CS1382 Discrete Computational Structures

Introduction to Discrete Probability

Spring 2019

Richard Matovu

TEXAS TECH UNIVERSITY

Probability of an Event

Pierre-Simon Laplace (1749-1827)

- Key terms:
  - **Experiment**: Procedure that yields one of a given set of possible outcomes.
  - **Sample space** of the experiment is the set of possible outcomes.
  - **Event** is a subset of the sample space.

- **Laplace Definition**:

  If *S* is a finite sample space of equally likely outcomes and *E* is an event, that is, a subset of *S*, then the ***Probability* of *E*** is **P(E) = |E|/|S|**.

- For every event E, we have $0 \leq P(E) \leq 1$
  - $0 \leq |E| \leq |S|$

    $0 \leq |E|/|S| \leq |S|/|S|$

    $0 \leq P(E) \leq 1$

Applying Laplace's Definition

- An container has four blue balls and five red balls. What is the probability that a ball chosen from the container is blue?

  - **Solution**:
    S = 4 + 5 = 9 balls , E = 4 balls
    P(E) = 4 / 9

- Jane tosses two fair coins. What is the probability of getting exactly one head?

  - **Solution:**
    S = { HH, HT, TH, TT }
    E = { HT, TH }
    P(E) = |E|/|S| = 2 / 4 = 1 / 2

Applying Laplace's Definition

- What is the probability that when two dice are rolled, the sum of the numbers on the two dice is 7?
  - **Solution**: 36 possible outcomes and six of these outcomes sum to 7.

    Hence, the probability of obtaining a sum of 7 is 6/36 = 1/6.
- In a lottery, a player wins a large prize when they pick four digits that match, in correct order, four digits selected by a random mechanical process. What is the probability that a player wins the prize?
  - **Solution**:

    10^4 = 10,000 ways to pick four digits.

    Only 1 way to pick the correct digits

    Therefore the probability of winning the large prize is 1/10,000 = 0.0001.

Applying Laplace's Definition

- There are many lotteries that award prizes to people who correctly choose a set of six numbers out of the first *n* positive integers, where *n* is usually between 30 and 60. What is the probability that a person picks the correct six numbers out of 40?

    - **Solution**:

      The number of ways to choose six numbers out of 40 is

      $C(40,6) = 40!/(34!6!) = 3,838,380$.

      Hence, the probability of picking a winning combination is $1/ 3,838,380 \approx 0.00000026$.

Applying Laplace's Definition

- What is the probability that the numbers 11, 4, 17, 39, and 23 are drawn in that order from a bin with 50 balls labeled with the numbers 1,2, …, 50 if

  a) The ball selected is not returned to the bin.

  b) The ball selected is returned to the bin before the next ball is selected.

  - **Solution**:
    Use the product rule in each case.

    a) Sampling without replacement:
    The probability is 1/254,251,200
    Since there are $50 \cdot 49 \cdot 47 \cdot 46 \cdot 45 = 254{,}251{,}200$ ways to choose the five balls.

    b) Sampling with replacement:
    The probability is $1/50^5 = 1/312{,}500{,}000$ since $50^5 = 312{,}500{,}000$.

The Probability of Complements

- **Theorem 1**: Let E be an event in sample space S.

  The probability of the event $\overline{E}$ = S − E, the complementary event of E, is given by

$$p(\overline{E}) = 1 - p(E).$$

- **Proof**: Using the fact that $|\overline{E}|$ = |S| − |E|,

$$p(\overline{E}) = \frac{|S|-|E|}{|S|} = 1 - \frac{|E|}{|S|} = 1 - p(E).$$

The Probability of Complements

- What is the probability of NOT getting two heads?

  - **Solution:**

    P(2 Heads) = | { HH } | / | S | = 1 / 4

    P(Not 2 Heads) = 1 – 1 / 4 = 3 / 4

- A sequence of 10 bits is chosen randomly. What is the probability that at least one of these bits is 0?

