Discrete Probability

Chapter 7

Discrete probability

Section 7.1
Finite probability

In a lottery, players win a large prize when they pick four digits that match, in the correct order, four digits selected by a random mechanical process. A smaller prize is won if only three digits are matched. What is the probability that a player wins the large prize? What is the probability that a player wins the smaller prize?

Solution 1:
- Only one way to choose all four digits correctly.
- By the product rule, there are $10 \times 10 \times 10 \times 10 = 10,000$ ways to choose four digits.
- The probability that a player wins the large prize is $1/10000 = 0.0001$.

Solution 2:
- By sum rule, $9 + 9 + 9 + 9 = 36$ ways to choose only three digits correctly.
- By the product rule, there are $10 \times 10 \times 10 \times 10 = 10,000$ ways to choose four digits.
- The probability that a player wins the large prize is $36/10000 = 0.0036$.

Probabilities of complements and unions of events

Let $E$ be an event in a sample space $S$. The probability of the event $\bar{E} = S \cdot E$, the complementary event of $E$, is given by

$$P(\bar{E}) = 1 - P(E)$$

Example:
- A sequence of 10 bits is randomly generated. What is the probability that at least one of these bits is 0?
- At least one bit is 0 = bits of length 10 - none of the 10 bits is 0
- $P = 1 - 1/2^{10} = 1023/1024$
Probabilities of complements and unions of events

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

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

Example:

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?

$E_1$: the integer selected at random is divisible by 2

$E_2$: the integer selected at random is divisible by 5

$$P(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) = \frac{50}{100} + \frac{20}{100} - \frac{10}{100} = \frac{3}{5}$$
Assigning Probabilities

- When there are \( n \) possible outcomes, \( x_1, x_2, \ldots, x_n \),
  1. \( 0 \leq p(x_i) \leq 1 \) \( i = 1, 2, \ldots, n \)
  2. \( \sum_{i=1}^{n} p(x_i) = 1 \)

What probabilities should we assign to these outcomes when the coin is biased so that heads comes up twice as often as tails?

- \( p(\text{H}) = 2p(\text{T}) \)
- \( p(\text{H}) + p(\text{T}) = 1 \)
- \( p(\text{H}) = \frac{2}{3}, \ p(\text{T}) = \frac{1}{3} \)

Assigning Probabilities

- Definition 2: The probability of the event \( E \) is the sum of the probabilities of the outcomes in \( E \). That is,
  \[ p(E) = \sum_{s \in E} p(s) \]

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:

- \( E = \{1, 3, 5\} \)
- \( p(1) = p(2) = p(4) = p(5) = p(6) = \frac{1}{2}p(3) \)
- \( p(1) + p(2) + p(3) + p(4) + p(5) + p(6) = 1 \)
- \( p(1) = p(2) = p(4) = p(5) = p(6) = \frac{1}{7}, \ p(3) = \frac{2}{7} \)
- \( p(E) = p(1) + p(3) + p(5) = \frac{4}{7} \)
Conditional Probability

- Definition 3: Let \( E \) and \( F \) be events with \( p(F) > 0 \). The conditional probability of \( E \) given \( F \), denoted by \( p(E \mid F) \), is defined as:
  \[
p(E \mid F) = \frac{p(E \cap F)}{p(F)}
  \]

- Example: A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0? (We assume that 0 bits and 1 bits are equally likely.)

  Solution:
  - \( E \): the event that a bit string of length four contains at least two consecutive 0s
  - \( F \): the event that the first bit of a bit string of length four is a 0.
  - The probability that a bit string of length four has at least two consecutive 0s, given that its first bit is a 0, equals \( p(E \mid F) = p(E \cap F)/p(F) \)
  - \( E \cap F = \{0000, 0001, 0010, 0011, 0100\} \), \( p(E \cap F) = 5/16 \).
  - Because there are eight bit strings of length four that start with a 0, \( p(F) = 8/16 = 1/2 \).
  - Consequently, \( p(E \mid F) = 5/16 / 1/2 = 5/8 \)

Independence

- The events \( E \) and \( F \) are independent \( p(E \cap F) = p(E) \cdot p(F) \)

- Example: Assume each of the four ways a family can have two children 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:
  - Two children : BB, BG, GB, GG
  - \( E = \{BB\} \), then \( p(E) = 1/4 \).
  - \( F = \{BB, BG, GB\} \), then \( p(F) = 3/4 \)
  - \( E \cap F = \{BB\} \), then \( p(E \cap F) = 1/4 \). But \( p(E)p(F) = 1/4 \cdot 3/4 = 3/16 \).
  - Therefore \( p(E \cap F) \neq p(E)p(F) \).
  - so the events \( E \) and \( F \) are not independent.
Bayes’ Theorem

