\text{СПИД}})P(\overline{\text{СПИД}})} \\
&= \frac{0.98 \cdot 0.03}{0.98 \cdot 0.03 + 0.01 \cdot 0.97} \\
&= 0.75
\end{aligned}
$$

То есть, **каждый четвертый** человек с положительным СПИД-тестом **здоров**