  - **Solution**:

    Let *E* be the event that at least one of the 10 bits is 0.

    Then E’ is the event that all of the bits are 1s.

    The size of the sample space *S* is 2^{10}. Hence,

$$p(E) = 1 - p(\overline{E}) = 1 - \frac{|\overline{E}|}{|S|} = 1 - \frac{1}{2^{10}} = 1 - \frac{1}{1024} = \frac{1023}{1024}.$$

The Probability of Unions of Events

- **Theorem 2**:

  Let E_1 and E_2 be events in the sample space S. Then

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

- **Proof**: Given the inclusion-exclusion formula, $|A \cup B| = |A| + |B| - |A \cap B|$, it follows that

$$p(E_1 \cup E_2) = \frac{|E_1 \cup E_2|}{|S|} = \frac{|E_1| + |E_2| - |E_1 \cap E_2|}{|S|}$$

$$= \frac{|E_1|}{|S|} + \frac{|E_2|}{|S|} - \frac{|E_1 \cap E_2|}{|S|}$$

$$= p(E_1) + p(E_2) - p(E_1 \cap E_2).$$

The Probability of Unions of Events

- What is the probability that a positive integer selected at random from the set of positive integers not exceeding 100 is divisible by either 2 or 5?

    - **Solution**:

        - E_1 be the event that the integer is divisible by 2 and E_2 be the event that it is divisible 5
        - $E_1 \cup E_2$ is the event that the integer is divisible by 2 or 5 is
        - $E_1 \cap E_2$ is the event that it is divisible by 2 and 5.

        It follows that:

        $$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

        $$= 50/100 + 20/100 - 10/100 = 3/5.$$

Exercise

- Given a deck of 52 cards, we draw 5. What is the probability of getting 3 Aces and 2 Jacks?

- I toss a coin and then roll a 6-sided die. Suppose we have the following events:

  A: A head appears , B: A 3 appears . What is

  - P(A)

  - P(B)

  - P(B’)

  - P(A n B)

  - P(A U B)

CS1382 Discrete Computational Structures

Probability Theory

Spring 2019

Richard Matovu

TEXAS TECH UNIVERSITY

Uniform Distribution

- Suppose that S is a set with n elements.
- The **uniform distribution** assigns the probability 1/n to each element of S.

  (Note that we could have used Laplace's definition)

  - **Example**:

    Consider again the coin flipping example, but with a fair coin.

    Now P(H) = P(T) = 1/2.

Assigning Probabilities

Laplace's definition: Assumes that all outcomes are equally likely.

More general definition of probabilities:

- Let S be a sample space of an experiment with a finite number of outcomes.

  We assign a probability $p(s)$ to each outcome s, so that:

  i. $0 \le p(s) \le 1$ for each $s \in S$

  ii. $\sum_{s\in S} p(s) = 1$

- The function p from the set of all outcomes of the sample space S is called a **Probability Distribution**.

Assigning Probabilities

- What probabilities should we assign to the outcomes H (heads) and T (tails) when a fair coin is flipped?
- What probabilities should be assigned to these outcomes when the coin is biased so that heads comes up twice as often as tails?
  - **Solution**:
    - For a fair coin, we have $P(H) = P(T) = ½$.
    - For a biased coin, we have $P(H) = 2P(T)$

      Because $P(H) + P(T) = 1$, it follows that

      $2P(T) + P(T) = 3P(T) = 1$.

      Hence, $P(T) = 1/3$ and $P(H) = 2/3$.

Example

- Suppose that a die is biased so that 3 appears twice as often as each other number, but that the other five outcomes are equally likely. What is the probability that an odd number appears when we roll this die?

- **Solution**:

  - We want the probability of the event E = {1,3,5}.

    We have P(3) = 2/7 and P(1) = P(2) = P(4) = P(5) = P(6) = 1/7.

  - Hence, P(E) = P(1) + P(3) + P(5)

    = 1/7 + 2/7 + 1/7 = 4/7.

