Lecture 5 Conditional Probability

BIO210 Biostatistics

Xi Chen

Spring, 2022

School of Life Sciences
Southern University of Science and Technology

南方科技大学生命科学学院
SUSTech · SCHOOL OF
LIFE SCIENCES

Probability axioms

1. **Nonnegativity:** $P(A) \geqslant 0$
2. **Normalisation:** $P(\boldsymbol{\Omega}) = 1$
3. **Additivity:** if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$

Extended version of Axiom 3 - **countable additivity axiom:**

If a sequence of events A_1, A_2, A_3, $\cdots$ are disjoint, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

Conditional probability

Team a Team b

vs

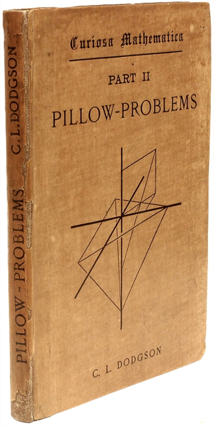
vs

vs

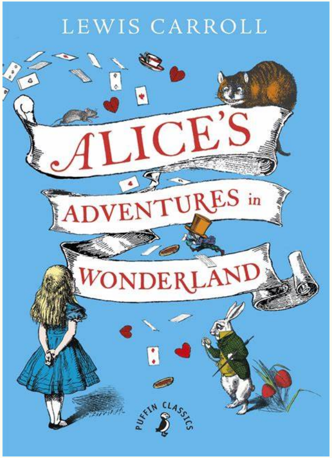
$A = \{$ Team a wins $\}$
$P(A) = ?$

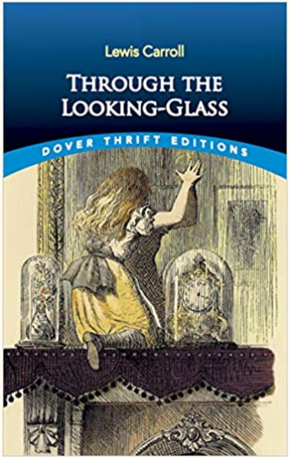
$B = \{$ key players from Team a injured $\}$
Konwing B, $P(A) = ?$

Lewis Carroll's Pillow Problems



by Charles Dodgson





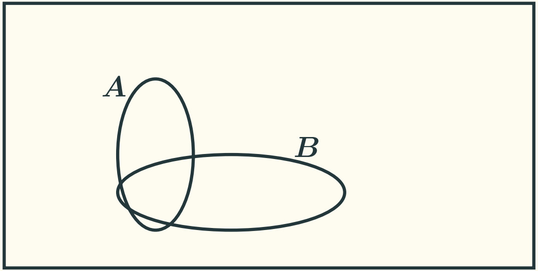
Question #5 (8th Sep 1887):

A bag contains one counter, known to be either white or black. A white counter is put in, the bag shaken, and a counter drawn out, which proves to be white. What is now the chance of drawing a white counter?

Conditional probabilities are just like normal probabilities in a different universe or sample space.

Conditional probability

Sample space (Ω)



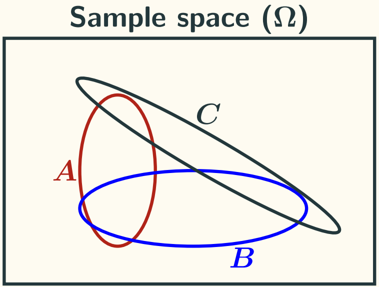
$P(A|B) =$ the probability of A, given that B has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, P(B) \neq 0$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, P(A) \neq 0$$

$$\begin{aligned} P(A \cap B) &= P(B)P(A|B) \\ &= P(A)P(B|A) \end{aligned}$$

Conditional version of the axioms



Conditional version of probability axioms:

1. **Conditional nonnegativity:** $P(A|C) \geqslant 0$
2. **Conditional normalisation:** $P(\Omega|C) = 1$
3. **Conditional additivity:**

$$\text{If } A \cap B \,|\, C = \emptyset, \text{ then}$$

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

Example 1

Experiment: flipping a fair coin three times.

Sample space Ω: { HHH, HHT, HTH, THH, HTT, THT, TTH, TTT }

$\boldsymbol{A}$ = { the first flip is H }
$\boldsymbol{B}$ = { the second flip is T }
$\boldsymbol{C}$ = { exactly two Hs }

$P(B|A) = ?$
$P(C|B) = ?$

Example 2

Results of a clinical depression study

	Treatment group			
Responses	Imipramine	Lithium	Combination	Placebo
Relapse	180	130	220	240
No relapse	220	250	160	100

$\boldsymbol{A} = \{$ the patient had a relapse $\}$
$\boldsymbol{B} = \{$ the patient received the placebo $\}$
$P(A) = ?$
$P(A|B) = ?$

Example 3

Experiment: using a virus detection kit on a person

- It is known that there are 0.5% people in general that carry the virus $\boldsymbol{V}$. A company has a kit to detect this specific virus. It is known that if a person does carry the virus, there are 99% of the chance that the kit will show a positive result; if a person does not carry the virus, there are 98% of the chance that the kit will show a negative result.

$\boldsymbol{A} = \{$ the person carries the virus $\}$

$\boldsymbol{B} = \{$ the test result is positive $\}$

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

$$P(B) = ?$$

$$P(A|B) = ?$$

The False Positive Puzzle

"Getting the goat" in the *Finance and economics* section from the Feb 20th 1999 edition of *The Economist* magazine

- **The false-positive puzzle.** You are given the following information. (a) In random testing, you test positive for a disease. (b) In 5% of cases, this test shows positive even when the subject does not have the disease. (c) In the population at large, one person in 1,000 has the disease. What is the probability that you have the disease?

Nearly everyone replies: 95%. This is not quite right. The answer is 2%. To see why, consider a population of 1,000 people. Of these, on average, one will have the disease, but 50 others will also test positive. Of those who test positive, therefore, only one in 51, about 2%, will turn out to have the disease. The key is to see that the information in (c) is crucial. Most people think it irrelevant: the test is "95% reliable", and that's that. Try this one on doctors. It deflates their egos wonderfully: they do hardly any better than laymen. In a study carried out in the 1970s, 80% of those questioned at a leading American hospital gave the wrong answer, most of them saying 95%.