Basic Probability

- Probability theory enables us to make rational decisions.
- Which mode of transportation is safer:
  - Car or Plane?
  - What is the probability of an accident?

Basic Probability Theory

- An **experiment** has a set of potential outcomes, e.g., throw a dice
- The **sample space** of an experiment is the set of all possible outcomes, e.g., {1, 2, 3, 4, 5, 6}
- An **event** is a subset of the sample space.
  - {2}
  - {3, 6}
  - even = {2, 4, 6}
  - odd = {1, 3, 5}

Probability as Relative Frequency

- An event has a probability.
- Consider a long sequence of experiments. If we look at the number of times a particular event occurs in that sequence, and compare it to the total number of experiments, we can compute a ratio.
- This ratio is one way of estimating the probability of the event.
- P(E) = (# of times E occurred)/(total # of trials)

- Example
  - 100 attempts are made to swim a length in 30 secs. The swimmer succeeds on 20 occasions therefore the probability that a swimmer can complete the length in 30 secs is:
    - 20/100 = 0.2
    - Failure = 1-.2 or 0.8
- The experiments, the sample space and the events must be defined clearly for probability to be meaningful
  - What is the probability of an accident?

Theoretical Probability

- **Principle of Indifference**— Alternatives are always to be judged equiprobable if we have no reason to expect or prefer one over the other.
- Each outcome in the sample space is assigned equal probability.
- Example: throw a dice
  - P({1})=P({2})= … =P({6})=1/6

Axioms of Probability Theory

- Suppose P(.) is a probability function, then
  1. for any event E, $0 \leq P(E) \leq 1$.
  2. $P(S) = 1$, where S is the sample space.
  3. for any two mutually exclusive events E1 and E2,

$$P(E1 \cup E2) = P(E1) + P(E2)$$

- Any function that satisfies the above three axioms is a probability function.

Joint Probability

- Let A, B be two events, the joint probability of both A and B being true is denoted by P(A, B).
- Example:

P(spade) is the probability of the top card being a spade.

P(king) is the probability of the top card being a king.

P(spade, king) is the probability of the top card being both a spade and a king, i.e., the king of spade.

P(king, spade)=P(spade, king)   ???

Properties of Probability

1. $P(\neg E) = 1 - P(E)$
2. If E1 and E2 are logically equivalent, then

   $P(E1)=P(E2)$.

   - E1: Not all philosophers are more than six feet tall.
   - E2: Some philosopher is not more that six feet tall.

   Then $P(E1)=P(E2)$.
3. $P(E1, E2) \leq P(E1)$.

Conditional Probability

- The probability of an event may change after knowing another event.

  The probability of A given B is denoted by P(A|B).

- Example
  - P(W=space) the probability of a randomly selected word from an English text is ‘space’
  - P(W=space | W’=outer) the probability of ‘space’ if the previous word is ‘outer’

Example

A: the top card of a deck of poker cards is a king of spade

$P(A) = 1/52$

However, if we know

B: the top card is a king

then, the probability of A given B is true is

$P(A|B) = 1/4$.

Independence: Intuition

- Events are independent if one has nothing whatever to do with others. Therefore, for two independent events, knowing one happening does change the probability of the other event happening.
  - one toss of coin is independent of another coin (assuming it is a regular coin).
  - price of tea in England is independent of the result of general election in Canada.

Independent or Dependent?

- Getting cold and getting cat-allergy
- Mile Per Gallon and acceleration.
- Size of a person's vocabulary the person's shoe size.

Independence: Definition

- Events A and B are independent iff:

$$P(A, B) = P(A) \times P(B)$$

which is equivalent to

$$P(A|B) = P(A) \text{ and}$$
$$P(B|A) = P(B)$$

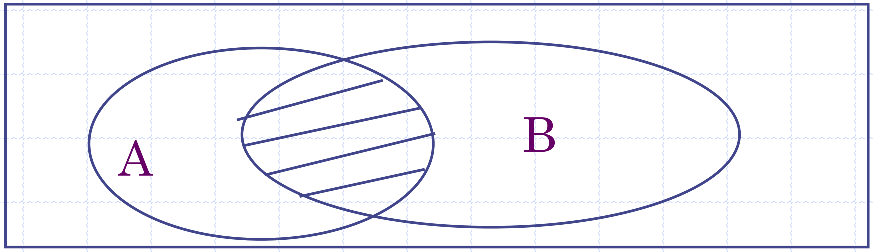
when $P(A, B) > 0$.

T1: the first toss is a head.

T2: the second toss is a tail.

$$P(T2|T1) = P(T2)$$

How to Compute P(A|B)?

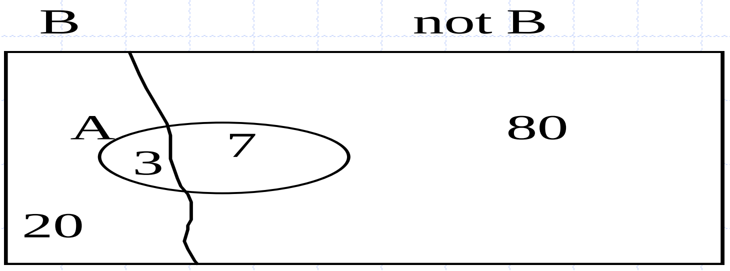


$$P(A \mid B) = \frac{N(A \text{ and } B)}{N(B)} = \frac{\frac{N(A \text{ and } B)}{N}}{\frac{N(B)}{N}} = \frac{P(A, B)}{P(B)}$$

$$P(\text{brown} \mid \text{cow}) = \frac{N(\text{brown-cows})}{N(\text{cows})} = \frac{P(\text{brown-cow})}{P(\text{cow})}$$

Business Students

Of 100 students completing a course, 20 were business major. Ten students received As in the course, and three of these were business majors., suppose A is the event that a randomly selected student got an A in the course, B is the event that a randomly selected event is a business major. What is the probability of A? What is the probability of A after knowing B is true?



Bayes Theorem

- If $P(E2)>0$, then $P(E1|E2)=P(E2|E1)P(E1)/P(E2)$
- This can be derived from the definition of conditional probability.

Example

A patient takes a lab test and the result comes back positive. The test has a false negative rate of 2% and false positive rate of 3%. Furthermore, 0.8% of the entire population have this cancer.

What is the probability of cancer if we know the test result is positive?

Probabilistic Reasoning

- Evidence
  - What we know about a situation.
- Hypothesis
  - What we want to conclude.
- Compute
  - P(Hypothesis | Evidence)

Credit Card Authorization

- E is the data about the applicant's age, job, education, income, credit history, etc,
- H is the hypothesis that the credit card will provide positive return.
- The decision of whether to issue the credit card to the applicant is based on the probability $P(H|E)$.

Application: Spam Detection

- Spam
  - Dear sir, We want to transfer to overseas ($ 126,000.000.00 USD) One hundred and Twenty six million United States Dollars) from a Bank in Africa, I want to ask you to quietly look for a reliable and honest person who will be capable and fit to provide either an existing ......
- Legitimate email
  - Ham: for lack of better name.

- Hypotheses: {Spam, Ham}
- Evidence: a document
  - The document is treated as a set (or bag) of words
- Knowledge
  - P(Spam)
    - The prior probability of an e-mail message being a spam.
    - How to estimate this probability?
  - P(w|Spam)
    - the probability that a word is w if we know w is chosen from a spam.
    - How to estimate this probability?