Probabilities of Complements and Unions of Events

- **Complements:**

$$p(\overline{E}) = 1 - p(E)$$

still holds. Since each outcome is in either E or $\overline{E}$, but not both,

$$\sum_{s \in S} p(s) = 1 = p(E) + p(\overline{E}).$$

- **Unions:**

$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$ also still holds under the new definition.

Combinations of Events

Theorem:

If E_1, E_2, … is a sequence of pair-wise disjoint events in a sample space S, then

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

Conditional Probability

Let E and F be events with P(F) > 0. The conditional probability of E given F, denoted by P(E|F), is defined as:

$$p(E|F) = \frac{p(E \cap F)}{p(F)}$$

Example:

- A fair coin is tossed three times, what is the probability of at least two heads, given that the first toss was a head?

  Let E be the event that you get two heads and F be the event that the first toss was a head

  - $P(E \cap F) = 3 / 8$ and $P(F) = 4 / 8 = 1 / 2$
  - $P(E \mid F) = P(E \cap F) / P(F) = 3/8 / 1/2 = 3 / 4$

Conditional Probability

- A bit string of length four is generated at random so that each of the 16 bit strings of length 4 is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0?

  - **Solution**:

    E be the event that the bit string contains at least two consecutive 0s

    F be the event that the first bit is a 0.

    - Since E ⋂ F = {0000, 0001, 0010, 0011, 0100}, P(E ⋂ F) = 5/16.

    - Because 8 bit strings of length 4 start with a 0, P(F) = 8/16= ½.

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{5/16}{1/2} = \frac{5}{8}.$$

Conditional Probability

- What is the conditional probability that a family with two children has two boys, given that they have at least one boy. Assume that each of the possibilities BB, BG, GB, and GG is equally likely where B represents a boy and G represents a girl.

- **Solution**: Let E be the event that the family has two boys and let F be the event that the family has at least one boy. Then E = {BB}, F = {BB, BG, GB}, and E ⋂ F = {BB}.
    - It follows that P(F) = 3/4 and P(E ⋂ F) = 1/4.

    Hence,

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Independence

The events E and F are independent if and only if $P(E \cap F) = P(E)\,P(F)$

Example:

- A – Roll a 6

  B – Roll an even

- $P(A) = 1 / 6$ and $P(B) = 3/6 = 1/2$

- $P(A \cap B) = P(A)\,P(B) = 1/6 * 1 / 2 = 1 / 12$

Independence

- Suppose E is the event that a randomly generated bit string of length four begins with a 1 and F is the event that this bit string contains an even number of 1s.
  Are E and F independent if the 16 bit strings of length four are equally likely?

- **Solution**:
  There are eight bit strings of length four that begin with a 1, and eight bit strings of length four that contain an even number of 1s.

  - Since the number of bit strings of length 4 is 16, $P(E) = P(F) = 8/16 = ½$.

  - Since $E \cap F = \{1111, 1100, 1010, 1001\}$, $P(E \cap F) = 4/16 = 1/4$.

  We conclude that E and F are independent, because

  $$P(E \cap F) = 1/4 = (½)\ (½) = P(E)\ P(F)$$

Independence

- Assume (as in the previous example) that each of the four ways a family can have two children (BB, GG, BG,GB) is equally likely. Are the events E, that a family with two children has two boys, and F, that a family with two children has at least one boy, independent?

- **Solution**:

  Because $E = \{BB\}$, $P(E) = 1/4$. We saw previously that that $P(F) = 3/4$ and $P(E \cap F) = 1/4$.

  The events E and F are not independent since

$$P(E)\,P(F) = 3/16 \neq 1/4 = P(E \cap F).$$

Pairwise and Mutual Independence

- The events E_1, E_2, …, E_n are *pairwise independent* if and only if

  $p(E_i \cap E_j) = p(E_i)\, p(E_j)$ for all pairs i and j with $i \leq j \leq n$.

