CHAPTER 3

PROBABILITY

Probability

Probability is used to evaluate the reliability and validity of our inference based on the sample in reference to the population (which is usually unknown).

Experiment is a process by which measurements are observed. Example:

- Flipping a coin
- Tossing a dice

Event

Simple event is the outcome of one time repeating an experiment. Example:

- Coin —$E_1 = \{H\}$, $E_2 = \{T\}$
- Dice — $E_1 = \{1\}$, $E_2 = \{2\}$, $E_3 = \{3\}$, $E_4 = \{4\}$, $E_5 = \{5\}$, $E_6 = \{6\}$

An event is any collection of simple events. Example:

- Tossing a dice and observing whether its even or odd
    - E = {2, 4, 6}
    - O = {1, 3, 5}

We say an event A happens if at least one its elements is observed.

Interpretation of Probability as Long Run Relative Frequency

$p(A) = \lim_{n \to \infty} \frac{f_A}{n}$

$0 \le p(A) \le 1$

Properties of Probability Function

1. For any event A, $0 \le p(A) \le 1$
2. The sum of probabilities of all simple events is 1

Compute Probability of an Event

The probability of event A as the sum of the probabilities of the simple event consisting A.

Example:

- In experiment of tossing a fair dice, what is the probability of observing an even number.

Sample Space S

The set of all possible outcomes of an experiment (the set of all simple events). Example:

- Tossing a dice, S = {1, 2, 3, 4, 5, 6}
- Flipping a coin, S = {H, T}

$p(S) = 1$

Some Elementary Probability Rules

Complement of an event A: $\overline{A}$ consists of all elements which are in S but not in A. Example:

- Rolling a dice
  - E = {2, 4, 6}, S = {1, 2, 3, 4, 5, 6}, $\overline{E}$ = {1, 3, 5}

$p(\overline{A}) = p(S) - p(A)$

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

Operators on the Events

A u B (A union B):

- A = {1, 2, 3}, B = {2, 4, 6}
  A u B = {1, 2, 3, 4, 6}

AnB (A intersect B):

- A = {1, 2, 3}, B = {2, 4, 6}
  A n B = {2}

A set is said to be mutually exclusive if A n B = 0 → Empty set

The Addition Rule

$p(A\ u\ B) = p(A) + p(B) - p(A\ n\ B)$

In particular, if the events A & B are M.E., then:

- $p(A\ u\ B) = p(A) + p(B)$

Conditional Probability

$p(A|B) = \frac{p(A\ n\ B)}{p(B)},\ p(B) \neq 0 \ \rightarrow\ p(A\ n\ B) = p(A|B) \cdot p(B)$

$p(B|A) = \frac{p(A\ n\ B)}{p(A)},\ p(A) \neq 0 \ \rightarrow\ p(A\ n\ B) = p(B|A) \cdot p(A)$

Note:

- $p(A|B) \neq p(B|A)$
- $A\ n\ B = B\ n\ A$

Example:

- Flip a coin and a fair dice at the same time
  - S = {1H, 2H, 3H, 4H, 5H, 6H, 1T, 2T, 3T, 4T, 5T, 6T}
  - Let A be the event that the dice show a number less than 3
    Let B be the event that the coin shows H

    A = {1H, 2H, 1T, 2T}
    B = {1H, 2H, 3H, 4H, 5H, 6H}

    $A \; n \; B = \{1H, \; 2H\}$
    $p(A \; n \; B) = 2/12$

A and B are independent events if:

- $p(A|B) = p(A) \rightarrow \; p(A \; n \; B) = p(A)p(B)$
- $p(B|A) = p(B) \rightarrow \; p(B \; n \; A) = p(A)p(B)$

<u>Factorials</u>

$n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 1 = n!$

Example:

- $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5!$
- 5 people, and I want to choose a team of 3 people:
  - $\frac{5 \cdot 4 \cdot 3}{3!} = \frac{5 \cdot 4 \cdot 3 \cdot 2!}{3! \cdot 2!} = \frac{5!}{3! \cdot 2!} = \binom{5}{3} \rightarrow$ combination 3 out of 5

In general $\binom{n}{r} = \frac{n!}{r! \, (n-r)!}, \; r \leq n$

<u>Properties of Sample Space</u>

We say the events $S_1, \; ..., \; S_K$ is a partition of the sample space S if:

I. $S_1 \bigcup S_2 \bigcup ... \bigcup S_K = S$

II. $S_i \bigcap S_j = 0, \; i.e. \; S_i \; and \; S_j \; are \; M.E.$

<u>The Bayes' Theorem</u>

Let $S_1, \; S_2, \; ..., \; S_K$ be a partition of the sample space $p(S_1), \; p(S_2), \; ..., \; p(S_K)$. Then the posterior probability of event S_j given A is:

- $p(S_j| \, A) = \frac{p(A| \, S_j) \, p(S_j)}{\sum_{i=1}^{k} p(A| \, S_j) \, p(S_j)}$

Example:

- 60% of carleton students are male (i.e. 40% are female)
- 30% of the male students are smokers and 25% of the female students are smokers.
    - If a chosen student is a smoker. Then what is probability that the chosen student is male:
        - $p(M) = 0.6$
          $p(F) = 0.4$
          $p(Sm|M) = 0.3$
          $p(Sm|F) = 0.25$
        - $p(M|Sm) = \frac{p(Sm|M)\,p(M)}{p(Sm|M)\,p(M) + p(Sm|F)\,p(F)} = \frac{0.3 \cdot 0.6}{(0.3 \cdot 0.6) + (0.25 \cdot 0.4)} \approx 0.642$