Probability and Random Process

Aimin Tang

The University of Michigan- Shanghai Jiao Tong University Joint Institute

Shanghai Jiao Tong University

Sep. 15 2020

Based on Lecture Notes by Prof. Yong Long

- 1. Introduction to Probability
  - Application example
  - Review of set and functions
  - Models of random experiments
  - Axioms and properties of probability
  - Conditional probability
  - Independence of events
  - Combinatorics and probability

Conditional probability

- The conditional probability $P(A|B)$ of event A occurring given that event B has occurred is a model for the conditional frequency with which A occurs when B has also occurred.

- Example: dice toss

  - $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = 2^{\Omega}$, $P(A) = \dfrac{|A|}{6}$, X

  - $A = \{3\}$, $B = \{1, 3, 5\}$

  - Find $P(A|B)$

x1	x2	x3	x4	x5	x6	x7	x8	x9	X10
3	2	6	4	2	5	4	1	1	3
A									A
B					B		B	B	B

- Frequency of A is 2/10
- Frequency of B is 5/10
- Frequency of A given B is 2/5

- $P(A|B) = 2/5$?

Conditional probability: relationship

- Relationship to "unconditional" probability

$$P(A|B) = \begin{cases} \dfrac{P(A \cap B)}{P(B)}, P(B) \neq 0 \\ \text{undefined}, P(B) = 0 \end{cases}$$

- usually apply

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

  – $P(A \cap B) =$ **joint probability** of A and B

  – $P(A/B) =$ conditional probability of A given B

- Previous examples:

  - $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{1/6}{3/6} = \dfrac{1}{3}$

- There are 200 light bulbs in one box. Each light bulb is good or defective. Three are three different powers: 50w(watt), 100w, or 150w

	50W	100W	150W
G	85	63	27
D	15	7	3

- Experiment: draw a bulb at random from the box

$$\Pr(D|50W) = ? \qquad \Pr(50W|D) = ?$$

$$\mathrm{P_r}(D|50w) = \frac{\mathrm{P_r}(\textit{defective of } 50w)}{\mathrm{P_r}(50w)} = \frac{15/200}{100/200} = \frac{15}{100} = \frac{3}{20}$$

$$\mathrm{P_r}(50w|D) = \frac{\mathrm{P_r}(\textit{defective of } 50w)}{\mathrm{P_r}(\textit{defective})} = \frac{15/200}{25/200} = \frac{15}{25} = \frac{3}{5}$$

- Conditional probability behaves like ordinary probability
  - It satisfies all the axioms and all the consequences of all axioms

$$P(A \cup B|C) = P(A|C) + P(B|C)$$

  - if A and B are disjoint
  - $P(A/B)$ is undefined if $(B) = 0$

- Question: which of the following is true
  - $P(A/B) \geq P(A)$
  - $P(A/B) \leq P(A)$

  - **Both wrong**

- The law of total probability gives an expression for computing the probability of an event that can occur in different ways

- Let $\{B_1, B_2, \ldots, B_n\}$ be a partition of Ω, and let $B_1, B_2, \ldots, B_n \in \mathcal{A}, \mathrm{P}(B_k) > 0, k = 1, 2, \ldots, n$, then we can write

$$\mathrm{P}(A) = \sum_{k=1}^{n} \mathrm{P}(A \cap B_k) = \sum_{k=1}^{n} \mathrm{P}(A|B_k)\,\mathrm{P}(B_k)$$

- This expression is the law of total probability.

$$A = A \cap \Omega = A \cap \left(\bigcup_{k=1}^{n} B_k \right)$$

$$= \bigcup_{k=1}^{n} (A \cap B_k) \quad (generalized distributive law)$$

Observe $(A \cap B_i)$ and $(A \cap B_j)$ are pairwise disjoint for $i \neq j$

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset$$

$(\{B_1, B_2, \ldots, B_n\}$ is a partition of $\Omega)$

$$P(A) = P\left(\bigcup_{k=1}^{n} (A \cap B_k) \right) = \sum_{k=1}^{n} P(A \cap B_k) \quad (Kolmogorov's\ axioms\ 3)$$

Example

- 80% of students who do HW pass the exam
- 10% of students who do not do HW pass the exam
- 60% of students who do HW
- suppose the student is randomly chosen. What is the probability that the student pass the exam?

$$P_r(HW) = 0.6, \quad P_r(P|HW) = 0.8, \quad P_r(P|no\ HW) = 0.1$$

$$
\begin{aligned}
P_r(P) &= P_r((P \cap HW)\ or\ (P \cap no\ HW)) \\
&= P_r(P \cap HW) + P_r(P \cap no\ HW) \\
&= P_r(HW)\,P_r(P|HW) + P_r(no\ HW)\,P_r(P|no\ HW) \\
&= 0.6 \cdot 0.8 + (1 - 0.6) \cdot 0.1 = 0.52
\end{aligned}
$$

- Bayes' rule provides an expression for the conditional probability of an event A given B

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|A^c)\,P(A^c)}$$

Proof. Combine the expression for the *conditional probability*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B)}$$

with *the law of total probability*

$$P(B) = P(B|A)\,P(A) + P(B|A^c)\,P(A^c)$$

- Let $A_1, A_2, \ldots, A_n$ be a set of **pairwise disjoint** events, $A_k \in \mathcal{A}, k = 1, 2, \ldots, n$, such that

$$\sum_{k=1}^{n} P(A_k) = 1$$

- Then for any event B

$$P(B) = \sum_{k=1}^{n} P(B|A_k)\, P(A_k)$$

- Do $A_1, A_2, \ldots, A_n$ form a partition of Ω?

