Sample Space and Probability

2023.7.10

Nemo

Contents

1 Sets (Quick Review)

Set, element, empety set $\emptyset$, finit set, countably finit set, uncountable set, sub set, equal, universal set Ω

1.1 Set opeartions:

1. Complement of a set S, with respect it the universe Ω, denoted by S^c
2. Union of two sets S, T, $S \cup T$
3. Intersection of two sets S, T, $S \cap T$
4. Union of several, $\bigcup_{n=1}^{\infty} S_n = S_1 \cup S_2 \cup \ldots$
5. Intersection of several, $\bigcap_{n=1}^{\infty} S_n = S_1 \cap S_2 \cap \ldots$
6. Sets are Disjoint if they share no element
7. A collection of sets is a partition of set S, if they are disjoint and the union of them are S

1.2 The Algebra of Sets:

De Morgan's laws:

$$\left(\bigcup_n S_n\right)^c = \bigcap_n S_n^c$$

$$\left(\bigcap_n S_n\right)^c = \bigcup_n S_n^c$$

2 Probabilistic Models

Elements of a Probabilistic Model

- The sample space Ω, the set of all possiable outcomes
- The probability law, which assigns any event A a non-negative number $P(A)$

*

2.1 Choosing an Appropriate Sample Space

The element of the sample space should be distinct and *mutually exclurive*, and the sample space should be collectively exhaustive.

2.2 Probability Axioms

1. **(Nonnegativity)** $P(A) \geq 0$, for every event A
2. **(Additivity)** A, B are disjoint, then $P(A \cup B) = P(A) + P(B)$
3. **(Normalization)** $P(\Omega) = 1$

2.3 Discrete Models

e.g. The toss of a coin several times *Like {HHH,HHT,HTH,HTT,THH,THT,TTH,TTT}(3 times) and the probability stuff*

Discrete Probability Law

The sample space $S = \{s_1, s_2, s_3, ..., s_n\}$ consists of finite number of elements, we have:

$$P(S) = P(\{s_1, s_2, s_3, ..., s_n\}) = P(s_1) + P(s_2) + P(s_3) + ... + P(s_n)$$

Discrete Uniform Probability Law

Ii the outcomes are equally likely, then the Probability of any single outcome A becomes:

$$P(A) = \frac{\text{number of elements of } A}{n}$$

2.4 Continuous Models

Like throughing a dart on a certian area or sth else ...

2.5 Properties of Probability Laws

1. If $A \in B$, then $P(A) \leq P(B)$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
3. $P(A \cup B) \leq P(A) + P(B)$
4. $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$

* Insight of Probability: The term "probability" should come with an event, like the probability of event A $P(A)$, which is further a outcome of the probability law and a part of the probabilistic model. And a valid probabilistic model should contain a sample space and a probability law which agree with the probability axioms.

3 Conditional Prabability

Conditional probability provides us with a way to reason about the outcome of an ekperiment, based on **parcial information**. (The experiment is done and the only have some parcial information about it.)

e.g. The experiment involving two successive rolls of a die, you are toled that the sum of the two rolls are 9. What's the probability of the first roll is a 6?

In precise terms, the conditional probability is when we know the is with in a given event B, we wish to know the probability of the event A. We call this *conditional probability of A given B*, denoted by $P(A \mid B)$

Definition$_{\textit{conditional probability}}$: $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

3.1 Verification of the Prabability Laws

1. Nonnegativity is clear since the original probability is nonnegative.
2. Additivity:

$$\begin{aligned} P(A_1 \cup A_2 \mid B) &= \frac{P((A_1 \cup A_2) \cap B)}{P(B)} \\ &= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} \\ &= P(A_1 \mid B) + P(A_2 \mid B) \end{aligned}$$

1. Normalization:

$$P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

3.2 Properties of Conditional Probability

- Conditional probability can be viewed as a normal probability on a new universe B.
- Furthermore, if all outcomes are equally likely, then $P(A \mid B) = \frac{\text{num of elements of } A \cup B}{\text{num of elements of } B}$

3.3 Using Conditional Probability for Modeling

A restatement of the definition of the conditional probability is $P(A \cup B) = P(B)P(A \mid B)$, which can be used to calculate a non-conditional probability.

Multiplication Rule

By definition, it's easy to get

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1)P(A_2 \mid A_3)P(A_3 \mid A_1 \cap A_2)...P\left(A_n \mid \bigcup_{i=1}^{n-1} A_i\right)$$

3.4 Total Probability Theorem and Bayes' Rule

Total Probability Theorem

Let $A_1, A_3, ..., A_n$ be disjoint events that *form a partition* of the sample space, and assume that $P(A_i) > 0$ for all i. Then far any event B, we have

$$P(B) = P(A_n \cup B) + ... + P(A_n \cup B)$$
$$= \sum_{i=1}^{n} P(A_i)P(B \mid A_i)$$

Bayes' Rule

Let $A_1, A_2, ..., A_n$ be disjoint events that form a partition of the sample space, and assume that $P(A_i) > 0$ for all i. Then, for any event B such that B that $P(B) > 0$, we have

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)}$$
$$= \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + ... + P(A_n)P(B \mid A_n)}$$

The Bayes' Rule reveals the relation between conditional probability of form $P(A \mid B)$ and $P(B \mid A)$, in which the order of conditioning is reversed.

e.g. An example in medicine

If there is a shade in someone's x-ray, and there are 3 possibilities:

1. It's a malignant tumor
2. It's a nonmalignant tumor
3. Not a tumor

Calculate the probability of each situation.