Section 7.3

Suppose that \( E \) and \( F \) are events from a sample space \( S \) such that \( p(E) \neq 0 \) and \( p(F) \neq 0 \). Then

\[
p(F \mid E) = \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) + p(E \mid F)p(F)}
\]

Note: \( F \) and \( \overline{F} \) are mutually exclusive and cover the entire sample space \( S \), that is \( F \cup \overline{F} = S \).
Example

- Suppose one person in 100,000 has a particular rare disease for which there is a fairly accurate diagnostic test. This test is correct 99% of the time when given to someone with the disease; it is correct 99.5% of the time when given to someone who does not have the disease.

- Only 0.2% of people who test positive actually have the disease.
- \( P(\text{someone who tests negative for the disease does not have the disease}) \)?

F: a person has the disease
E: a person tests positive

\[
p(F \mid E) = \frac{p(E \mid F) p(F)}{p(E \mid F) p(F) + p(E \mid F) p(F)}
= \frac{0.99 \cdot 0.0001}{0.99 \cdot 0.0001 + 0.005 \cdot 0.99999} \approx 0.002
\]

Example

- Suppose one person in 100,000 has a particular rare disease for which there is a fairly accurate diagnostic test. This test is correct 99% of the time when given to someone with the disease; it is correct 99.5% of the time when given to someone who does not have the disease.

- 99.99999% of people who test negative really do not have the disease.
- F: a person has the disease
- E: a person tests positive

\[
p(\overline{F} \mid E) = \frac{p(\overline{E} \mid \overline{F}) p(\overline{F})}{p(\overline{E} \mid \overline{F}) p(\overline{F}) + p(\overline{E} \mid F) p(F)}
= \frac{0.995 \cdot 0.99999}{0.995 \cdot 0.99999 + 0.01 \cdot 0.00001} \approx 0.9999999
Expected value

- The expected value of a random variable is the sum over all elements in a sample space of the product of the probability of the element and the value of the random variable at this element.

- The weighted average of the values of a random variable

\[
E(X) = \sum_{s \in S} p(s)X(s)
\]

Example
Let \( X \) be the number that comes up when a die is rolled. What is the expected value of \( X \)?

\[
E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{7}{2}
\]
Linearity of Expectations

• If $X_i$ ($i = 1, 2, \ldots, n$) are random variables on $S$, and if $a$ and $b$ are real numbers, then
  
  i. $E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$
  
  ii. $E(aX+b)=aE(X)+b$

Example

Example: Find the expected value of the sum of the numbers that appear when a pair of fair dice is rolled.

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{1}{18} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{9} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{1}{6} + 8 \cdot \frac{1}{9} + 9 \cdot \frac{1}{10} + 10 \cdot \frac{1}{12} + 11 \cdot \frac{1}{18} + 12 \cdot \frac{1}{36} = 7.$$
Variance and Standard Deviation

- How widely the values of a random variable are distributed?
- Let $X$ be a random variable on a sample space $S$. The variance of $X$, denoted by $V(X)$, is
  \[ V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s). \]
- The standard deviation of $X$, denoted by $\delta(X)$, is defined to be $\sqrt{V(X)}$.
- If $X$ is a random variable on a sample space $S$, then $V(X) = E(X^2) - E(X)^2$.
- If $X$ and $Y$ are two independent random variables on a sample space $S$, then $V(X+Y) = V(X) + V(Y)$. Furthermore, if $X_i$ ($i=1, 2, \ldots, n$) are pairwise independent random variables on $S$, then $V(X_1+X_2+\ldots+X_n) = V(X_1)+V(X_2)+\ldots+V(X_n)$.

Example

- The variance of the random variable $X$ whose value is the number of heads when tossing a coin twice.
- Solution:
  - a coin is tossed twice, the number of heads is: 0 with probability 0.25, 1 with probability 0.5 and 2 with probability 0.25.
  - $X$: the number of heads
  - $E(X) = 0.25 \times 0 + 0.5 \times 1 + 0.25 \times 2 = 1$,
  - variance is $V(X) = 0.25 \times (0 - 1)^2 + 0.5 \times (1 - 1)^2 + 0.25 \times (2 - 1)^2 = 0.25 + 0 + 0.25 = 0.5$. 