    - Since we do not require $\bigcup_{k=1}^n A_k = \Omega$, then A_k do not, strictly speaking, for a partition. The reminder set $A^c = (\bigcup_{k=1}^n A_k)^c$ has probability zero

Ve501 2020-2021 Fall

- Applying the general form of the law of total probability yields the general form of Bayes' rule

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}$$

  – The probabilities $P(A_k), k = 1, 2, \ldots, n$ are called prior probabilities of events $A_k, k = 1, 2, \ldots, n$

  – The probabilities $P(A_k|B), k = 1, 2, \ldots, n$ are called posterior probabilities of events $A_k, k = 1, 2, \ldots, n$ given B

- It is useful when there is a cause & effect relationship.
  - Some event happens that we do not observe, but for which we have some statistical information- prior probabilities $P(A_k)$
  - Based on the event B we observed and the conditional probability $P(B|A_k)$ we know , we would like to find the posterior probabilities $P(A_k|B)$ of event A_k conditioned on the observation of B

Example

- Blood test for disease *X*
    - test is 98% accurate for people who have *X*
    - test is 97% accurate for people who do not have *X*
    - P = test is positive, NP = test is negative
    - 1% of people have disease *X*
- Your test result is P, what's the probability you have disease *X*?

- D=have disease X, ND= do no have disease X

- Solve for $P(D|P)$

- We have $P(D) = 0.01, P(P|D) = 0.98, P(NP|ND) = 0.97$

$$P(D|P) = \frac{P(P|D)P(D)}{P(P)} = \frac{P(P|D)P(D)}{P(P|D)P(D) + P(P|ND)P(ND)}$$

$$= \frac{0.98 * 0.01}{0.98 * 0.01 + (1 - 0.97)0.99} = 0.248$$

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

$P(A_1 \cap A_2 \ldots \cap A_n) = P(A_1)P(B|A_1)$, where $B \triangleq A_2 \cap \cdots \cap A_n$

Example

- In a factory there are 100 units of a certain product, 5 of which are defective. We pick three units from the 100 units at random. What is the probability that none of them are defective?

- Solution:
  - Let us define A_i as the event that the i th chosen unit is not defective for i = 1, 2, 3. We are interested in $P(A_1 \cap A_2 \cap A_3)$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \, P(A_2|A_1) \, P(A_3|A_2 \cap A_1)$$

$$= \frac{95}{100} \frac{94}{99} \frac{93}{98}$$

$$\boxed{= 0.8560}$$

Independence of events

Independence derivation

- If events A and B satisfy

$$P(A|B) \ = \ P(A|B^c) \qquad\qquad (1)$$

then we say that **_A does not depend on B_**

- We can rewrite (1) as

$$\frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B^c)}{P(B^c)}$$

- Using

$$P(B^c) = 1 - P(B)$$
$$P(A) = P(A \cap B) + P(A \cap B^c)$$

- We get

$$\frac{P(A \cap B)}{P(B)} = \frac{P(A) - P(A \cap B)}{1 - P(B)}$$

$$P(A \cap B)\big(1 - P(B)\big) = P(B)(P(A) - P(A \cap B))$$

$$P(A \cap B) = P(A)P(B)$$

- The above sequence of calculation is reversible

$$P(A \cap B) = P(A)P(B) \Rightarrow P(A|B) = P(A|B^c)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A) - P(A \cap B)}{1 - P(B)} = \frac{P(A) - P(A)P(B)}{1 - P(B)} = P(A)$$

$$\Rightarrow P(A|B) = P(A|B^c)$$

- Interchanging the role of A and B in the above equations, we get that if B does not depend on A then
$$P(A \cap B) = P(A)P(B)$$

- Since sequence of calculation is reversible and $P(A \cap B) = P(A)P(B)$ is symmetric in A and B it follows that

- *A* does not depend on *B* iff *B* does not depend on *A*

- Events A and B are said to be statistically independent or simply independent if their joint probability equals the product of their probabilities

$$P(A \cap B) = P(A)P(B)$$

- Equivalently:

  - $P(A) = 0$, or $P(A) \neq 0$ & $P(B/A) = P(B)$

  - $P(B) = 0$, or $P(B) \neq 0$ & $P(A/B) = P(A)$

- Interpretation: A and B independent means that the occurrence of one tells nothing about the occurrence of the other.

- Show that events A and B are independent if $P(B) = 0$

- Solution: Suppose that A and B are any two events. If P(B) = 0, then we claim that A and B are independent. We must show that $P(A \cap B) = P(A)P(B) = 0$.