Ans Let A_1, A_2, A_3 be the three events, and B be the probability of there being a shade. Assume that we know the probabilities $P(A_i)$ and $P(B \mid A_i)$ (these data can be actually found in practise). So we due to Bayes' Rule, we have

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(A_1 \mid B) + P(A_2)P(A_2 \mid B) + P(A_3)P(A_3 \mid B)}$$

As above, the Bayes Rule is often used for inference where we need to infer the "causes" from "effects". The events A_1, A_2, A_3 are the causes and the shade event B is the effect by the causes. In a lot of situations, we have already collected the data of the effects, and we want to evaulate the probability of the cause A_i is present, that's when Bayes' Rule come into use.

And just like the example above, the A_i stands for the cause, the B stands for the effect, we give the definition of *posterior probability* and *prior probability*:

Posterior probability $P(A_i \mid B)$

Prior probability $P(A_i)$

e.g. The False-Positive Puzzle

A test for a certian rare disease is assumed to be 95% correct, and a random person drawn from a cortain population has the probability 0.001 of having the disease. Then, if a person tested positive, what is the probability of the person having the disease?

Ans Now we know the effect, we want to evaulate the probability of the cause is present – Apply the Bayes' Rule! Let A be the event the person have the disease, B be the event of tested positive. So we want $P(A \mid B)$, and we have $P(A) = 0.001, P(B \mid A) = 0.95$. So

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)} = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)} = 0.0187$$

Less than 2%!!!

3.5 Independence

Definition$_{\textit{independence}}$: If $P(A \cap B) = P(A)P(B)$, then we say that A is *independent* of B.

The equation above is also equivalent to $P(A \mid B) = P(A)$.

3.6 Conditional Independence

Definition$_{\textit{conditional independence}}$: If $P(A \cap B \mid C) = P(A \mid C)P(B \mid C)$, we say thar events A and B are conditionally independent (given C).

Summary

- Two events A, B are independent if $P(A \cap B) = P(A)P(B)$. In addition, if $P(B) > 0$, independence is equivalent to $P(A \mid B) = P(A)$
- If A, B are independent, so are A and B^c
- Two events A, B are conditionally independent if $P(A \cap B \mid C) = P(A \mid C)P(B \mid C)$. In addition, if $P(B) > 0$, independence is equivalent to $P(A \mid B \cap C) = P(A \mid C)$.
- Independence does not imply conditional independence and vice versa.

Independence of a collection of events

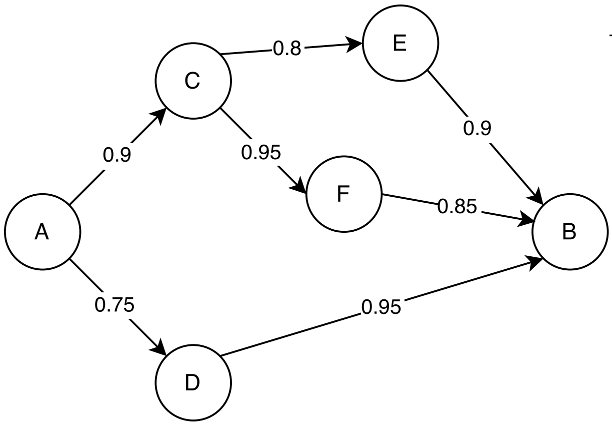
Events $A_1, A_2, ..., A_n$ are independent if

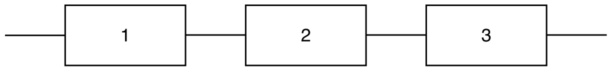
$$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i), \text{for every subset } S \text{ of } \{1, 2, ..., n\}$$

3.7 Reliability

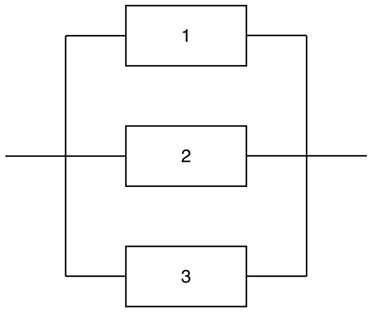
e.g. Network connectivity

The following graph is a network with the prabability of link of the connected nodes is up. Evaulate the probability of the successful connection from A to B.(Classic problem ,omitting ans)





Series Connection



Parallel Connection

3.8 Independent Trials and the Binomial Probabilities