- The events are *mutually independent* if

$$p(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_m}) = p(E_{i_1})p(E_{i_2}) \cdots p(E_{i_m})$$

  whenever i_j, $j = 1,2,\ldots, m$, are integers with

  $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ and $m \geq 2$.

CS1382 Discrete Computational Structures

Bayes' Theorem

Spring 2019

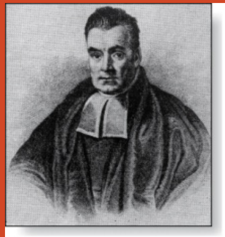
Richard Matovu

TEXAS TECH UNIVERSITY

Bayes' Theorem

Thomas Bayes

(1702-1761)



- Suppose that E and F are events from a sample space S such that P(E) ≠ 0 and P(F) ≠ 0. Then:

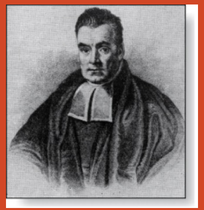
$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Comes from the Law of Total Probability

P(E)	=	P(F ∩ E) + P(F' ∩ E)
	=	P(F) P(E | F) + P(F ') P(E | F ')

Bayes’ Theorem

Thomas Bayes
(1702-1761)



Example:

We have two boxes. The first box contains two green balls and seven red balls. The second contains four green balls and three red balls. Bob selects one of the boxes at random. Then he selects a ball from that box at random. If he has a red ball, what is the probability that he selected a ball from the first box.

- Let *E* be the event that Bob has chosen a red ball
- *F* be the event that Bob has chosen the first box.
- By Bayes’ theorem the probability that Bob has picked the first box is:

$$p(F|E) = \frac{(7/9)(1/2)}{(7/9)(1/2) + (3/7)(1/2)} = \frac{7/18}{38/63} = \frac{49}{76} \approx 0.645.$$

Applying Bayes' Theorem

Example:

Suppose that one person in 100,000 has a particular disease. There is a test for the disease that gives a positive result 99% of the time when given to someone with the disease. When given to someone without the disease, 99.5% of the time it gives a negative result. Find

a) the probability that a person who test positive has the disease.

b) the probability that a person who test negative does not have the disease.

Should someone who tests positive be worried?

Applying Bayes' Theorem

Solution: Let D be the event that the person has the disease, and E be the event that this person tests positive. We need to compute P(D|E) from P(D), P(E|D), P(E | $\overline{D}$), P($\overline{D}$).

$$p(D) = 1/100,000 = 0.00001$$

$$p(E|D) = .99 \qquad p(\overline{E}|D) = .01$$

$$p(\overline{D}) = 1 - 0.00001 = 0.99999$$

$$p(E|\overline{D}) = .005 \qquad p(\overline{E}|\overline{D}) = .995$$

$$p(D|E) = \frac{p(E|D)p(D)}{p(E|D)p(D) + p(E|\overline{D})p(\overline{D})}$$

$$= \frac{(0.99)(0.00001)}{(0.99)(0.00001) + (0.005)(0.99999)}$$

$$\approx 0.002$$

So, don't worry too much, if your test for this disease comes back positive.

Applying Bayes' Theorem

- What if the result is negative?

$$p(\overline{D}|\overline{E}) = \frac{p(\overline{E}|\overline{D})p(\overline{D})}{p(\overline{E}|\overline{D})p(\overline{D}) + p(\overline{E}|D)p(D)}$$

So, the probability you have the disease if you test negative is

$$= \frac{(0.995)(0.99999)}{(0.995)(0.99999) + (0.01)(0.00001)}$$

$$\approx 0.9999999$$

- So, it is extremely unlikely you have the disease if you test negative.

Exercise

- In a group of 100 students, 40 are taking algebra, 30 are taking biology and 20 are taking both algebra and biology. If a student, chosen at random, is taking algebra, what is the probability that he or she is taking biology?

Questions?

Thank You!