Probability summary

Week 1

Introduction

Random experiment: An experiment that can result in different outcomes, even if repeated in the same manner.

Sample space: Set of all possible outcomes of a random experiment. Denoted as S.

- Example: $S = 0 \cup [2.5, 4.0]$
- Can be *discrete* or *continuous*.

Event: Subset of the sample space of a random experiment.

- $E_1 = \{\text{\# of students} \geq 100\}$
- You can do union, intersection and other set operations with them.
  - They act as sets, can use distributivity rule and De Morgan's laws.
- **Mutually exclusive events:** $E_1 \cap E_2 = \emptyset$
- **Venn Diagrams** are useful for visualization

Probability

Number of permutations of n different elements is $n!$.

Number of permutations of subsets of r elements selected from a set of n elements is $P_r^n = \frac{n!}{(n-r)!}$. Since there are n possibilities for the first element, $n-1$ for the second element and so forth, up to the $(n-r)^{\text{th}}$ element.

Number of subsets of r elements out of a set of n is called the number of combinations, given by:
$C_r^n = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

Sampling

Sampling with replacements: return element back to whole set.
Sampling without replacement: remove element from the whole set.

Probability

Probability of event E, denoted by $P(E)$, expresses the likelihood or chance of the occurrence of event E.

Axioms of Probability

- $P(S) = 1$
- $0 \leq P(E) \leq 1$
- For two events, E_1 and E_2 such that $E_1 \cap E_2 = \emptyset$, then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Extra rules

- $P(A') = 1 - P(A)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional Probability

The probability of event B given A. Assume $P(A) > 0$. Conditional probability is given as:

$P(B|A) = \frac{P(B \cap A)}{P(A)}$, just check the Venn diagram.

Multiplication rule: $P(A \cap B) = P(B|A)P(A) = P(A|B)P(B)$

Total probability rule: $P(B) = P(B \cap A) + P(B \cap A') = P(B|A)P(A) + P(B|A')P(A')$

For multiple events: Let A be an event, and let $E_1, \ldots, E_k$ be k mutually exclusive events, i.e. such that $\bigcup_{i=1}^{k} E_k = S$ and $\forall i \neq j : E_i \cap E_j = \emptyset$, then we know that:

- $P(A) = P(A \cap E_1) + P(A \cap E_2) + \cdots + P(A \cap E_k)$ or
  $P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_k)P(E_k)$

Independence of Events: Two events are independent if any of the following equivalent statements are true:

- $P(A|B) = P(A)$, so the occurrence event B does not affect event A
- $P(B|A) = P(B)$
- $P(A \cap B) = P(A)P(B)$

Jointly independent: If $P(E_1 \cap E_2 \cap \cdots \cap E_k) = P(E_1)P(E_2) \cdots P(E_k)$

Bayes' rule

Bayes' rule we can measure what conditional probabilities are.

$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

For multiple events: Let A be an event, and let $E_1, \ldots, E_k$ be k mutually exclusive events, i.e. such that $\bigcup_{i=1}^k E_k = S$ and $\forall i \neq j : E_i \cap E_j = \emptyset$, then we know that:

$$P(E_1|A) = \frac{P(A|E_1)P(E_1)}{P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + \cdots + P(A|E_k)P(E_k)}$$

Rest

Random variable: A function that assigns a real number to each outcome in the sample space of a random experiment. Can be discrete or continuous.

- Capital letters are used for random variables, X
- Lower case letters are used for the observed value, x.

Week 2

Random variable: A random experiment whose outcome is a real number. E.g. the temperature of the room is X (experiment), however, once we measure the value, it'll become x (value).

Discrete variables

Probability Mass Function (p.m.f.): Given a discrete random variable X with possible values $x_1, x_2, \ldots$, the p.m.f. is $f : \{x_1, x_2, \ldots\} \to [0, 1]$ such that:

- $f(x_i) \geq 0$
- $\sum_{i=1}^{\infty} f(x_1) = 1$
- $P(X = x_i) = f(x_i)$

Basically, the p.m.f. a function is a description of the probabilities associated with each possible outcome of X.

> Example
>
> Let sample space S={Print, save, cancel} for possible requests in a GUI.
> Identify each request with a number (0, 1 and 2 respectively).
> Let X be the random variable for this experiment.
>
> Now we can construct the p.m.f. w.r.t. X: $P(X = 0) = 0.2$, $P(X = 1) = 0.5$ and $P(X = 2) = 0.3$.

Cumulative Distribution Function (c.m.f.): The c.m.f. of a random variable X is denoted by $F(x) : \mathbb{R} \to [0, 1]$ and is given by:

- $F(x) = P(X \leq x)$, where $x \in \mathbb{R}$
- This is extremely powerful and is properly defined for any random variable, unlike p.m.f.

More concretely, let X be a discrete random variable with p.m.f. given by f.
For any $x \in \mathbb{R}$ we have that:

- $F(X) = P(X \leq x) = \sum_{i : x_i \leq x} f(x_i)$

  $$\implies 0 \leq F(X) \leq 1$$

  $$\implies (x \leq y) \Rightarrow (F(x) \leq F(y))$$

Thus, it's non-decreasing.

Also, note that $P(a < X \leq b) = F(b) - F(a)$.

Why is this useful? We are often interested in probabilities that is less or equal than a certain number, for example, that the number of customers in my shop is ≤ 50.

Mean (expected value) and Variance

There are certain "summaries" of the distribution of a random variable that can give a lot of information about it.

Let X be a discrete random variable taking values in $\{x_1, x_2, \ldots\} \in \mathbb{R}$.

The **mean** or **expected value** of X is denoted by μ_X or $\mathbb{E}(x)$ and is defined as:

- $\mathbb{E}(X) = \sum_{x \in \{x_1, x_2, \ldots\}} x f(x)$
- The weighted average of the possible values of X. The "center" of the distribution.

The **variance** of X is denoted by σ_X^2 or $V(X)$ and is defined as:

- $\sigma^2(X) = V(X) = \sum_{x \in \{x_1, x_2, \ldots\}} (x - \mu_X)^2 f(x)$
- Alternatively: $\sigma^2(x) = (\sum_{x_1, x_2, \ldots} x^2 f(x)) - \mu_X^2$ [Easier by hand, harder numerically]
- The dispersion of X around the mean. If the variance is large, then X varies a lot.
- Is always non-negative $V(X) \geq 0$.

Standard deviation: $\sqrt{V(X)}$

Function of Random Variables

Functions of random variables are *also* random variables!

Let X be a random variable, and let $h : \mathbb{R} \to \mathbb{R}$ be an arbitrary function.
Then $Y = h(X)$ is also a random variable.