- To show that the left-hand side is zero, observe that since probabilities are nonnegative, and since $A \cap B \subset B$,
$$0 \leq P(A \cap B) \leq P(B) = 0$$

- If the events A and B are independent, are
  - A and B^c
  - A^c and B
  - A^c and B^c

  independent?

We prove A and B^c are independent.

$$P(A) = P(A \cap B^c) + P(A \cap B)$$

$$\implies P(A \cap B^c) = P(A) - P(A \cap B)$$

$$\implies P(A \cap B^c) = P(A) - P(A)P(B) \quad (A, B \text{ independent})$$

$$\implies P(A)(1 - P(B)) = P(A)P(B^c)$$

Similarly, one can easily prove A^c and B are independent and A^c and B^c are independent.

Ve501 2020-2021 Fall

- toss 3 coins

$$\Omega = \{HHH, HHT, HTH, \ldots, TTT\},$$
$$\mathcal{A} = 2^{\Omega}, P(A) = |A|/8$$

  – A = even numbers of H's and B = first toss of H
  – Are A and B independent?

$$A = \{HHT, HTH, THH, TTT\}, \quad B = \{HHH, HHT, HTH, HTT\}$$

$$P(A) = \frac{4}{8}, P(B) = \frac{4}{8}$$

$$A \cap B = \{HHT, HTH\}, \quad \Rightarrow P(A \cap B) = \frac{2}{8}$$

$$so \ P(A) P(B) = P(A \cap B)$$

- Events $A_1, A_2, \ldots, A_n$ are mutual independent if the following relation holds

$$P\left(A_{k_1} \cap A_{k_2} \ldots \cap A_{k_l}\right) = \prod_{i=1}^{l} P\left(A_{k_i}\right)$$

for any subset $\{k_1, k_2, \ldots, k_l\}$ of $\{1, 2, \ldots, \mathrm{n}\}$.

  – The above definition implies that any set of two events, three events, and so on, must satisfies this equation.

Important remarks

- Pairwise independence does not imply mutual independence

- Example: Let $\Omega = \{1,2,3,4,5,6,7\}, \mathcal{A} = 2^\Omega$

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = \frac{1}{8}, P(\{7\}) = \frac{1}{4}$$

Consider

$$A = \{1,2,7\}, B = \{3,4,7\}, C = \{5,6,7\}$$

- Note that $A, B, C \in \mathcal{A}$ and
$$P(A) = P(B) = P(C) = 1/2$$

- Note that

$$P(A \cap B) = P(\{7\}) = \frac{1}{4}$$

$$P(A \cap C) = P(\{7\}) = \frac{1}{4}$$

$$P(B \cap C) = P(\{7\}) = \frac{1}{4}$$

- Then

$$\frac{1}{4} = P(A \cap B) = P(A)P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

$$\frac{1}{4} = P(A \cap C) = P(A)P(C) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

$$\frac{1}{4} = P(B \cap C) = P(B)P(C) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

So, A,B,C are pairwise independent.

- But $P(A \cap B \cap C) = P(\{7\}) = \frac{1}{4}, P(A)P(B)P(C) = \frac{1}{8}$

- So $P(A \cap B \cap C) \neq P(A)P(B)P(C)$, A, B, C are not mutually independent

- Three events A, B, C such that $P(A \cap B \cap C) = P(A)P(B)P(C)$ does not imply A, B, C are pairwise independent.

- Example: let $\Omega = \{1,2,3,4\}, \mathcal{A} = 2^{\Omega}$ and

$$P(\{1\}) = P(\{2\}) = P(\{3\}) = p = \frac{3 - \sqrt{3}}{4}$$

$$P(\{4\}) = q = \frac{-5 + 3\sqrt{3}}{4}$$

Let $A = \{1,4\}, B = \{2,4\}, C = \{3,4\}$

- Note that as required

$$3p + q = \frac{9 - 3\sqrt{3} - 5 + 3\sqrt{3}}{4} = 1$$

- further

$$(p + q)^3 = q, (p + q)^2 \neq q$$

- Then

$$A \cap B \cap C = \{4\}, \quad A \cap B = B \cap C = A \cap C = \{4\}$$

$$P(A \cap B \cap C) = P(\{4\}) = q$$

$$P(A) = P(B) = P(C) = p + q$$

$$P(A) P(B) P(C) = (p + q)^3 = q$$

$$P(A \cap B \cap C) = P(A) P(B) P(C)$$

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = q$$

$$P(A) P(B) = P(B) P(C) = P(A) P(C) = (p + q)^2 \neq q$$

So A, B, C are not pairwise independent.

Important remarks: disjoint and independent

- Difference between disjoint events and independent events

- Recall that events A and B are disjoint if $A \cap B = \emptyset$

  – This concept does not involve the probability measure P in any way

  – To determine if A and B are disjoint only knowledge of A and B is required

  – The notation of independence of events A and B does involve the probability measure P

    - Note that $P(A \cap B) = P(A)P(B)$ does depend on P and not just on A and B

      若 $P(A)\,P(B) \neq 0$.

      disjoint � 不 为 independent

Thank You!